Uniform Medical Plan coverage limits

Updates effective 09/1/2022

The benefit coverage limits listed below apply to these UMP plans:

- Uniform Medical Plan (UMP) Classic (PEBB)
- UMP Select (PEBB)
- UMP Consumer-Directed Health Plan (UMP CDHP) (PEBB)
- UMP Plus–Puget Sound High Value Network (UMP Plus–PSHVN) (PEBB)
- UMP Plus–UW Medicine Accountable Care Network (UMP Plus–UW Medicine ACN) (PEBB)
- UMP Achieve 1 (SEBB)
- UMP Achieve 2 (SEBB)
- UMP High Deductible Plan (SEBB)
- UMP Plus–Puget Sound High Value Network (UMP Plus–PSHVN) (SEBB)
- UMP Plus–UW Medicine Accountable Care Network (UMP Plus–UW Medicine ACN) (SEBB)

Some services listed under these benefits have coverage limits. These limits are either determined by a Health Technology Clinical Committee (HTCC) decision or a Regence BlueShield medical policy. **The table below does not include every limit or exclusion under this benefit. For more details, refer to your plan’s Certificate of Coverage.**

Uniform Medical Plan Pre-authorization List

The Uniform Medical Plan (UMP) Pre-authorization List includes services and supplies that require pre-authorization or notification for UMP members.
**Laboratory**

<table>
<thead>
<tr>
<th>These services or supplies have coverage limits</th>
<th>The rules or policies that define the coverage limits</th>
<th>Limit applies to these codes (chosen by your provider to bill for services)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement of Serum Antibodies to Selected Biologic Agents</td>
<td>Regence Medical Policy Lab65</td>
<td>• 80145, 80230, 80280</td>
</tr>
</tbody>
</table>

**Maternity**

Elective early delivery, prior to 39 weeks' gestation is not a covered benefit (not applicable to emergency delivery or spontaneous labor).

**Medicine**

<table>
<thead>
<tr>
<th>These services or supplies have coverage limits</th>
<th>The rules or policies that define the coverage limits</th>
<th>Limit applies to these codes (chosen by your provider to bill for services)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioengineered Skin and Soft Tissue Substitutes and Amniotic Products</td>
<td>Regence Medical Policy Med170</td>
<td>• A4100, A6460, A6461, C1849, Q4100, Q4101, Q4102, Q4105, Q4106, Q4107, Q4114, Q4116, Q4122, Q4128, Q4132, Q4133, Q4151, Q4154, Q4159, Q4186, Q4187</td>
</tr>
<tr>
<td>Confocal Laser Endomicroscopy</td>
<td>Regence Medical Policy Med151</td>
<td>• 43206, 43252, 88375</td>
</tr>
<tr>
<td>Coverage of Treatments Provided in a Clinical Trial</td>
<td>Regence Medical Policy Med150</td>
<td>• S9990, S9991, S9988</td>
</tr>
<tr>
<td>Digital Health Products</td>
<td>Regence Medical Policy Med175</td>
<td>• 0702T, 0703T, A9291</td>
</tr>
<tr>
<td>Digital Health Products for Attention Deficit Hyperactivity Disorder</td>
<td>Regence Medical Policy Med175.01</td>
<td>• 0702T, 0703T, A9291</td>
</tr>
<tr>
<td>Digital Health Products for Substance Use Disorders</td>
<td>Regence Medical Policy Med175.02</td>
<td>• 0702T, 0703T, A9291</td>
</tr>
</tbody>
</table>

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
<table>
<thead>
<tr>
<th>Service Description</th>
<th>Regence Medical Policy</th>
<th>Criteria</th>
</tr>
</thead>
</table>
| Hyperbaric Oxygen Therapy for Tissue Damage, Including Wound Care and Treatment of Central Nervous System Conditions | **HTCC decision**     | • 99183, G0277
*Regence Medical Policy Med14* is used only to determine units of treatment, criteria for diabetic "standard wound therapy" and to address any conditions not addressed in the HTCC decisions under the HTCC "limitations of coverage" or "non-covered indicators". |
| In Vivo Analysis of Colorectal Polyps                                           | **Regence Medical Policy Med104** | • 88375                                                                                                                                   |
| Intensity Modulated Radiotherapy (IMRT)                                           | **HTCC decision**     | • 77301, 77338, 77385, 77386, G6015, G6016                                                                                                   |
| Laser Interstitial Thermal Therapy                                                | **Regence Medical Policy Med177** | • 61736, 61737                                                                                                                                  |
| Orthopedic Applications of Stem-Cell Therapy, Including Bone Substitutes Used with Autologous Bone Marrow | **Regence Medical Policy Med142** | • 38206, 38232, 38241                                                                                                                            |
| Progenitor Cell Therapy for the Treatment of Damaged Myocardium Due to Ischemia   | **Regence Medical Policy Med100** | • 38205, 38206, 38240, 38241                                                                                                                   |
| Charged-Particle (Proton or Helium Ion) Radiotherapy                              | **HTCC decision**     | • 77520, 77522, 77523, 7525
  o Pre-authorization is not required for members under 21 years of age

When the following codes are used for Charged-Particle (Proton or Helium Ion) Radiotherapy with SRS or SBRT, use [HCTT Decision (PDF)] criteria: 32701, 61796, 61797, 61798, 61799, 61800, 63620, 63621, 77301, 77338, 77371, 77372, 77373, 77432, 77435, G0339, G0340 |
| Radioembolization, Transarterial Embolization (TAE) and Transarterial Chemoembolization (TACE) | **Regence Medical Policy Med140** | • 37243, 79445, S2095
*Note: Regence Medical Policy Ovarian and Internal Iliac Vein Embolization as a Treatment of Pelvic Congestion Syndrom is considered investigational.* |
| **Tinnitus: Non-invasive, non-pharmacologic treatments** | **HTCC decision** | • UMP is subject to [HTCC Decision](https://example.com/HTCCDecision) (PDF) for codes 0552T, 90832, 90833, 90834, 90836, 90837, 90838, 90867, 90868, 90869, 96156, 96158, 96159, 96160, 96161, 96164, 96165, 96167, 96168, 96169, 96170, 96171, S8948
• Codes 0552T and S8948, when billed without a tinnitus diagnosis will be denied as investigational based on Regence Medical Policy Low Level Laser Therapy
• Note: Codes 90867 and 90868, when billed with chronic migraine and chronic tension headaches, is not a covered benefit per [HTCC Decision](https://example.com/HTCCDecision)

| **Gender Affirming Interventions for Gender Dysphoria** | **Regence Medical Policy Med153** | • 15775, 15776, 17380, 55970, 55980 Codes 55970 and 55980 are non-specific. The specific procedure code(s) must be requested in place of these non-specific codes. 17999, 19303, 19316, 19318, 19325, 19350, 53400, 53405, 53410, 53415, 53420, 53425, 53430, 54125, 54400, 54401, 54405, 54520, 54660, 54690, 55175, 55180,

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Measurement of Serum Antibodies to Selected Biologic Agents

Effective: June 1, 2022

Next Review: April 2023
Last Review: April 2022

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Anti-drug antibodies to drugs such as infliximab, adalimumab, ustekinumab, and vedolizumab may be found in patients undergoing treatment for inflammatory diseases including inflammatory bowel disease, psoriasis, ankylosing spondylitis, or rheumatoid arthritis and are thought to be associated with a loss of treatment response.

MEDICAL POLICY CRITERIA

I. Measurement of serum antibodies to infliximab (Remicade, Inflectra, Renflexis) or adalimumab (Humira), either alone or as a combination test that includes serum drug levels, may be considered medically necessary for patients with inflammatory bowel disease (i.e., Crohn’s disease or ulcerative colitis), when there is documentation of a loss of response to one of these medications.

II. Measurement of serum antibodies to infliximab (Remicade, Inflectra, Renflexis) or adalimumab (Humira), either alone or as a combination test that includes serum drug levels, is considered not medically necessary when there has not been a loss of response to the medication.

III. Measurement of serum antidrug antibodies, either alone or as a combination test that includes serum drug levels, is considered investigational for all of the following:
A. For any chronic inflammatory condition other than inflammatory bowel disease (i.e., Crohn’s disease or ulcerative colitis), including but not limited to rheumatoid arthritis and psoriasis, and
B. For quantification of antibodies to ustekinumab, vedolizumab, certolizumab, etanercept, or golimumab for any condition.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Medication Policy Manual, Note: Do a find (Ctrl+F) and enter drug name in the find bar to locate the appropriate policy.

BACKGROUND

INFLIXIMAB, ADALIMUMAB, USTEKINUMAB, AND VEDOLIZUMAB IN AUTOIMMUNE DISEASE

Therapy with monoclonal antibodies has revolutionized treatment of patients with inflammatory diseases such as inflammatory bowel disease (IBD; Crohn's disease [CD] and ulcerative colitis [UC]), rheumatoid arthritis and psoriasis. These agents are generally given to patients who fail conventional medical therapy, and they are typically highly effective for induction and maintenance of clinical remission. However, not all patients respond, and a high proportion of patients lose response over time. An estimated one-third of patients do not respond to induction therapy (primary nonresponse), and among initial responders, response wanes over time in approximately 20% to 60% of patients (secondary nonresponse). The reasons for therapeutic failures remain a matter of debate but include accelerated drug clearance (pharmacokinetics) and neutralizing agent activity (pharmacodynamics) due to anti-drug antibodies (ADA).[1]

Infliximab (Remicade® by Janssen Biotech, Inflectra® by Pfizer, and Renflexis® by Merck Sharp & Dohme) is an intravenous tumor necrosis factor alpha (TNFα) blocking agent approved by the U.S. Food and Drug Administration (FDA) for the treatment of rheumatoid arthritis, CD, ankylosing spondylitis, psoriatic arthritis, plaque psoriasis, and ulcerative colitis (UC). Infliximab is a chimeric (mouse/human) anti-TNFα monoclonal antibody. Adalimumab (Humira® AbbVie) is a subcutaneous TNFα inhibitor that is FDA-approved for treatment of the above indications (CD and UC in adults only) plus juvenile idiopathic arthritis (JIA). Adalimumab is a fully human monoclonal antibody to TNFα. Certolizumab (Cimzia® by UCB) is a subcutaneous TNFα inhibitor that is FDA-approved for treatment of rheumatoid arthritis, CD, ankylosing spondylitis, psoriatic arthritis, plaque psoriasis, and non-radiographic axial spondyloarthritis (nr-axSpA). Etanercept (Enbrel®, Immunex) is a TNFα inhibitor that is FDA-approved for the treatment of rheumatoid arthritis, JIA, ankylosing spondylitis, psoriatic arthritis, and plaque psoriasis. Golimumab (Simponi® by Janssen Biotech) is a subcutaneous TNFα inhibitor that is FDA-approved for the treatment of rheumatoid arthritis, ankylosing spondylitis, UC, and psoriatic arthritis. Vedolizumab (Entyvio®, Millennium Pharmaceuticals) is an intravenous blocking agent for integrin α4β7 and is FDA-approved for adults with CD or UC. Ustekinumab (Stelara®, Janssen Biotech) is an antibody that blocks interleukins IL-12 and IL-23 and is FDA-approved to treat psoriasis and certain patients with Crohn’s disease.
Following primary response to these medications, some patients become nonresponders (secondary nonresponse). The development of anti-drug antibodies (ADA) is thought to be a cause of secondary nonresponse. ADA are also associated with injection site reactions (adalimumab), and acute infusion reactions and delayed hypersensitivity reactions (infliximab). As a fully human antibody, adalimumab is considered less immunogenic than chimeric antibodies, such as infliximab.

DETECTION OF ANTI-DRUG ANTIBODIES

The detection and quantitative measurement of ADA has been fraught with difficulty owing to drug interference and identifying when antibodies are likely to have a neutralizing effect. First-generation assays, (i.e., enzyme-linked immunosorbent assays [ELISA]) can measure only ADA in the absence of detectable drug levels, due to interference of the drug with the assay. Other techniques available for measuring antibodies include the radioimmunoassay (RIA) method, and more recently, the homogenous mobility shift assay (HMSA) using high-performance liquid chromatography. Disadvantages of the RIA method are associated with the complexity of the test and prolonged incubation time, and safety concerns related to the handling of radioactive material. The HMSA has the advantage of being able to measure ADA when infliximab is present in the serum. A reporter-gene assay (RGA) is also available, which allows for the measurement of ADAs capable of neutralizing drug activity.[2] Cell-based assays typically have difficulty in standardization, take up to two days to complete, and with effects from the serum matrix. However, the RGA can quantify the anti-drug neutralizing antibody independent of matrix effects within two hours. Application of the RGA has recently been assessed for use in a clinical laboratory setting, and found to be a precise and high-throughput robust platform for detection of ADA.[3] Large randomized studies are still necessary to establish relevant clinical cut-off levels. Studies evaluating the validation of results among different assays are lacking, making inter-study comparisons difficult. One retrospective study in 63 patients demonstrated comparable diagnostic accuracy between two different ELISA methods in patients with IBD (i.e., double antigen ELISA and antihuman lambda chain-based ELISA).[4] This study did not include an objective clinical and endoscopic scoring system for validation of results. A 2013 review by Seow and Panaccione, noted that the variability and lack of standardization in current assay tests has important implications for subsequent studies which report associations between antibodies-to-infliximab (ATIs) and infliximab levels and utilize these assays to predict treatment response.[5] These findings highlight the need for a validated gold standard test and established diagnostic parameters with which to measure levels of infliximab and ATIs.

TREATMENT OPTIONS FOR PATIENTS WITH SECONDARY LOSS OF RESPONSE TO ANTI-TNF THERAPY

A diminished or suboptimal response to infliximab or adalimumab can be managed in several ways: shortening the interval between doses, increasing the dose, switching to a different anti-TNF agent (in patients who continue to have loss of response after receiving the increased dose), or switching to a non-anti-TNF agent.

REGULATORY STATUS

Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests (LDTs) must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments (CLIA). Laboratories that offer LDTs must be...
licensed by the CLIA for high complexity testing. To date, the U.S. Food and Drug Administration has chosen not to require regulatory review of these tests.

Prometheus® Laboratories Inc., a College of American Pathologists–accredited lab under CLIA, offers non-radiolabeled fluid-phase HMSA tests called the Anser® IFX test for infliximab. Anser® ADA for adalimumab, Anser® UST for ustekinumab, and Anser® VDZ for vedolizumab. None of these tests are ELISA-based and they can measure anti-drug antibodies in the presence of detectable drug levels, improving upon a major limitation of the ELISA method. All tests measure serum concentrations and anti-drug antibodies.

LabCorp has a portfolio of tests called DoseASSURE™ including DoseASSURE™ ADL for adalimumab, DoseASSURE™ UST for ustekinumab, DoseASSURE™ IFX for infliximab, DoseASSURE™ CTZ for certolizumab, DoseASSURE™ ETN for etanercept, and DoseASSURE™ GOL for golimumab. These tests are electrochemiluminescence immunoassay (ECLIA) and/or ELISA-based and report drug concentration and anti-drug antibody levels.

EVIDENCE SUMMARY

Validation of the clinical use of any diagnostic test focuses on analytic validity, diagnostic validity, and clinical utility. Analytic validity demonstrates technical feasibility as compared to a gold standard, including assessment of test reproducibility and precision. For comparison among studies, a common standardized protocol is necessary. Diagnostic utility is evaluated by the ability of a test to accurately predict the clinical outcome in appropriate populations of patients. For accurate interpretation of study results, sensitivities, specificities, and positive and negative predictive values compared to a gold standard must be known. Clinical utility is established when the evidence demonstrates that the diagnostic information obtained from a test can be used to benefit patient management and improve health outcomes. This evidence review focuses on the clinical validity and clinical utility.

Most studies evaluating antibodies to infliximab, adalimumab, ustekinumab, or vedolizumab report serum drug levels together with anti-drug antibody (ADA) levels, and correlate levels to disease response. Serum drug levels and disease response will not be addressed in this section and therefore the data reported on ADA will be highlighted from the identified studies. Most evidence concerning testing for ADA is derived from the data available for patients with inflammatory bowel disease (IBD) and rheumatoid arthritis (RA). Less literature exists concerning other diseases comprising psoriasis and spondyloarthropathies (SpA; i.e., ankylosing spondylitis, psoriatic arthritis, IBD-associated arthritis, reactive arthritis, and undifferentiated and juvenile SpA). There is also a lack of literature on the measurement of anti-vedolizumab and anti-ustekinumab antibodies for patient management.

CLINICAL VALIDITY

There is a substantial body of evidence examining associations of ADA with nonresponse and injection or infusion site reactions; numerous systematic reviews and meta-analyses have been published. Accordingly, the review of evidence concerning clinical validity focuses on the most current systematic reviews (see Tables 3 through 5) and studies published since those reviews,[6] as well as relevant studies not included in identified reviews (e.g., those focusing on adverse reactions and ADA).

Systematic Reviews
A systematic review (SR) published by Vermeire (2018) evaluated studies on immunogenicity to adalimumab (ADM), certolizumab pegol (CZP), golimumab, infliximab (IFX), ustekinumab, and vedolizumab in patients being treated for inflammatory bowel disease (IBD). Although 122 publications covering 114 studies were noted as included in the review, all study designs and abstracts from conference proceedings were included. Greater than 90% of studies involved administration of ADM or IFX. Of the studies involving IFX administration, only 12 were RCTs and 62 were non-randomized or observational studies. Across these studies, rates of ADA formation were highly variable, ranging from 0.0–65.3% in patients with IBD. While the authors reported that the proportion of patients achieving and maintaining a response to treatment with IFX was “generally lower” for patients with detected ADA than those without detected ADA, no pooled analyses were reported for any study outcomes. No analysis informing clinically useful thresholds or timing of antibody testing was provided. This review was funded by Pfizer, Inc, a manufacturer of Inflectra, which is an infliximab biosimilar and multiple study authors are employees and/or stakeholders in Pfizer, Inc.

Six SRs published from 2012 through 2017 were identified. The number of studies included ranged from 11 to 68, varying according to review objectives and conditions of interest. Although not detailed here, there was considerable overlap in included studies across reviews.

A SR with meta-analysis by Pecoraro (2017) selected 34 studies (total n=4,273 patients), including randomized controlled trials (RCTs, n=4), prospective observational (n=22), retrospective observational (n=6), and cross-sectional studies (n=2). Studies evaluated RA (n=18), ulcerative colitis (n=2), CD (n=5), psoriatic arthritis (n=4), ankylosing spondylitis (n=5), plaque psoriasis (n=4), spondyloarthritis (n=1). Most of the patients (45%) received infliximab, 35% received adalimumab, and 21% received etanercept. None received golimumab or certolizumab. Reviewers identified studies published through August 2016 and rated study quality as good (n=17), fair (n=16), and poor (n=1). The effect of ADA was evaluated in 19 studies, showing a significant (p<0.05) reduction of response (relative risk [RR] 0.43, 95% confidence interval [CI] 0.3 to 0.63) in ADA-positive patients relative to ADA-negative patients, with adalimumab therapy demonstrating a greater reduction (RR 0.40, 95% CI 0.25 to 0.65, p<0.001) than infliximab (RR 0.37, 95% CI 0.2 to 0.7, p<0.001). Measures of heterogeneity were 84%, 57%, and 79%, respectively. Fourteen studies reported on the effect of ADA on clinical response (see Table 1). Eleven studies found the risk of developing ADA to be significantly (p=0.03) lower in patients treated with concomitant methotrexate therapy relative to those without methotrexate (RR 0.65, 95% CI 0.47 to 0.9). Studies comparing treatment response with nonresponse (n=15) found responders to have a significantly (p<0.001) lower risk of developing ADA relative to nonresponders (RR 0.31, 95% CI 0.18 to 0.52). The presence of ADA was associated with a significant reduction of anti-tumor necrosis factor α (TNF-α) serum concentration (see Table 2). Of the 20 studies (n>2,800 patients) reporting data on adverse events, 31% (n=2 studies) developed infections, 18% (n=12 studies) developed injection-site reactions, 8% (n=11 studies) discontinued treatment due to adverse events, and 5% (n=1 study) developed serious adverse events (5%). Although ADA significantly reduced TNF-α response, the results should be viewed cautiously due to reported study limitations, including small numbers of studies included and considerable heterogeneity.

Freeman (2017) published a SR with meta-analysis evaluating the test accuracy estimates of levels of anti-tumour necrosis factor (anti-TNF) and antibodies to anti-TNF to predict loss of response or lack of regaining response in patients with anti-TNF managed Crohn’s disease (CD). Studies of patients with CD treated with infliximab or adalimumab as well as studies
with mixed Crohn’s and ulcerative colitis populations were included if the proportion of Crohn’s patients was at least 70%. Twenty-four full-test reports and seven conference abstracts were included in the SR; eleven of the 31 studies examined infliximab trough levels, 20 examined levels of antibodies to infliximab and five and six studies, respectively, investigated adalimumab levels and antibodies to adalimumab. The greatest identified threat to validity of the studies was high risk of bias in patient selection, which was present in nearly 80% of the included studies. The studies were heterogeneous with respect to the type of test used (e.g., commercial or in-house ELISA, radioimmunoassay (RIA), homogeneous mobility shift assay (HMSA)), criteria for establishing response or lack of regaining response (e.g., use of the Crohn’s Disease Activity Index score or the physician’s global assessment score) and population examined (responders or patients with secondary loss of response). Summary point estimates for sensitivity and specificity were 56% and 79% for antibodies to infliximab, respectively, and results for antibodies to adalimumab were similar. Positive and negative predictive values across all pooled studies ranged between 70% and 80%, implying that between 20% and 30% of both positive and negative test results may be incorrect in predicting loss of response. The authors concluded that “higher quality head-to-head test accuracy studies are required to enable differentiation between different types of tests and cut-offs, with consistent outcome measurement in the same population” and “more clinical trial evidence from test–treat studies is required before the clinical utility of the tests can be reliably evaluated.”

A SR and meta-analysis by Thomas (2015) included 68 studies (14,651 patients) with patients with RA (n=8,766), SpA (n=1,534), and IBD (n=4,351) and examined the immunogenicity of infliximab (39 comparisons), adalimumab (15), etanercept (5), golimumab (14), and certolizumab (8).[12] The review identified studies published through December 2013 and included 38 RCTs and 30 observational studies (study quality rated as good [n=32], moderate [n=26], or poor [n=10]). The pooled prevalence of ADA varied with disease and drug (see Table 3, highest with infliximab: 25.3%). Duration of exposure (reported in 60 studies) was examined for its potential effect on the development of ADA and most studies employed ELISA assays. The presence of ADA was associated with lower odds of response across most drugs and diseases (see Table 4). An exception was in studies of IBD (similar to that reported by Lee [2012]). The use of immunosuppressive agents substantially decreased the risk of ADA (odds ratio [OR] 0.26, 95% CI 0.21 to 0.32). Finally, infusion reactions and injection site reactions were more common (see Table 5) when ADA were detectable (OR 3.25, 95% CI 2.35 to 4.51). Evaluation of potential publication bias or overall assessment (e.g., GRADE or similar) for the body of evidence was not reported. Additionally, no measures of heterogeneity were reported.

A SR by Meroni (2015) included 57 studies of infliximab (n=34), adalimumab (n=18), and etanercept (n=5).[8] Studies included primarily patients with IBD and RA, but also SpA and psoriasis. Most studies were prospective cohort designs (n=42) and a formal assessment of study quality (bias) was not reported. The authors noted considerable variability in the time from drug administration to ADA and drug bioavailability testing across studies. Varied antibody testing assay methods were used and included solid-phases RIA, traditional ELISA, fluid-phase RIA, and bridging ELISA; cutoffs for positive test results were also inconsistently reported. The ranges of patients with detectable ADA varied substantially (see Table 3) but were consistent with other reviews. Qualitatively, the presence of ADA was associated with lower levels of infliximab and lower risk of disease control or remission. The presence of ADA also increased the risk of infusion reactions. When ascertained, the time to development of ADA varied from as little as 16 weeks to over a year. The time to ADA positivity varied – fifty percent of patients with detectable ADA at 28 weeks to a median time of one year. Finally, for
both infliximab and adalimumab, immunosuppression was associated with less ADA positivity. The authors concluded that “…the lack of homogeneity in study design and methodologies used in the studies analyzed limited the opportunity to establish the time-course and clinical consequences of anti-drug antibody development….” Although qualitative, the authors included many studies, and provided a detailed review of each study not reported by the other meta-analyses. The author’s conclusions are consistent with the meta-analyses but with emphasis on important aspects of heterogeneity across studies.

Hsu (2014) published a SR of ADA in psoriasis that included 25 studies (n=7,969). Inclusion criteria for the studies were: having at least 15 patients, documentation of serial assessments of psoriasis severity, and reporting ADA in patients with psoriasis receiving infliximab, etanercept, adalimumab, or ustekinumab. Ten of these studies reported on infliximab ADA: three found an association between ADA and lower serum infliximab levels, and five found an association between ADA and clinical response. Of the five studies that evaluated antiadalimumab antibodies, four found lower treatment efficacy for those with ADA. Six studies reported on ustekinumab ADA, and two of these found an association between ADA and Psoriasis Area and Severity Index (PASI) response. The remaining six studies in the review focused on anti-etanercept antibodies.

Nanda (2013) conducted a meta-analysis of studies that reported on clinical outcomes according to the presence or absence of ADA in patients with IBD. MEDLINE, Web of Science, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Scopus databases were searched to February 2012, EMBASE to August 2012; 11 studies involving 707 patients were included. Six of these studies (two RCTs, one prospective cohort study, three retrospective cohort studies) were included in the meta-analysis by Lee (2012) outlined below. In at least one quality domain (study eligibility criteria, measurement of exposure and outcome, control for confounders, completeness of follow-up), all the included studies had high risk of bias. The prevalence of detectable ADA in the included studies ranged from 22.4% to 46% (see Table 3). The outcome of interest was loss of response to infliximab, defined as “relapse of clinical symptoms in patients who were in clinical remission from, or had responded to, infliximab.” Measures of loss of response varied across studies and included clinician assessment, standardized scales (Crohn’s Disease Activity Index [CDAI], Harvey-Bradshaw Index, Simple Clinical Colitis Activity Index), and requirement for surgery or presence of nonhealing fistula. Patients with ATIs had a three-fold greater risk of loss of response than those without ATIs (RR 3.2, 95% CI 2.0 to 5.0) (shown in Table 3 as the RR of clinical response in treated vs. untreated patients to allow comparison with other meta-analyses). This result was influenced primarily by 532 patients with CD (RR 3.2, 95% CI 1.9 to 5.5); pooled results for 86 patients with ulcerative colitis (UC) were not statistically significant (pooled RR 2.2, 95% CI 0.5 to 9.0). Eighty-nine patients with unspecified IBD also were included in the meta-analysis. In addition to potential bias in included studies and heterogeneity in outcome assessment, the meta-analysis is limited by variability in the method of ADA detection (double-antigen ELISA, antihuman lambda chain-based ELISA, fluid-phase RIA). Study investigators stated, “[t]he true incidence of ADA in IBD patients treated with infliximab remains unknown due to the different administration schedules, timing of ADA measurements, methods used in ADA detection, and the presence of serum infliximab.” Finally, although the authors noted that the funnel plot “suggested the presence of publication bias,” the small number of studies and plot appearance (only two of 11 studies suggesting asymmetry) preclude conclusions.

Garces (2013) performed a meta-analysis of studies of infliximab and adalimumab used to treat RA, IBD, SpA, and psoriasis. Databases were searched to August 2012, and 12...
prospective cohort studies included involving 860 patients (540 with RA, 132 with SpA, 130 with IBD, 58 with psoriasis). The outcome of interest was response, assessed by using standard assessment scales for rheumatologic diseases (e.g., European League Against Rheumatism criteria for RA; Assessment in Ankylosing Spondylitis 20% response criteria, or ASDAS for spondyloarthritis; Psoriasis Area and Severity Index for psoriasis) and clinician assessment for IBD. Overall, detectable ADA were associated with a 68% reduction in drug response (pooled RR=0.32, 95% CI 0.22 to 0.48). Significant heterogeneity was introduced by varying use of immunosuppressant therapy (e.g., methotrexate) across studies. To assess ADA, most studies used RIA, which is less susceptible than ELISA to drug interference and may be more accurate.

Lee (2012) conducted a meta-analysis of patients with IBD receiving infliximab to estimate the prevalence of ADA, effect of ADA on the prevalence of infusion reactions, and the effect of ADA on disease remission rates.[10] Databases were searched through October 2011, and 18 studies involving 3,326 patients were included. Studies included nine RCTs, five prospective cohort studies, and four retrospective cohort studies. The prevalence of ADA was 45.8% when episodic infusions of infliximab were given and 12.4% when maintenance infliximab was given (see Table 3). Patients with ADA were less likely to be in clinical remission (Table 4), but this was not statistically significant (RR, 0.90, 95% CI 0.79 to 1.02, p=0.10). The rates of infusion reactions were significantly higher in patients with ADA (RR 2.07 [see Table 5], 95% CI 1.61 to 2.67). Immunosuppressants resulted in a 50% reduction in the risk of developing ADA (p<0.001). The meta-analysis concluded that patients with IBD who test positive for ATIs are at an increased risk of infusion reactions, but have similar rates of remission compared with patients who test negative for ATIs.

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>No. Studies</th>
<th>MD, mg/L</th>
<th>95% Confidence Interval</th>
<th>I², %</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease Activity Score 28</td>
<td>9</td>
<td>-7.07</td>
<td>-1.65 to -5.25</td>
<td>98</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BASDAI</td>
<td>2</td>
<td>-2.02</td>
<td>-1.01 to -1.03</td>
<td>0</td>
<td>0.17</td>
</tr>
<tr>
<td>ASDAS</td>
<td>2</td>
<td>-0.62</td>
<td>-1.09 to -0.14</td>
<td>0</td>
<td>0.13</td>
</tr>
<tr>
<td>Psoriasis Area Severity Index</td>
<td>1</td>
<td>4.7</td>
<td>-1.15 to 9.25</td>
<td>NR</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Adapted from Pecoraro (2017).[13]

ASDAS: Ankylosing Spondylitis Disease Activity Score; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; I²: heterogeneity measure; MD: mean difference; NR: not reported.

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>No. Studies</th>
<th>MD, mg/L</th>
<th>95% Confidence Interval</th>
<th>I², %</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADA-positive vs ADA-negative</td>
<td>8</td>
<td>-7.07</td>
<td>-1.65 to -5.25</td>
<td>98</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Responders vs no responders</td>
<td>13</td>
<td>2.77</td>
<td>1.97 to 3.58</td>
<td>82</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adalimumumab therapy</td>
<td>6</td>
<td>5.07</td>
<td>3.77 to 6.36</td>
<td>62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Infliximab</td>
<td>4</td>
<td>2.74</td>
<td>0.59 to 4.89</td>
<td>62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Etanercept</td>
<td>3</td>
<td>0.85</td>
<td>0.41 to 1.23</td>
<td>82</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DAS28 change from baseline</td>
<td>8</td>
<td>-2.18</td>
<td>-2.91 to -1.44</td>
<td>97</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Adapted from Pecoraro (2017).[13]

ADA: anti-drug antibodies; DAS28: Disease Activity Score in 28 joints; I²: heterogeneity measure; MD: mean difference; TNF: tumor necrosis factor.
## Table 3. Estimated Prevalence of Anti-drug Antibodies from Meta-Analyses

<table>
<thead>
<tr>
<th>Author</th>
<th>Included Studies</th>
<th>Drugs</th>
<th>Disease</th>
<th>Prevalence of ADA</th>
<th>Pooled (95% CI)</th>
<th>Range in Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee (2012)</td>
<td>18</td>
<td>IFX</td>
<td>IBD</td>
<td>20.8%</td>
<td>(19.2 to 22.5)</td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>5</td>
<td>ADL</td>
<td>RA</td>
<td>45.8%</td>
<td>(41.7 to 50.0)</td>
<td></td>
</tr>
<tr>
<td>Nanda (2013)</td>
<td>11</td>
<td>Other</td>
<td>SpA</td>
<td>12.4%</td>
<td>(10.8 to 14.1)</td>
<td></td>
</tr>
<tr>
<td>Thomas (2015)</td>
<td>39</td>
<td>IFX</td>
<td>IBD</td>
<td>25.3%</td>
<td>(19.5 to 32.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>ADL</td>
<td>RA</td>
<td>6.9%</td>
<td>(3.4 to 13.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Other</td>
<td>SpA</td>
<td>15.8%</td>
<td>(9.6 to 24.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>IFX</td>
<td>IBD</td>
<td>12.1%</td>
<td>(8.1 to 17.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>ADL</td>
<td>RA</td>
<td>8.9%</td>
<td>(3.8 to 19.2)</td>
<td></td>
</tr>
</tbody>
</table>

ADL: adalimumab; CI: confidence interval; IBD: inflammatory bowel disease; IFX: infliximab; RA: rheumatoid arthritis; SpA: spondyloarthropathy.

\(^a\) Includes etanercept, golimumab, certolizumab.

\(^b\) Includes three studies including both maintenance and episodic therapy.

\(^c\) Number of comparisons in table; did not report studies for pooled prevalence.

\(^d\) Also psoriasis.

## Table 4. Results from Meta-Analyses of Anti-drug Antibodies and Clinical Response

<table>
<thead>
<tr>
<th>Author</th>
<th>Included Studies</th>
<th>Drugs</th>
<th>Disease</th>
<th>Clinical Response: ADA vs None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee (2012)</td>
<td>18</td>
<td>IFX</td>
<td>IBD</td>
<td>0.90 (0.79 to 1.02)</td>
</tr>
<tr>
<td>Nanda (2013)</td>
<td>11</td>
<td>ADL</td>
<td>RA</td>
<td>0.33 (0.20 to 0.40)</td>
</tr>
<tr>
<td>Garces (2013)</td>
<td>12</td>
<td>Other</td>
<td>SpA</td>
<td>0.32 (0.22 to 0.48)</td>
</tr>
<tr>
<td>Thomas (2015)</td>
<td>14</td>
<td>IFX</td>
<td>IBD</td>
<td>1.16 (0.66 to 2.03)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>ADL</td>
<td>RA</td>
<td>0.27 (0.20 to 0.36)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Other</td>
<td>SpA</td>
<td>0.18 (0.09 to 0.37)</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>IFX</td>
<td>IBD</td>
<td>0.42 (0.30 to 0.58)</td>
</tr>
</tbody>
</table>

ADL: adalimumab; CI: confidence interval; IBD: inflammatory bowel disease; IFX: infliximab; NR: not reported; OR: odds ratio; RA: rheumatoid arthritis; RR: relative risk; SpA: spondyloarthropathy.

\(^a\) Includes etanercept, golimumab, certolizumab.

\(^b\) Also psoriasis.

## Table 5. Increased Risk of Adverse Reaction Associated With the Presence of Anti-drug Antibodies

<table>
<thead>
<tr>
<th>Author</th>
<th>Included Studies</th>
<th>Drugs</th>
<th>Disease</th>
<th>Adverse Reactions: ADA vs None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee (2012)</td>
<td>18</td>
<td>IFX</td>
<td>IBD</td>
<td>2.07 (1.61 to 2.67)</td>
</tr>
<tr>
<td>Thomas (2015)</td>
<td>NR</td>
<td>ADL</td>
<td>RA</td>
<td>3.25 (2.35 to 4.51)</td>
</tr>
</tbody>
</table>

ADL: adalimumab; CI: confidence interval; IBD: inflammatory bowel disease; IFX: infliximab; NR: not reported; OR: odds ratio; RA: rheumatoid arthritis; RR: relative risk; SpA: spondyloarthropathy.

\(^a\) Infusion reaction.

## Nonrandomized Studies

Recent publications not included in the SRs above are included, below.
A multicenter prospective cohort study of 137 patients with plaque-type psoriasis was published by De Keyser (2019). Serum samples and Psoriasis Area and Severity Index scores were obtained at baseline, week 16, 28, 40, 52, and/or ≥64 of ustekinumab treatment. Presence of anti-ustekinumab antibodies (prevalence of 8.7%) was significantly associated with a diminished clinical response (p=0.032). The median ustekinumab trough concentration was 0.3 mcg/mL (<0.02-3.80). No differences in serum concentrations were observed between moderate to good responders and nonresponders (p=0.948). Although the authors found that the presence of anti-ustekinumab antibodies was associated with treatment response in this patient population, serial measurements were collected in less than half (43.8%) of the patients. Anti-ustekinumab antibodies was reported to have developed during the first 52 weeks of treatment, however, the number of observations in the first year of treatment (n=191) was significantly higher than the number of observations in patients on treatment more than one year (n=38). This may underestimate the prevalence of anti-ustekinumab antibody formation after long-term treatments. Ultimately, the authors concluded that while measurement of anti-ustekinumab antibodies should be considered if treatment response is unsatisfactory, additional research is needed to identify tools for TDM in psoriasis patients on ustekinumab treatment.

As part of a RCT of treatment strategies in rheumatoid arthritis (RA), Hambardzumyan (2019) analyzed serum infliximab (sIFX) and anti-drug antibodies (ADAs) levels in study participants randomized to methotrexate + infliximab therapy and for whom serial serum sampling data at three, nine, and 21 months were available (n=101). The primary and secondary outcome measures were low disease activity [LDA = 28-joint Disease Activity Score (DAS28) ≤ 3.2] and remission (DAS28 < 2.6). The frequencies of very low sIFX levels increased over time, with 15%, 23%, and 28% at 3, 9, and 21 months from IFX start, respectively, and the majority of patients with very low sIFX levels were ADA positive at these time-points [71% (10/14), 82% (18/22), and 68% (19/28), respectively]. The proportion of patients with LDA was numerically higher at all follow-up time-points among those with sIFX ≥ 0.2 µg/mL compared with patients who had sIFX < 0.2 µg/mL and positive ADAs, although only significant at 21 months (67% and 26%, p=0.002). Similar results were observed when remission was the outcome measure (47% vs 11%, p=0.004). The authors concluded that these findings support the monitoring of serum drug levels, however, these findings require validation in larger populations and for dose-adjustment studies.

Van den Berghe (2018) published a small study evaluating ADA to vedolizumab in a cohort of 40 patients with IBD. This study included the development of an ELISA-based test to measure ADA in the presence of the drug. Antivedolizumab antibodies and vedolizumab trough levels were measured after six weeks of treatment and after treatment discontinuation. At the six-week follow-up, three (8%) of the patients were positive for ADA, but this appeared to be transient. None of the patients who discontinued vedolizumab were positive for ADA at the time of their last infusion or after discontinuation. The authors concluded that immunogenicity did not appear to play a major role in vedolizumab treatment failure.

Cludts (2017) conducted a single-center retrospective cohort analysis of patients with RA (n=18), psoriatic arthritis (n=9), or ankylosing spondylitis (n=12) in Italy. Serum samples were taken prior to adalimumab therapy and after 12 and 24 weeks of treatment. Psoriatic arthritis and ankylosing spondylitis patients were grouped together (SpA) due to axial involvement in all psoriatic arthritis patients. Although adalimumab levels varied among patients (0 to 30 µg/mL), median levels were significantly lower at 12 and 24 weeks in ATA-positive samples, and antibody formation was associated with decreasing levels of circulating
adalimumab. A reporter gene assay detected neutralizing antibodies against TNF antagonists in ATA-positive, therapeutic-negative patients; however, neutralization could not be confirmed in all ATA-positive samples due to adalimumab interference. There was a negative correlation between ADA levels and adalimumab in all groups, with 43.6% and 41% of the adalimumab-treated patients developing antibodies at 12 and 24 weeks, respectively. These percentages increased to 48.7% and 46% after subjecting the samples to acid treatment. There was a negative correlation between adalimumab trough levels and DAS28 and Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) scores (p<0.001). There were no significant differences between BASDAI in ATA-positive compared with ATA-negative patients at 12 or 24 weeks. The study is consistent with others suggesting that adalimumab levels can serve as an indicator of ATA; however, limitations included small sample size, retrospective research design, and failure to confirm neutralization in all ATA-positive samples.

Using an observational, cross-sectional study design, Ara-Martin (2017) analyzed the impact of immunogenicity on response to anti-TNF therapy in 137 adults with moderate-to-severe plaque psoriasis at 35 centers in Spain between 2012 and 2014.[20] All patients experienced secondary nonresponse to adalimumab (n=65), etanercept (n=47), and infliximab (n=19) after six or more months of treatment. Serum ADA was identified in 48%, 0%, and 42% of patients of patients treated with adalimumab, etanercept, and infliximab, respectively. Loss of efficacy was assessed using the PASI (PASI >5), 75% improvement in PASI score from baseline (PASI75), and/or the Physician Global Assessment (PGA, >2). PGA values for ADA-positive vs ADA-negative patients were significantly worse in the adalimumab group (3.7 vs 3.2, p=0.02) but not in the infliximab group. There was a significant negative linear correlation between serum drug concentrations and ADA in both the adalimumab group (p=0.001) and among the three groups combined (p=0.001), and a significant (p=0.019) correlation between serum ADA titer and body surface area. Unlike the other studies, in this study, the use of concomitant antirheumatic drugs was not associated with anti-TNF immunogenicity in any of the groups. This study provided evidence of antibody development against adalimumab and infliximab (not against etanercept) in patients with psoriasis, with ADA formation accounting for half of the secondary nonresponse associated with these therapies. However, conclusions were limited due to the cross-sectional study design, use of ELISA to detect ADAs due to drug interference, the potential presence of neutralizing antibodies as confounding factors, and limited information about patients’ health status prior to the study period.

A case-control, longitudinal study by Lombardi (2016) excludes possible confounding factors by analyzing adalimumab treatment for psoriasis in five distinct groups, including individuals who received: biologic therapies after switching from adalimumab (n=20); ongoing adalimumab therapy (n=30); novel adalimumab therapy (n=30); biologic therapies other than adalimumab (n=15); and no treatment with immunosuppressants or biologics (n=15), serving as a quasi-control.[21] The clinical severity of psoriasis was scored using the Psoriasis Area Severity Index (PASI). At 12-month follow-up, ADA was highest (87%) in patients who received biologic therapies after switching from adalimumab. The false-positive rate was 23% for adalimumab detection and 22% for anti-adalimumab antibodies in individuals who were never treated with adalimumab. There was no significant difference in median PASI score between the anti-adalimumab antibody-negative patients (1.1) and the anti-adalimumab antibody-positive patients (4.0). There was no association between PASI score or TNF-α concentration and the presence of anti-adalimumab antibodies in patients receiving adalimumab. Additionally, there were no significant differences in TNF-α and C-reactive protein concentrations. Study limitations included its observational design, small sample size, use of ELISA to measure ADA, and high variability of results. The authors concluded that the assay has limited clinical utility.
Chiu (2015) published a prospective observational study investigating the role of ustekinumab ADA in psoriasis.[22] The study included 76 individuals with plaque psoriasis who were treated with ustekinumab for at least seven months (mean 13 months). Antibodies to ustekinumab were found in five (6.5%) of the patients, and the presence of these antibodies was associated with lower serum levels of the drug (p<0.001) and lower PASI 50 response (p=0.004). Among the 15 patients who switched to ustekinumab from adalimumab, no difference in ustekinumab ADA was found between patients who had previously developed adalimumab ADA and those who did not.

Menting (2015) reported on the association between serum ustekinumab trough levels, ADA, and treatment efficacy in a small prospective study that included 41 patients with RA.[23] The mean follow-up time was 32 weeks (range 4 to 52 weeks), and during this period ADA to ustekinumab were detected in three patients. No correlations were seen between ustekinumab trough levels and clinical response to the medication.

While many studies have evaluated clinical validity using single ADA measurements, at least one study assessed their persistence over time. Vande Casteele (2013) analyzed infliximab trough and ADA levels using an HMSA assay with banked serum obtained from 90 IBD patients treated between May 1999 and August 2011.[24] ADA levels had been previously assayed using an ELISA-based test. A total of 1,232 samples were evaluated (mean 14 per patient). Treatment decisions were made solely on clinical evaluation and C-reactive protein levels. ADA were detected in 53 of 90 (59%) of patients but subsequently were nondetectable in 15 of the 53 (28%). Persistent ATIs were associated with discontinuation of infliximab (RR 5.1, 95% CI 1.4 to 19.0), but the wide confidence interval reflects considerable uncertainty. Although transience of ADA in IBD has not been carefully scrutinized, if replicated, these results suggest interpreting a single ADA result cautiously.

**Section Summary: Clinical Validity**

A large body of evidence has evaluated the clinical validity of ADA testing. ADA has been associated with secondary nonresponse in RA, SpA, and IBD. The presence of ADA has been consistently associated with an increased risk of an infusion-site reaction related to infliximab and injection-site reactions related to adalimumab. A concomitantly administered immunosuppressant agent may reduce the risk of developing ADA. Although ADA significantly reduced TNF-α response in a recent meta-analysis, considerable heterogeneity limits those findings. In addition, a recent observational study found no association between concomitant immunosuppressants and anti-TNF immunogenicity in patients with psoriasis; and a second cohort study found no association between PASI score or TNF-α concentration and the presence of anti-adalimumab antibodies in patients receiving adalimumab to treat psoriasis.

**CLINICAL UTILITY**

Several algorithms have been developed for management of patients with irritable bowel disease (IBD)[25-27] or rheumatoid arthritis (RA)[28] who have relapsed during TNF-inhibitor therapy. These algorithms are generally based on evidence that has indicated an association between ADA, reduced serum drug levels, and relapse. None has included evidence demonstrating improved health outcomes, such as reduced time to recovery from relapse (response), using algorithmic rather than dose-escalation approaches.

Syversen (2021) reported results of a randomized, parallel-group, open-label trial of 411 adults with RA, spondyloarthritis, psoriatic arthritis, ulcerative colitis, Chron’s disease (CD), or
psoriasis who received either proactive therapeutic drug monitoring of infliximab therapy based on serum infliximab level and ADA testing, or standard therapy without serum infliximab level or ADA testing.\textsuperscript{[29]} Serum trough infliximab levels and ADA levels were measured at each infusion in the therapeutic drug monitoring group. The infliximab dose or interval could be adjusted based on the therapeutic range during induction and during treatment. If ADA level was greater than 50 mcg/L at any point, therapy with infliximab was switched to a different agent. No significant difference between the therapeutic drug monitoring group and standard therapy group in clinical remission at week 30 was found (50.5\% versus 53\% of patients, respectively; \( p = 0.78 \)). During infliximab treatment, 36 (18\%) patients in the therapeutic drug monitoring group and 34 (17\%) in the standard therapy group developed ADAs ≥15 mcg/L. Antidrug antibodies ≥50 mcg/L (the threshold for discontinuation) occurred in 20 (10\%) of patients in the therapeutic drug monitoring group and 30 (15\%) in the standard therapy group. The remission rate in patients who developed ADAs was 56\% in the therapeutic drug monitoring group and 35\% in the standard therapy groups. The trial was limited by the small sample size of subjects who developed ADAs.

In a study of patients with IBD, Fernandex (2019) compared proactive monitoring of infliximab ADA and trough levels (\( n = 56 \)) to a retrospective control cohort (\( n = 149 \)).\textsuperscript{[30]} The primary outcomes were hospital admission, surgery, treatment discontinuation, and rates of mucosal healing. A composite “unfavorable outcome” comprised of all of these was also analyzed. There was an association between treatment escalation rates and proactive monitoring (60.7\% vs. 16.8\% of controls, \( p < 0.001 \)). After two years of follow-up, surgery rates were lower in the proactive group (8.9\% vs. 20.8\%, \( p = 0.030 \)) and mucosal healing was more common (73.2\% vs. 38.9\%, \( p < 0.0001 \)). No significant differences were seen in hospitalization rate or treatment discontinuation.

A similar retrospective study by Papamichael (2019) evaluated proactive monitoring of serum adalimumab levels and ADA (\( n = 53 \)) with standard of care, defined as empirical dose escalation (\( n = 279 \)) or reactive monitoring (\( n = 50 \)).\textsuperscript{[31]} Patients with early treatment failure (within eight weeks) were not included. After a median follow up of 3.1 years, fewer patients in the proactive monitoring group experienced treatment failure (hazard ratio [HR] 0.4, 95\% CI 0.2 to 0.9). No significant difference was found for the probability of IBD-related surgery.

Kamperidis (2019) published retrospective observational study on the impact of therapeutic drug level monitoring (TDM) on outcomes of 291 patients with Crohn's disease treated with Infliximab (IFX).\textsuperscript{[32]} Primary outcomes were clinicians' response to each TDM result and the rate of IFX discontinuation due to secondary loss of response or serious adverse event. Secondary outcomes included the intestinal surgery rate after IFX initiation and remission six months after TDM. Two hundred thirty-eight (81.8\%) patients were tested for TDM at least once during their follow-up with 672 TDM results. 95/238 patients (39.9\%) had undetectable levels and 76 (31.9\%) had positive antibodies to infliximab (ATI) at least once. IFX was discontinued in 109 patients (37.5\%). TDMs results were not followed by altered patient management in 526/672 (78.3\%) of the observations. Treatment was discontinued in 40 (75.5\%) patients never tested for TDM compared with 69 (29.0\%) of those tested (\( p < 0.01 \)). Fewer TDM tested patients (29; 12.2\%) required intestinal surgery post IFX initiation compared with those not TDM tested (15; 28.3\%). In this retrospective study, data collected on clinical outcomes relied on record keeping and physician response was taken as the measure of clinical remission. These methods may be subject to interpretation bias.
Dong (2019) reported an observational study of 60 patients with ankylosing spondylitis (AS) taking a biosimilar of etanercept. Serum drug levels and anti-drug antibody levels, as well as clinical measures of disease activity were assessed at baseline and after four, 12, and 24 weeks of treatment. The authors found that anti-drug antibodies had no effect on the Assessment of Spondylosis Arthritis International Society (ASAS) remission rates but reported that patients with ADA had lower drug levels and higher TNF-α levels.

Steenholdt (2014) reported results of a noninferiority trial and cost-effectiveness analysis of 69 patients with CD who relapsed (CDAI ≥220 and/or ≥1 draining perianal fistula) during infliximab therapy. Patients were randomized to infliximab dose intensification (5 mg/kg every four weeks) or algorithmic treatment based on serum infliximab level and ATI: Patients with subtherapeutic infliximab level (<0.5 μg/mL) had infliximab dose increased if ADA were undetectable or were switched to adalimumab if ADA were detectable; patients with therapeutic infliximab level underwent repeat testing of infliximab and ADA levels if ADA were detectable or diagnostic reassessment if ADA were undetectable. Serum infliximab and ADA levels were measured in all patients using RIA in single-blind fashion (patients unaware but investigators aware of test results). Randomized groups were similar at baseline; overall, 55 (80%) of 69 patients had nonfistulizing disease. Most patients (70%) had therapeutic serum infliximab levels without detectable ATI; revised diagnoses in 6 (24%) of 25 such patients in the algorithm arm included bile acid malabsorption, strictures, and IBS. In both intention-to-treat (ITT) and per-protocol analyses, similar proportions of patients in each randomized group achieved clinical response at week 12, defined as a minimum 70-point reduction from baseline CDAI for patients with nonfistulizing disease and a minimum 50% reduction in active fistulas for patients with fistulizing disease (ITT 58% in the algorithm group vs 53% in the control group, p=0.810; per-protocol, 47% in the algorithm group vs 53% in the control group, p=0.781). Only the ITT analysis fell within the prespecified noninferiority margin of -25% for the difference between groups.

Conclusions on the noninferiority of an algorithmic approach compared with dose intensification from this trial are limited. The noninferiority margin was arguably large and was exceeded in the conservative per-protocol analysis. Dropouts were frequent and differential between groups; 17 (51%) of 33 patients in the algorithm group and 28 (78%) of 36 patients in the control group completed the 12-week trial. A large proportion of patients (24%) in the algorithmic arm were potentially misdiagnosed (i.e., CD flare was subsequently determined not to be the cause of relapse); the comparable proportion in the control arm was not reported. In most patients (80% who had nonfistulizing disease), only a subjective measure of treatment response was used (minimum 70-point reduction from baseline CDAI).

Roblin (2014) conducted a single-center, prospective observational study of 82 patients with IBD (n=45 CD, n=27 UC) with clinical relapse (CDAI >220 or Mayo Clinic >5) during treatment with adalimumab 40 mg every two weeks. For all patients, trough adalimumab levels and ADA were measured in a blinded fashion using ELISA, and adalimumab dose was optimized to 40 mg weekly. Those who did not achieve clinical remission (CDAI <150 or Mayo score <2) within four months underwent repeat trough adalimumab and anti-adalimumab antibody testing and were switched to infliximab. Clinical and endoscopic responses after adalimumab optimization and after infliximab therapy for six months were compared across three groups: (1) those with a therapeutic adalimumab level (>4.9 μg/mL), (2) those with a subtherapeutic adalimumab level and undetectable ATA; and (3) those with a subtherapeutic adalimumab level and detectable ADA. After adalimumab optimization, more group 2 patients achieved clinical remission (16 [67%] of 24 patients) than group 1 (12 [29%] of 41 patients; p<0.01 vs
group 2) and group 3 (2 [12%] of 17 patients, p<0.01 vs group 2) patients. Duration of remission was longest in group 2 (mean 15 months) compared with group 1 (mean five months) and group 3 (mean, four months, p<0.01 for both comparisons vs group 2). At one year, 13 (52%) of 24 patients in group 2 maintained clinical remission compared with no patients in groups 1 or 3 (p<0.01 for both comparisons vs group 2). Results were similar when remission was defined using calprotectin levels (<250 μg/g stool) or endoscopic Mayo score (<2).

Fifty-two patients (n=30 CD, n=22 UC) who failed to achieve clinical remission after adalimumab optimization were switched to infliximab. More patients in group 3 achieved clinical remission (12 [80%] of 15 patients) than in group 1 (2 [7%] of 29 patients) or group 2 (2 [25%] of 8 patients, p<0.01 for both comparisons vs group 3). Duration of response after switching to infliximab was longest in group 3 (mean, 14 months) compared with group 1 (mean, three months) and group 2 (mean, five months, p<0.01 for both comparison vs group 3). At one year, 8 (55%) of 15 patients in group 3 maintained clinical remission compared with no patients in groups 1 or 2 (p<0.01 for both comparisons vs group 3). Results were similar using objective measures of clinical remission (calprotectin level, endoscopic Mayo score).

These results suggested that patients with IBD who relapse on adalimumab and have subtherapeutic serum adalimumab levels may benefit from a higher adalimumab dose if ADA are undetectable or from a change to another TNF inhibitor if ADA are detectable. Relapsed patients who have therapeutic serum adalimumab levels may benefit from change to a different drug class. Strengths of the study include its use of subjective and objective measures of remission and blinded serum drug level and ADA monitoring. However, results were influenced by the small sample size, use of ELISA for antibody testing, and lack of ADA levels for decision making. Studies comparing management using the algorithm proposed with usual care are needed.

Affi (2010) evaluated the clinical utility of measuring ADA (referred to as human antichimeric antibodies [HACA] in the study) and infliximab concentrations by retrospectively reviewing patient medical records.[39] Record review from 2003 to 2008 identified 155 patients who had had ADA, had data on infliximab concentrations, and met the study inclusion criteria. A single physician ordered 72% of the initial tests. The authors retrospectively determined clinical response to infliximab. Forty-seven percent of patients were on concurrent immunosuppressive medication. The main indications for testing were loss of response to infliximab (49%), partial response after initiation of infliximab (22%), and possible autoimmune or delayed hypersensitivity reaction (10%). ADA were identified in 35 (23%) patients and therapeutic infliximab concentrations in 51 (33%) patients. Of 177 tests assessed, the results impacted treatment decisions in 73%. In ATI-positive patients, change to another anti-TNF agent was associated with a complete or partial response in 92% of patients, whereas dose escalation occurred in 17%. The authors concluded that measurement of ADA and infliximab concentration had a clinically useful effect on patient management. The strategy of increasing infliximab dose in patients with ADA was ineffective whereas in patients with subtherapeutic infliximab concentrations this strategy was a good alternative to changing to another anti-TNF agent. Study limitations included the retrospective design and using ELISA testing for ADA. Because there was no control group, one cannot determine what changes in management would have been made absent ADA measurement. Because clinicians are likely to change management for patients who do not achieve or maintain a clinical response, it is important to understand how these management decisions differ when ADA are measured.
Section Summary: Clinical Utility

Significant evidence for the clinical utility of ADA testing is currently lacking. Uncontrolled retrospective studies in IBD have demonstrated the impact of ADA testing on treatment decisions but cannot demonstrate improved patient outcomes compared with a no-testing strategy. Additional limitations of these studies included a lack of clinical follow-up after treatment decisions were made and a lack of clinical assessments to guide treatment decisions. Additionally, the determination of a clinically relevant threshold for the ADA level is complicated by the use of various assay methods. A small, nonrandomized prospective study suggested that ADA levels may be informative in relapsed patients with IBD who have low serum adalimumab levels, but this finding requires confirmation in larger, randomized trials. Methodologic flaws, including relapse misclassification, limit conclusions from the RCT in patients with relapsed IBD. Direct or indirect evidence for clinical utility in patients with RA or SpA was not identified.

PRACTICE GUIDELINE SUMMARY

AMERICAN COLLEGE OF GASTROENTEROLOGY

In 2019, the American College of Gastroenterology published a guideline on ulcerative colitis (UC).\[40\] The guideline stated: "In patients with moderately to severely active UC who are responders to anti-TNF therapy and now losing response, we suggest measuring serum drug levels and antibodies (if there is not a therapeutic level) to assess the reason for loss of response (conditional recommendation, very low quality of evidence)."

In 2018, the American College of Gastroenterology published a guideline on Crohn’s disease (CD).\[41\] Although acknowledging that a detailed review of therapeutic drug monitoring was beyond the scope of the guideline, it stated: "If active CD is documented, then assessment of biologic drug levels and antidrug antibodies (therapeutic drug monitoring) should be considered."

AMERICAN GASTROENTEROLOGICAL ASSOCIATION

In 2017, the American Gastroenterological Association published an evidence-based clinical practice guideline on therapeutic drug monitoring (TDM) in inflammatory bowel disease (IBD).\[42\] The guideline was developed according to the GRADE framework to evaluate certainty of evidence, and a Technical Review was published to accompany the recommendations.\[43\] Regarding measurement of anti-drug antibodies, the Association made the following statement:

“In adults with active IBD treated with anti-TNF agents, the AGA suggests reactive therapeutic drug monitoring to guide treatment changes.” Conditional recommendation, very low quality of evidence.

According to the GRADE method, very low quality is defined as: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

The guideline also stated:
In adult patients with quiescent IBD treated with anti-TNF agents, the AGA makes no recommendation regarding the use of routine proactive therapeutic drug monitoring. No recommendation, knowledge gap.

**SUMMARY**

Antibodies to drugs for chronic inflammatory diseases including, but not limited to infliximab, adalimumab, ustekinumab, certolizumab, etanercept, golimumab, and vedolizumab, are present in a substantial number of patients treated with these medications. A correlation between the level of these antibodies and clinical response has been identified in patients with some chronic inflammatory conditions.

There is some evidence that, in patients with inflammatory bowel disease who have lost response to infliximab or adalimumab, measurement of serum drug antibodies can impact patient care decisions. Evidence-based clinical practice guidelines recommend reactive monitoring of serum drug levels and anti-drug antibodies to guide treatment changes in patients with active inflammatory bowel disease who are being treated with an anti-TNF agent. Therefore, measurement of serum antibodies to infliximab (Remicade, Inflectra, Renflexis) or adalimumab (Humira), either alone or as a combination test that includes serum drug levels, may be considered medically necessary for patients with inflammatory bowel disease (i.e., Crohn’s disease or ulcerative colitis) when there is documentation of a loss of response to these medications.

There is not enough evidence to show that measurement of serum drug antibodies, either alone or as a combination test that includes serum drug levels, improves net health outcomes when there has not been a loss of response to the medication. No evidence-based clinical practice guidelines recommend the measurement of serum drug antibodies when there has not been a loss of response to medication. Therefore, measurement of serum drug antibodies, either alone or as a combination test that includes serum drug levels, is considered not medically necessary when there has not been a loss of response to the medication.

There is not enough research to determine whether measurement of serum anti-drug antibodies can be used in patient management to improve net health outcomes for all conditions. The optimal timing of when to measure antibody levels and measurement cutoff levels has not been established. No evidence-based clinical practice guidelines recommend testing for serum drug antibodies in the treatment of chronic inflammatory conditions other than anti-TNF agents in the treatment of inflammatory bowel disease. Therefore, measurement of serum drug antibodies, either alone or as a combination test that includes serum drug levels, other than infliximab or adalimumab in the treatment of inflammatory bowel disease, is considered investigational.

**REFERENCES**


---

**CODES**

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>80145</td>
<td>Adalimumab</td>
</tr>
</tbody>
</table>

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80230</td>
<td>Infliximab</td>
</tr>
<tr>
<td></td>
<td>80280</td>
<td>Vedolizumab</td>
</tr>
<tr>
<td></td>
<td>80299</td>
<td>Quantitation of therapeutic drug, not elsewhere specified</td>
</tr>
<tr>
<td></td>
<td>84999</td>
<td>Unlisted chemistry procedure</td>
</tr>
<tr>
<td>HCPCS</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

*Date of Origin: January 2013*
IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Hyperbaric oxygen therapy (HBOT) is a technique of delivering higher pressures of oxygen to the tissues. Two methods of administration are available, systemic and topical.

MEDICAL POLICY CRITERIA

I. Systemic hyperbaric oxygen therapy may be considered medically necessary when both of the following criteria (A. and B.) are met:

   A. Systemic hyperbaric oxygen therapy services must comply with the following guidelines which are consistent with the Undersea and Hyperbaric Medical Society criteria:

      1. Patient must breathe 100% oxygen intermittently or continuously while the pressure of the treatment chamber is increased above one atmosphere absolute; and

      2. Systemic hyperbaric oxygen pressurization should be at least 1.4 atmospheres absolute (atm abs) (20.5 psi); and

      3. Treatment is provided in a hospital or clinic setting; and

   B. Treatment meets one or more of the following conditions:
1. Acute carbon monoxide poisoning (*Recommended treatment review threshold: 5 treatments*); or
2. Acute traumatic ischemia (*Recommended treatment review threshold:* Reperfusion injury – 1 treatment; Crush injury – 12 treatments (3 times per day for 2 days, then twice a day for 2 days, then daily for 2 days); Compartment syndrome – 3 treatments (twice a day for 1 day, then 1 treatment on day 2); or
3. Chronic refractory osteomyelitis (*Recommended treatment review threshold: 40 treatments*); or
4. Cyanide poisoning, acute (*Recommended treatment review threshold: 5 treatments*); or
5. Decompression sickness (*Recommended treatment review threshold: 10 treatments*); or
6. Gas or air embolism, acute (*Recommended treatment review threshold: 10 treatments*); or
7. Gas gangrene (i.e., clostridial myositis and myonecrosis; *Recommended treatment review threshold: 10 treatments*); or
8. Non-healing diabetic wounds of the lower extremities as an adjunct to ongoing conventional wound care in patients who meet all of the following Criteria (a. – c.) (*Recommended treatment review threshold: 30 treatments (one or two treatments daily):*
   a. Patient has type I or type II diabetes and has a lower extremity wound that is due to diabetes; and
   b. Patient has a wound classified as Wagner grade 3 or higher (see Policy Guidelines); and
   c. Patient has no measurable signs of healing after 30 days of an adequate course of standard wound therapy including all of the following:
      i. Assessment of vascular status and correction of any vascular problems in the affected limb if possible; and
      ii. Optimal glycemic control; and
      iii. Optimal nutritional status; and
      iv. Topical wound treatment (e.g., saline, hydrogels, hydrocolloids, alginates) with maintenance of a clean, moist bed of granulation tissue; and
      v. Debridement to remove devitalized tissue, any technique; and
      vi. Pressure reduction or offloading; and
      vii. Treatment to resolve infection (e.g., antibiotics); or
9. Pre- and post-treatment for patients undergoing dental surgery (non-implant-related) of an irradiated jaw; or
10. Profound anemia with exceptional blood loss: only when blood transfusion is impossible or must be delayed (*Recommended treatment review threshold:*
HBOT should be continued with taper of both time and frequency until red blood cells have been satisfactorily replaced by patient regeneration or the patient can undergo transfusion; or

11. Soft-tissue radiation necrosis (e.g., radiation enteritis, cystitis, proctitis) and osteoradionecrosis (Recommended treatment review threshold for mandibular osteoradionecrosis: 60 treatments); or

12. Idiopathic Sudden Sensorineural Hearing Loss of greater than or equal to 41 decibels and an onset of treatment within 14 days (Recommended treatment review threshold: 20 treatments.); or

13. Necrotizing soft tissue infections; or

14. Actinomycosis; or

15. Central retinal artery occlusion; or

16. Compromised skin grafts and flaps where hypoxia or decreased perfusion has compromised viability acutely (Recommended treatment review threshold: 30 treatments.)

II. Systemic hyperbaric oxygen for non-healing diabetic wounds of the lower extremities as an adjunct to conventional wound care is considered **not medically necessary** when Criterion I.B.8 is not met.

III. Systemic hyperbaric oxygen therapy is considered **investigational** for all other indications including but not limited to other ophthalmologic conditions, non-diabetic wounds, and acute thermal burns.

IV. Topical hyperbaric and topical normobaric oxygen therapies are considered **investigational**.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

**POLICY GUIDELINES**

**WAGNER CLASSIFICATION**

- Grade 0: No open lesion
- Grade 1: Superficial ulcer without penetration to deeper layers
- Grade 2: Ulcer penetrates to tendon, bone, or joint
- Grade 3: Lesion has penetrated deeper than grade 2 and there is abscess, osteomyelitis, pyarthrosis, plantar space abscess, or infection of the tendon and tendon sheaths
- Grade 4: Wet or dry gangrene in the toes or forefoot
- Grade 5: Gangrene involves the whole foot or such a percentage that no local procedures are possible and amputation (at least at the below the knee level) is indicated

**LIST OF INFORMATION NEEDED FOR REVIEW**

It is critical that the list of information below is submitted for review to determine if the policy criteria are met. If any of these items are not submitted, it could impact our review and decision.
out

- History and physical/chart notes
- Indication for the requested service including type of HBOT planned
- Treatment plan including the following:
  - Percent of oxygen that the patient will breathe while receiving therapy
  - Pressurization (atm abs, psi)
  - Treatment setting
- Condition being treated including how many treatments being requested
  - If a diabetic wound is being treated then the request must include the following:
    - Type of diabetes
    - Location of wound
    - Wagner Classification
    - Measurable signs of healing following standard wound therapy including therapy length of time with documentation of the following:
      - Vascular assessment and correction, if possible, of vascular problems to affected area
      - Glycemic data for patient (e.g., A1C)
      - Nutritional status
      - Topical wound treatments utilized including wound bed description
      - Debridement
      - Pressure reduction or offloading
      - Any infection treatment utilized
    - If dental surgery, include description and diagnosis
    - If anemia, include blood loss and ability to transfuse patient
    - If necrosis, include type
    - If idiopathic sudden sensorineural hearing loss, include decibels of loss and onset of treatment

CROSS REFERENCES

None

BACKGROUND

SYSTEMIC HBOT

In systemic or large chamber hyperbaric oxygen, the patient is entirely enclosed in a pressure chamber and breathes oxygen at a pressure greater than 1 atmosphere (atm, the pressure of oxygen at sea level). Thus, this technique relies on systemic circulation to deliver highly oxygenated blood to the target site, typically a wound. In addition, systemic hyperbaric oxygen therapy can be used to treat systemic illness, such as air or gas embolism, carbon monoxide poisoning, clostridial gas gangrene, etc. Treatment may be carried out either in a monoplace chamber pressurized with pure oxygen or in a larger, multiplace chamber pressurized with compressed air, in which case the patient receives pure oxygen by mask, head tent, or endotracheal tube.

Mild Hyperbaric Oxygen Therapy

Oxygen therapy delivered via soft-sided chambers is referred to as mild hyperbaric oxygen therapy. While this implies that these chambers provide HBOT, the therapy is not considered...
hyperbaric as they provide pressurization of only about 4.5 psi, compared with true HBOT which is defined as pressurization of 20.5 psi or higher.

**TOPOCAL OXYGEN THERAPY**

**Topical Hyperbaric Oxygen Therapy**

Topical hyperbaric oxygen therapy is a technique of delivering 100% oxygen directly to an open, moist wound at a pressure slightly higher than atmospheric pressure. It is hypothesized that the high concentrations of oxygen diffuse directly into the wound to increase the local cellular oxygen tension, which in turn promotes wound healing. This therapy has been investigated as a treatment of skin ulcerations resulting from diabetes, venous stasis, postsurgical infection, gangrenous lesion, decubitus ulcers, amputations, skin graft, burns, or frostbite.

Topical hyperbaric oxygen devices consist of an appliance to enclose the wound area (frequently an extremity) and a source of oxygen; conventional oxygen tanks may be used. Topical hyperbaric oxygen therapy may be performed in the office, clinic, or may be self-administered by well-trained patients in the home. Typically, the therapy is offered for 90 minutes per day for 4 consecutive days. After a 3-day break, the cycle may be repeated. The regimen may last for 8 to 10 weeks.

**Topical Normobaric Oxygen Therapy**

Devices that deliver topical oxygen to a wound at normal atmospheric pressure (normobaric) are not considered hyperbaric oxygen therapy. These devices may also be called low dose tissue oxygenation systems. An example of a normobaric oxygen delivery system is the TransCu O2™, a small handheld device with an attached cannula. According to the manufacturer, the TransCu O2 is “intended for use with wound dressings to treat the following: skin ulcerations due to diabetes, venous stasis, post-surgical infections and gangrenous lesions; pressure ulcers; infected residual limbs; skin grafts; burns; and frostbite.” The device concentrates room air to 99.9% oxygen which is delivered via the cannula which is placed under the wound dressing.

**REGULATORY STATUS**

In 2013, U.S. Food and Drug Administration (FDA) published a statement warning that non-FDA approved uses of HBOT may endanger the health of patients.[1] “Patients may incorrectly believe that these devices have been proven safe and effective for uses not cleared by FDA, which may cause them to delay or forgo proven medical therapies. In doing so, they may experience a lack of improvement and/or worsening of their existing condition(s).”

The following are examples of oxygen therapy devices:

In February 1999, the Numobag™ Kit (Numotech, Inc) for application of topical hyperbaric therapy was cleared for marketing by the FDA through the 510(k) process. The FDA determined that this device was substantially equivalent to existing devices. Another example is the AOTI Hyper-Box™ (AOTI Ltd., Galway, Ireland) which was cleared by FDA in 2008.

In August 2009, the TransCu O2 (Electrochemical Oxygen Concepts, Inc.) was cleared for marketing by the FDA through the 510(k) process as substantially equivalent to existing...
devices.

There are numerous FDA-approved hyperbaric oxygen chambers. In May 2005, the ATA Monoplace Hyperbaric System (ATA Hyperbaric Chamber Manufacturing, Inc.) was cleared for marketing by the FDA through the 510(k) process. The FDA determined that this device was substantially equivalent to existing hyperbaric devices.

**EVIDENCE SUMMARY**

Current evidence is sufficient to determine the effectiveness of hyperbaric oxygen therapy (HBOT) for the indications that meet the above medical necessity criteria. Assessing the effectiveness and safety of HBOT for the investigational indications requires randomized controlled trials comparing HBOT with the conventional treatments for each indication. Therefore, the following literature review on HBOT focuses on randomized controlled trials (RCTs) and systematic reviews of RCTs for the investigational indications.

Assessment of efficacy for therapeutic interventions involves a determination of whether the intervention improves health outcomes. The optimal study design for a therapeutic intervention is a randomized controlled trial (RCT) that includes clinically relevant measures of health outcomes. Intermediate outcome measures, also known as surrogate outcome measures, may also be adequate if there is an established link between the intermediate outcome and true health outcomes. When the primary outcomes are subjective (e.g., pain, depression), sham-controlled RCTs are needed to assess the effect of the intervention beyond that of a placebo effect.

**TOPICAL HYPERBARIC OXYGEN**

Due to their different methods of delivery, topical and systemic hyperbaric oxygen are distinct technologies such that they must be examined separately.\[2\] There is minimal published literature regarding topical hyperbaric oxygen therapy. A 2015 Cochrane review of interventions for treating gas gangrene evaluated the safety and efficacy of topical HBOT and Chinese herbs as treatments options.\[3\] Re-analysis if cure rate did not show beneficial effects from either treatment. In 1984, Heng published a controlled study of topical hyperbaric oxygen therapy in 6 patients with 27 ulcers compared to no treatment in 5 patients with 10 ulcers.\[4\] Although a greater improvement was noted in the treated group, the results were calculated according to the number of ulcers rather than based on individual patients. Leslie reported on a trial that randomly assigned 18 patients with diabetic foot ulcers to receive either topical hyperbaric oxygen therapy plus standard wound care or standard wound care alone.\[5\] Changes in ulcer size and depth did not differ between the 2 groups. Other studies consist of anecdotal reports or uncontrolled case series.\[6\]

**SYSTEMIC HYPERBARIC OXYGEN THERAPY (HBOT)**

**In-home Hyperbaric Oxygen**

A position statement from the National Board of Diving & Hyperbaric Medical Technology on in-home HBOT has been published on the Web site for The Undersea and Hyperbaric Medicine Society (UHMS).\[7\] The statement indicates that in-home HBOT “is inherently unsafe and cannot be condoned.” This position is based on concern for the safety and well-being of patients as well as those people in proximity to the HBOT delivery system because in-home provision of HBOT is likely to:
- Bypass otherwise mandatory federal, state, and local codes related to design, construction, installation, and operation of these devices; and
- Occur without adequate physician oversight and the operational support of appropriately qualified HBOT providers.

**Acute Coronary Syndromes**

**Systematic Reviews**

A 2012 Cochrane review by Bennett identified 6 trials with a total of 665 patients evaluating HBO for acute coronary syndrome.[8] All of the studies included patients with acute myocardial infarction (MI); one study also included individuals presenting with unstable angina. Additionally, all trials used HBOT as an adjunct to standard care. Control interventions varied; only 1 trial described using a sham therapy to blind participants to treatment group allocation. In a pooled analysis of data from 5 trials, there was a significantly lower rate of death in patients who received HBOT compared to a control intervention (RR: 0.58; 0.36 to 0.92). Due to variability of outcome reporting in the studies, few other pooled analyses could be conducted. A pooled analysis of data from 3 trials on improvements in left ventricular function did not find a statistically significant benefit of HBOT (RR: 0.09; 95% CI: 0.01 to 1.4). The authors noted that, although there is some evidence from small trials that HBOT is associated with a lower risk of death, larger trials with high methodologic quality are needed in order to determine which patients, if any, can be expected to derive benefit from HBOT. Therefore, HBOT is considered investigational in the treatment of acute coronary syndromes.

**Autism Spectrum Disorders (ASD)**

**Systematic Reviews**

A 2016 systematic review on hyperbaric oxygen therapy for treatment of children with autism identified one RCT[9] with a total of 60 children. The study quality was rated as low using GRADE criteria with small sample size and wide confidence intervals. The results indicated no improvement in social interaction and communication, behavioral problems, communication and linguistic abilities, or cognitive function. The authors reported minor-grade ear barotrauma as adverse events.

A 2012 systematic review[10] of RCTs on hyperbaric oxygen therapy for treatment of children with autism identified two RCTs[11, 12] with a total of 89 participants. In both RCTs the active hyperbaric treatment was 24% oxygen delivered at an atmospheric pressure of 1.3 atmospheres (atm). Although this regimen was referred to as HBOT in the article, it differed from standard HBOT which uses 100% oxygen and a pressure of at least 1.4 atm. A detailed analysis of these RCTs is provided below. Briefly, one of the two RCTs found better outcomes after hyperbaric oxygen compared with placebo treatment, and the other did not find significant differences in outcomes. The author concluded that additional sham-controlled trials with rigorous methodology are needed in order to draw conclusions about the efficacy of HBOT for treating autism.

**Randomized Controlled Trials (RCTs)**

The following is a summary of the 2 RCTs reported in the above systematic review:

- One of the above two RCTs was by Rossignol.[11] This study was a double-blind RCT that included 62 children, ages 2-7, meeting DSM-IV criteria for autistic disorder. The
active treatment was hyperbaric treatment at 1.3 atmospheres (atm) and 24% oxygen in a hyperbaric chamber. (This regimen differs from standard HBOT which uses 100% oxygen and a pressure of at least 1.4 atm). The other group received a sham treatment consisting of 1.03 atm and ambient air (21% oxygen). Both groups received 40 sessions of active or sham treatment lasting 60 minutes each over a period of 4 weeks. The equipment, procedures, etc. in the two groups were as similar as possible to maintain blinding. The investigators, participants, parents, and clinic staff were blinded to treatment group. Only the hyperbaric technician, who had no role in outcome assessment, was aware of group assignment. After completion of the 4-week study, families with children in the control group were offered the active intervention. When asked at the end of the study, there was no significant difference in the ability of parents to correctly guess the group assignment of their child.

The outcomes were change compared to baseline after 4 weeks on the following scales: Aberrant Behavior Checklist (ABC) total score and 5 subscales; Autism Treatment Evaluation Checklist (ATEC) total score and 4 subscales; and Clinical Global Impression-Improvement (CGI) overall functioning score and 18 subscales. P values of <0.05 were considered statistically significant; there was no adjustment for multiple comparisons. The analysis included all children who completed at least one complete session. Of the 33 children assigned to active treatment, 30 were included in the analysis and 29 completed all 40 treatments. Of the 29 children assigned to the control treatment, 26 completed all 40 sessions and were included in the analysis.

There was no significant between-group improvement on the ABC total score, any of the ABC subscales, or on the ATEC total score. Compared to the control group, the treatment group had a significant improvement in 1 of 4 subscales of the ATEC, the sensory/cognitive awareness subscale. The change from baseline on this subscale was a mean of 16.5 in the treatment group and a mean of 5.4 in the control group, a difference of 11.1 (p=0.037). (Note: due to an administrative error, baseline ATEC was not collected at one site, and thus data were not available for 23 children in the treatment group and 21 children in the control group). On the physician-rated CGI total score, 9/30 (30%) children in the treatment group had a score of 1 (very much improved) or 2 (much improved) compared to 2/26 (8%) in the control group (p=0.047). On the parental-rated CGI total score, 9/30 (30%) children in the treatment group had a score of 1 or 2 compared to 4/26 (15%) in the control group (p=0.22, not statistically significant). (The exact numbers receiving scores of 1 vs. 2 were not reported). Change in mean CGI scores were also reported, but this may be a less appropriate way to analyze these data. Among the parental-rated CGI subscales, significantly more children were rated as improved in the treatment group compared to control on 2 out of 18 subscales, receptive language (p=0.017) and eye contact (p=0.032).

A key limitation of this study was that the authors reported only outcomes at 4 weeks, directly after completion of the intervention. It is not known whether there are any long-term effects. Additional follow-up data cannot be obtained because members of the control group crossed over to the intervention after 4 weeks. Other limitations included lack of adjustment for multiple comparisons and unclear clinical significance of the statistically significant outcomes. The Undersea and Hyperbaric Medical Society (UHMS) issued a position paper after publication of the Rossignol et al. study stating that they still did not recommend routine treatment of autism with HBOT. [13]
The other RCT included in the systematic review was a double-blind RCT that began with 46 children with autism, ages 2-14 years, who were matched in pairs according to age and the number of hours of Applied Behavior Analysis (ABA) treatment they were receiving at the start of the study. Randomized treatment allocation of the matched pairs was by coin toss. Both groups received 80 1-hour sessions of active treatment (24% oxygen at 1.3 atm) or sham treatment (room air at ambient pressure) for up to 15 weeks. Participants were allowed to undergo ABA, take any supplements, pharmacological interventions, and dietary modifications. Twelve patients withdrew from the trial, leaving 18 patients in the treatment group and 16 in the control group.

The primary outcome of change in symptoms was based on direct observation and the scales noted in the Rossignol et al. study above in addition to the Behavior Rating Inventory of Executive Functioning (BRIEF), Parent Stress Index (PSI), Peabody Picture Vocabulary Test (PPVT-III), Repetitive Behavior Scale (RBS), and the Vineland Adaptive Behavior Scales (VABS-II). Direct observation and intention to treat analysis of test scores found no significant difference on any outcome measures between the treatment and sham groups. No participants experienced adverse effects attributable to barotrauma (e.g., pressure injury to tympanic membranes or sinuses).

A limitation of this study was the small sample size which was determined to be adequate to detect only large effects, which were not present in this study. In addition, since some patients in both groups received intensive ABA interventions during the study period, any potential effects of HBOT could not be isolated. The authors concluded that the active treatment had no significant beneficial effect on ASD and was not recommended for the treatment of ASD symptoms.

One additional RCT not included in the systematic review above was identified:

A 2012 RCT published after the systematic review randomly assigned 60 children with autism to receive 20 one-hour sessions with HBOT or sham air treatment (n=30 per group). The primary outcome measures were change in the ATEC and CGI, evaluated separately by clinicians and parents. There were no statistically significant differences between groups on any of the primary outcomes. For example, post-treatment clinician-assessed mean scores on the ATEC were 52.4 in the HBOT group and 52.9 in the sham air group.

Conclusion

There is insufficient evidence from well-designed RCTs that HBOT improves health outcomes for patients with autism spectrum disorder; therefore, HBOT therapy for this indication is considered investigational.

Bell’s Palsy

Systematic Review

In 2012, Holland published a Cochrane review evaluating HBOT in adults with Bell’s palsy. The authors identified one RCT with 79 participants, and this study did not meet the Cochrane review methodologic standards because the outcome assessor was not blinded to treatment allocation. Therefore, the evidence is insufficient to permit conclusions and HBOT is considered investigational for the treatment of Bell’s palsy.
Randomized Controlled Trials (RCTs)

No RCTs have been published since the 2012 Cochrane review.

Bisphosphonate-related Osteonecrosis of the Jaw (BRONJ)

Randomized Controlled Trials (RCTs)

An unblinded RCT was published by Freiberger in 2012 on use of HBOT as an adjunct therapy for patients with bisphosphonate-related osteonecrosis of the jaw.\[16\] Forty-nine patients were randomly assigned to HBOT in addition to standard care (n=22) or standard care alone (n=27). Five patients in the standard care group received HBOT and 1 patient assigned to the HBOT group declined HBOT. The investigators decided to do a per protocol (PP) analysis (actual treatment received) because of the relatively large degree of crossover. Participants were evaluated at 3, 6, 12 and 18 months. Data were available on 46 patients, 25 received HBOT in addition to standard care and 21 received standard care alone. The primary outcome measure was change in oral lesion size or number. When change from baseline to last available follow-up was examined, 17 of 25 (68%) of HBO-treated patients had improvement in oral lesion size or number compared to 8 of 21 (38%) in the standard care group, p=0.043. When change from baseline to 6, 12 or 18 months was examined, there was not a statistically significant difference between groups in the proportion of patients with improvement. In addition, the proportion of patients who healed completely did not differ significantly between groups at any time point. This single trial does not report consistent findings of benefit across outcome measures. It also has a number of methodologic limitations, e.g., unblinded, cross-over, and analysis performed on a per-protocol basis rather than intention to treat. A disadvantage of the per-protocol analysis is that randomization is not preserved, and the two groups may differ on characteristics that affect outcomes. As a result, this trial is insufficient to conclude that HBOT improves health outcomes for patients with bisphosphonate-related osteonecrosis of the jaw.

Conclusion

Current evidence is insufficient to determine the safety and efficacy of HBOT in the treatment of bisphosphonate-related osteonecrosis of the jaw. Therefore, HBOT is considered investigational for this indication.

Cancer Treatment

Randomized Controlled Trials (RCTs)

In an RCT of 32 patients, Heys found no increase in 5-year survival in patients treated with HBOT prior to chemotherapy for locally advanced breast carcinoma to increase tumor vascularity.\[17\] This approach is being studied since studies in animal models have suggested that HBOT increases tumor vascularity and thus may make chemotherapy more effective. In a Cochrane review, Bennett concluded that HBOT given with radiotherapy may be useful in tumor control; however, the authors expressed caution since significant adverse effects were common with HBOT and indicated further study would be useful.\[18\]

Conclusion

Current evidence is insufficient to determine the safety and efficacy of HBOT in the treatment of cancer of any type and location. Therefore, HBOT is considered investigational for this indication.
Cerebral Palsy

Randomized Controlled Trials (RCTs)

- In 2012, Lacey published a double-blind RCT that included 49 children age 3-8 years with spastic cerebral palsy. Participants were randomized to receive 40 treatments with either HBOT (n=25) or hyperbaric air to simulate 21% oxygen at room air (n=24). The primary efficacy outcome was change in the Gross Motor Function Measure (GMFM-88) global score after the 8-week treatment period. The study was stopped early due to futility, when an interim analysis indicated that there was less than a 2% likelihood that a statistically significant difference between groups would be found. At the time of the interim analysis, there was no significant between-group difference in the post-treatment GMFM-88 global score (p=0.54).

- In the largest RCT to date, Collet et al. randomly assigned 111 children with cerebral palsy to 40 treatments over a 2-month period of either HBOT (n=57) or slightly pressurized room air (n=54). The authors found HBOT and slightly pressurized air produced similar improvements in both groups for outcomes such as gross motor function and activities of daily living.

Conclusion

HBOT is considered investigational as a treatment for cerebral palsy because it has not been shown to provide additional health benefits in this patient population.

Compromised Skin Grafts and Flaps

Systematic Reviews

- In a 2010 Cochrane review, Estes found a lack of high quality evidence regarding HBOT in the treatment of skin grafts and flaps. The authors found one randomized controlled trial (RCT) on skin grafts for burn wounds (n=48) which reported significantly higher graft survival with HBOT, and one RCT on flap grafting (n=135) which reported no significant differences in graft survival with HBOT compared with dexamethasone or heparin. However, these data are unreliable due to various methodologic limitations such as biased analysis, omitted data, and small size.

- In 2006, Friedman published a systematic review of literature on use of HBOT for treating skin flaps and grafts. No RCTs were found. The authors identified 2 retrospective case series on use of HBOT for clinically compromised skin grafts and flaps. The series had sample sizes of 65 and 26, respectively; both were published in the 1980s based on treatment provided in the 1970s and 1980s.

Randomized Controlled Trials (RCTs)

No RCTs have been published since the above systematic reviews.

Conclusion

Although the study of HBOT for compromised skin grafts and flaps goes back several decades, the clinical trial data is limited to noncomparative case series and a single
randomized controlled trial. This evidence is insufficient to determine the safety and efficacy of HBOT in the treatment of compromised skin grafts and flaps. Therefore, HBOT is considered investigational for these indications.

**Carbon Monoxide Poisoning**

A 2011 Cochrane review of seven RCTs concluded that the available evidence is insufficient to determine whether adverse neurologic outcomes in patients with carbon monoxide poisoning are reduced with HBOT.[24] In 2008, the American College of Emergency Physicians (ACEP) published a clinical policy on critical issues in carbon monoxide poisoning.[25] Their literature review indicated there was only level C evidence (preliminary, inconclusive, or conflicting evidence) for treatment of acute carbon monoxide poisoning. The 2008 UHMS guidelines, however, list carbon monoxide poisoning as an indication for HBOT.

Two blinded randomized trials were discussed in both the Cochrane and ACEP reviews. One is a study by Scheinkestel, a double-blind, RCT comparing HBOT with normobaric oxygen in patients with carbon monoxide poisoning.[26] The authors reported that HBOT did not benefit patient outcomes of neuropsychologic performance when HBOT was completed and at 1-month follow-up. This study was limited, however, by a high rate (46%) of patients who were lost to follow-up. Moreover, the trial has been criticized for administering 100% normobaric oxygen for at least 72 hours between treatments, which has been called a toxic dose of oxygen.[27] The critiques also mention that there was an unusually high rate of neurologic sequelae after the treatment period, which could be due in part to the high dose of oxygen and/or the high rate of cognitive dysfunction in the study population (69% were poisoned by carbon monoxide through suicide attempts).

The other blinded trial, by Weaver, also compared hyperbaric and normobaric oxygen.[28] Patients received either 3 sessions of HBOT or 1 session of normobaric oxygen plus 2 sessions of exposure to normobaric room air. The primary outcome was the rate of cognitive sequelae at 6 weeks. Cognitive function was assessed using a battery of neuropsychological tests. At the 6-week follow-up, the intention-to-treat analysis found that 19 of 76 (25.0%) in the HBOT group and 35 of 76 (46.1%) in the control group had cognitive sequelae; the difference was statistically significant (p=0.007). There was a high rate of follow-up at 6 weeks, 147 of 152 (97%) of randomized patients. Enrollment in the study was stopped early because an interim analysis found HBOT to be effective. A follow-up study, which included 147 patients from the randomized trial and 75 who had been eligible for the trial but had not enrolled, was published in 2007.[29] Of the group treated with HBOT (n=75), cognitive sequelae were identified in 10 of 58 (17%) at 6 months and 9 of 62 (14%) at 12 months. Of the group not treated with HBOT (n=163), 44 of 146 (30%) at 6 months and 27 of 149 (18%) at 12 months had cognitive sequelae. (The follow-up rate was higher at 12 months because the investigators received additional funding for data collection.)

**Delayed-Onset Muscle Soreness**

**Systematic Review**

In a 2005 Cochrane review, Bennett concluded that available evidence is insufficient to demonstrate beneficial outcomes with HBOT for delayed-onset muscle soreness and closed soft-tissue injury.[30] It was noted that HBOT possibly even increases pain initially and further studies are needed. Therefore, use of HBOT for this indication is considered investigational.

**Randomized Controlled Trials (RCTs)**
No RCTs have been published since the 2005 Cochrane review.

**Dementia**

**Systematic Review**

A 2012 Cochrane review identified 1 RCT evaluating HBOT for the treatment of vascular dementia. The 2009 study compared HBOT plus donepezil to donepezil-only in 64 patients. The HBOT and donepezil group had significantly better cognitive function after 12 weeks of treatment, as assessed by the Mini-Mental State Examination. However, the Cochrane investigators judged the trial to be of poor methodologic quality because it was not blinded and the methods of randomization and allocation concealment were not discussed.

**Randomized Controlled Trials (RCTs)**

No RCTs have been published since the 2012 Cochrane review.

**Conclusion**

The current evidence for HBOT as a treatment of dementias of any cause is limited to a single short-term clinical trial on vascular dementia. This evidence is insufficient to permit conclusions about the safety and efficacy of HBOT on vascular dementia. No other randomized controlled trials were found for HBOT as a treatment of demential from any cause. Due to the lack of sufficient evidence, HBOT is considered investigational for treatment of dementias.

**Femoral Neck Necrosis, Idiopathic**

**Randomized Controlled Trials (RCTs)**

In 2010, Camporesi published the results of a double-blind RCT that evaluated HBOT in 20 adult patients with idiopathic unilateral femoral head necrosis. Patients received 30 treatments over 6 weeks with either HBOT at 2.5 ATA (n=10) or a sham treatment consisting of hyperbaric air (n=10). The mean severity of pain on a 0-to-10 scale was significantly lower in the HBOT group than the control group after 30 sessions (p<0.001) but not after 10 or 20 sessions. (The article did not report exact pain scores). Several range-of-motion outcomes were also reported. At the end of the initial treatment period, extension, abduction and adduction, but not flexion, were significantly greater in the HBOT group compared to the control group. Longer-term comparative data were not available because the control group was offered HBOT at the end of the initial 6-week treatment period.

**Conclusion**

The current evidence is limited to a single, small short-term RCT. Thus, there is insufficient data on which to draw conclusions about the efficacy of HBOT for treating femoral head necrosis, and it is considered investigational for this indication.

**Fibromyalgia**

One quasi-randomized trial and 1 delayed-treatment RCT on HBOT for fibromyalgia were identified. In 2004, a study by Yildiz included 50 patients with fibromyalgia who had ongoing symptoms despite medical and physical therapy. On an alternating basis, patients were assigned to HBOT or a control group. The HBOT consisted of fifteen 90-minute sessions at 2.4
ata (1 session per day, 5 d/wk). The control group breathed room air at 1 ata on the same schedule. Baseline values on the 3 outcomes were similar in the 2 groups. After the course of HBOT treatment, the mean (SD) number of tender points were 6.04 (1.18) in the HBO group and 12.54 (1.10) in the control group. The mean (SD) pain threshold was 1.33 kg (0.12) and 0.84 kg (0.12), respectively, and the mean VAS was 31.54 (8.34) and 55.42 (6.58), respectively. In the study abstract, the authors stated that there were statistically significant differences between the HBO and control groups after 15 therapy sessions, but the table presenting outcomes lacked the notation used to indicate between-group statistical significance. It is not clear whether the control group actually received a sham intervention that would minimize any placebo effect whether or not the control intervention was delivered in a hyperbaric chamber. The authors stated that the study was double-blind but did not specify any details of patient blinding.

In 2015, Efrati published an RCT that included 60 female patients who had fibromyalgia for at least 2 years and were symptomatic. Patients were randomized to an immediate 2 month course of HBOT or delayed HBOT after 2 months. The HBOT protocol was forty 90-minute sessions of 100% oxygen at 2 ata (1 session per day, 5 d/wk). Forty-eight of 60 patients (80%) completed the study and were included in the analysis. After the initial 2 months, outcomes including number of tender points, pain threshold, and quality of life (SF-36) were significantly better in the immediate treatment group compared with the delayed treatment group (which received no specific intervention during this time). After the delayed treatment group had undergone HBOT, outcomes were significantly improved compared with scores prior to HBOT treatment. These findings are consistent with a clinical benefit of HBOT, but also with a placebo effect. A sham-control is needed to confirm the efficacy of HBOT in the treatment of fibromyalgia and other conditions where primary end points are pain and other subjective outcomes.

The above studies were few in number with relatively small sample sizes and had methodological limitations, e.g., quasi-randomization and no or uncertain sham control for a condition with subjective outcomes susceptible to a placebo effect. Moreover, the HBO protocol varied (e.g., 15 HBOT sessions vs 40 HBOT sessions). Thus, the evidence is insufficient to draw conclusions about the impact of HBOT on health outcomes for patients with fibromyalgia.

**Fracture Healing**

**Systematic Review**

In 2012, Bennett published a Cochrane review on HBOT to promote fracture healing and treat non-union fractures. The investigators did not identify any published RCTs on this topic that compared HBOT to no treatment, sham treatment, or another intervention and reported bony union as an outcome.

**Randomized Controlled Trials (RCTs)**

No RCTs have been published since the 2012 Cochrane review.

**Conclusion**

Due to the lack of RCTs, it is not possible to conclude whether the use HBOT to promote fracture healing improves outcomes; therefore, the use of HBOT for this indication is considered investigational.
Headaches

When assessing any treatment focused on pain relief, randomized, placebo-controlled trials are necessary to investigate the extent of any placebo effect and to determine whether any improvement with the treatment exceeds that associated with a placebo.

The following is a summary of the available evidence:

Migraine headaches

- Systematic Review

A 2008 Cochrane review by Bennett identified RCTs that evaluated the effectiveness of systemic hyperbaric oxygen therapy for preventing or treating migraine headache compared to another treatment or a sham control.[36] Five trials with a total of 103 patients were identified that addressed treatment of acute migraine with HBOT. A pooled analysis of 3 trials (total of 43 patients) found a statistically significant increase in the proportion of patients with substantial relief of migraine within 45 minutes of HBOT (relative risk [RR] 5.97, 95% confidence interval [CI] 1.46-24.38, p=0.001). No other pooled analyses were conducted due to variability in the outcomes reported in the trials. The meta-analysis did not report data on treatment effectiveness beyond the immediate post-treatment period, and the methodologic quality of trials was moderate to low, e.g., randomization was not well-described in any trial. There was no evidence that HBOT could prevent episodes of migraine headache.

- Randomized Controlled Trials (RCTs)

In 2004 Eftedal reported the results of a randomized, double-blind, placebo-controlled trial to assess whether HBOT had a prophylactic effect on migraine headache. [37] Forty patients were randomly assigned to either a treatment group receiving 3 sessions of HBOT or a control group receiving 3 hyperbaric treatments with room air. Thirty-four patients completed the study. Efficacy was measured as the difference between pre- and post-treatment hours of headache per week. There was no significant reduction in hours of headache with HBOT compared with hyperbaric air treatments. Nor was there a significant difference in either group in pre- and post-treatment levels of endothelin-1 in venous blood. The authors concluded that HBOT had no significant prophylactic effect on migraine headache or on the endothelin-1 level in venous blood.

Cluster headaches

- Systematic Reviews

Two 2008 systematic reviews, including the Cochrane review noted above, reported few studies comparing HBOT with sham treatment for cluster headaches.[36, 38] Available randomized, placebo-controlled trials measuring effect on symptoms are unreliable due to very small size.[39, 40]

- Randomized Controlled Trials (RCTs)

No RCTs have been published since the 2008 systematic reviews.

- Conclusion
Due to the lack of sufficient evidence from well-designed clinical trial, HBOT for the treatment of headaches from any cause is considered investigational.

**Herpes Zoster**

**Randomized Controlled Trial (RCT)**

In 2012, Peng published an RCT evaluating HBOT as a treatment of herpes zoster.\[^41^\] Sixty-eight patients with herpes zoster diagnosed within the previous 2 weeks were randomized to 30 sessions of HBOT (n=36) or medication treatment (n=32). Pharmacotherapy included antiviral, pain, nerve nutritive and antidepressive medication. Therapeutic efficacy was calculated at the end of the 3-week treatment period and included the proportion of patients who were healed (i.e., complete subsidence of pain and rash) or improved (i.e., significant pain relief and rash subsistence). Rates of therapeutic efficacy were 97.2% in the HBOT group and 81.3% in the medication group (p<0.05). Limitations of the study included a lack of blinding and lack of long-term follow-up.

**Conclusion**

The evidence from the single randomized controlled trial is insufficient to permit conclusions about the effect of HBOT on health outcomes for patients with herpes zoster; therefore, HBOT is considered investigational for this indication.

**Inflammatory Bowel Disease**

**Systematic Reviews**

Singh (2021) published a systematic review on the efficacy of HBOT in patients with ulcerative colitis and Chron’s disease.\[^42^\] A total of 18 studies were included in the review consisting mainly of observational studies. The overall response rate of HBOT in ulcerative colitis was 83.24% (95% CI: 61.90-93.82), while the response in Crohn's disease was 81.89 (95% CI: 76.72-86.11). The results of randomized trials for HBOT as adjuvant therapy in ulcerative colitis were conflicting within the review. The complete healing of fistula in fistulizing Crohn's disease was noted 47.64% (22.05-74.54), while partial healing was noted in 34.29% (17.33-56.50%). This review is limited by inclusion of inadequately powered studies and lack of randomized trials.

McCurdy (2021) published a systematic review evaluating the efficacy of HBOT on various inflammatory bowel disease phenotypes.\[^43^\] There were 19 studies included in the review with 809 patients in three randomized trials and 16 case series. Rates of clinical remission included 87% (95% CI, 10-100) for ulcerative colitis (n = 42), 88% (95% CI, 46-98) for luminal Crohn’s disease (CD, n = 8), 60% (95% CI, 40-76) for perianal CD (n = 102), 31% (95% CI, 16-50) for pouch disorders (n = 60), 92% (95% CI, 38-100) for pyoderma gangrenosum (n = 5), and 65% (95% CI, 10-97) for perianal sinus/metastatic CD. This review is limited by the inclusion of primarily case studies and studies with inadequate descriptions of the interventions and outcomes.

A 2014 systematic review by Dulai examined the evidence on HBOT for inflammatory bowel disease (Crohn disease and ulcerative colitis).\[^44^\] The review was not limited by study design. The authors included 17 studies: 1 RCT, 2 case-control studies, 3 case series, and 11 case reports. The studies reported on a total of 613 patients, 286 with Crohn disease and 327 with ulcerative colitis. The only RCT identified was published in 2013; it was open-label and
included 18 patients with ulcerative colitis. Patients were randomized to treatment with standard medical therapy only (n=8) or medical therapy plus HBOT (n=10) consisting of 90-minute treatments at 2.4 atm, 5 days a week for 6 weeks (total of 30 sessions). The primary outcome was the self-reported Mayo score which has a potential range of 0 to 12. Patients with a score of 6 or more are considered to have moderate to severe active disease. At 6 months follow-up there was no significant difference between groups in the Mayo score, with a median score of 0.5 in the HBOT group and 3 in the control group (exact p value not reported). In addition, there were no significant differences in any of the secondary outcomes including laboratory tests and fecal weight. Overall, the authors found that the studies had a high risk of bias, particularly in the areas of attrition and reporting bias, and further study in well-controlled, blinded RCTs was recommended.

**Randomized Controlled Trials (RCTs)**

No RCTs have been published since the 2014 systematic review.

**Conclusion**

There is insufficient evidence that HBOT is effective for treating inflammatory bowel disease. Only 1 small RCT has been published, and this study did not find a significant improvement in health outcomes when HBOT was added to standard medical therapy.

**In Vitro Fertilization**

In a 2005 nonrandomized pilot study, Van Voorhis reported that HBOT was well tolerated in women undergoing ovarian follicular stimulation for in vitro fertilization; however no outcomes were reported. Therefore, current evidence is insufficient to permit conclusions and HBOT is considered investigational for this indication.

**Mental Illness**

A Rapid Response Report from the Canadian Agency for Drugs and Technologies in Health (CADTH) searched the literature through July 2014 on the clinical effectiveness of hyperbaric oxygen therapy for treatment of adults with posttraumatic stress disorder, generalized anxiety disorder, and/or depression.

The review’s inclusion criteria were health technology assessments, systematic reviews, meta-analyses, RCTs or nonrandomized studies comparing HBOT to any active treatment and reporting clinical outcomes. No eligible studies were identified.

**Multiple Sclerosis**

A Cochrane review of RCTs on HBOT for multiple sclerosis was published by Bennett in 2004. The authors identified 9 RCTs, with a total of 504 participants that compared the effects of HBOT with placebo or no treatment. The primary outcome of the review was score on the Expanded Disability Status Scale (EDSS). A pooled analysis of data from 5 trials (N=271) did not find a significant difference in change in the mean EDSS after 20 HBOT treatments versus control (mean difference [MD], -0.07; 95% CI, -0.23 to 0.09). Moreover, a pooled analysis of data from 3 trials (n=163) comparing HBOT and placebo did not find a significant difference in mean EDSS after 6 months of follow-up (MD = -0.22; 95% CI, -0.54 to 0.09).

**Osteomyelitis**
No prospective clinical trials on chronic refractory osteomyelitis or acute refractory osteomyelitis were identified in updated searches. The justification for the use of HBOT in chronic osteomyelitis has been primarily based on case series. Among the larger case series, Maynor reviewed the records of all patients with chronic osteomyelitis of the tibia seen at one institution.\(^5\) Follow-up data were available on 34 patients who had received a mean of 35 adjunctive HBO treatments (range, 6-99). Of the 26 patients with at least 2 years of follow-up after treatment, 21 (81%) remained drainage-free. Twelve of 15 (80%) with follow-up data at 60 months had remained drainage-free. A study by Davis reviewed outcomes for 38 patients with chronic refractory osteomyelitis treated at another U.S. institution.\(^6\) Patients received HBOT until the bone was fully recovered with healthy vascular tissue; this resulted in a mean of 48 daily treatments (range, 8-103). After a mean posttreatment follow-up of 34 months, 34 of 38 (89%) patients remained clinically free of infection (i.e., drainage-free and no tenderness, pain, or cellulitis). Success rates from several smaller case series, all conducted in Taiwan, are 12 of 13 (92%) patients, 11 of 14 (79%) patients, and 13 of 15 (86%) patients.\(^7\) A high percentage of refractory patients in these series had successful outcomes.

**Radiotherapy Adverse Effects**

**Systematic Review**

- A 2017 systematic review on the effectiveness of HBOT for the treatment of radiation-induced skin necrosis included eight articles with five case series studies, two case reports, and one observational cohort.\(^8\) The authors investigated the change in symptoms and alteration in wound healing and reported that HBOT was a safe intervention with promising outcomes. However, the authors recommended additional high quality evidence in order for HBOT to be considered as a relevant treatment for this indication.

- A 2014 systematic review on the safety and effectiveness of HBOT for the treatment of non-neurological soft tissue radiation-related injuries (STRI) included 41 articles, 11 of which compared regimens with and without HBOT.\(^9\) Serious adverse effects were rare and the more common adverse effects were minor and self-limiting. Evidence of a beneficial effect of HBOT was reported radiation proctitis and STRI of the head and neck, but not for post-radiation soft tissue edema or radiation cystitis. The authors recommended further studies to validate the use of HBOT as both a definitive and adjunctive treatment for individual STRI.

- In 2010, Spiegelberg conducted a systematic review of studies on HBOT to prevent or treat radiotherapy-induced head and neck injuries associated with treatment of malignant tumors.\(^9\) The authors identified 20 studies. Eight of the studies included control groups; their sample sizes ranged from 19 to 78 individuals. Four (50%) of the studies with a control group concluded that HBOT was effective, and the other 4 did not conclude that the HBOT was effective. The authors noted a paucity of RCTs but did not state the number of RCTs identified in their review.

**Randomized Controlled Trials**

- Teguh reported on 17 patients with oropharyngeal or nasopharyngeal cancer who were treated with radiation therapy.\(^9\) Eight patients were randomly assigned to receive 30 sessions of HBOT, beginning within 2 days of completing radiation therapy, and 9 patients received no additional treatment. All patients were included in the analysis.
Quality of life outcomes were assessed and the primary outcome was specified as xerostomia at 1 year. Quality of life measures did not differ significantly between groups in the acute phase (first 3 months). For example, 1 month after treatment, the mean visual analog scale (VAS) score for xerostomia (0-to-10 scale) was 5 in the HBOT group and 6 in the control group. However, at 1 year, there was a statistically significant difference between groups; the mean VAS score for xerostomia was 4 in the HBOT group and 7 in the control group (p=0.002). Also at 1 year, the mean quality of life score for swallowing (0-to-100 scale) was 7 in the HBOT group and 40 in the control group (p=0.0001). The study is limited by the small sample size and the wide fluctuation over the follow-up period in quality-of-life ratings.

- In 2010, Gothard randomized 58 patients with arm lymphedema (at least 15% increase in arm volume) following cancer treatment in a 2:1 ratio to receive HBOT (n=38) or usual care without HBOT (n=20). Fifty-three patients had baseline assessments and 46/58 (79%) had 12-month assessments. No statistically significant difference was found in the change in arm volume from baseline to 12-month follow-up. The median change from baseline was -2.9% in the treatment group and -0.3% in the control group. The study protocol defined response as at least an 8% reduction in arm volume relative to the contralateral arm. According to this definition, 9 of 30 (30%) patients in the HBOT group were considered responders compared with 3 of 16 (19%) in the control group; the difference between groups was not statistically significant. Other outcomes, e.g., quality-of-life scores on the Short-Form (SF)-36, were also similar between groups.

Conclusion

Due to the lack of sufficient evidence from well-designed clinical trial, HBOT for the treatment of adverse effects related to radiation therapy is considered investigational.

Radionecrosis and Osteoradionecrosis

Several systematic reviews of RCTs have been published. A 2008 Cochrane review by Esposito reviewed the use of HBOT in patients requiring dental implants. The authors identified 1 randomized trial involving 26 patients. The authors concluded that despite the limited amount of clinical research available, it appears that HBOT in irradiated patients requiring dental implants may not offer any appreciable clinical benefits. They indicated that there is a need for more RCTs to ascertain the effectiveness of HBOT in irradiated patients requiring dental implants.

In 2012, Bennett published a Cochrane review on HBOT for late radiation tissue injury. The authors identified 11 RCTs; there was variability among trials, and study findings were not pooled for the primary outcomes of survival, complete resolution of necrosis or tissue damage, and improvement in a late effects symptom scale. In a pooled analysis of 3 studies, a significantly higher proportion of patients with osteoradionecrosis achieved complete mucosal cover after HBOT compared with controls (RR=1.30; 95% CI, 1.09 to 1.55). From their review of the literature, the authors concluded that data from small trials “suggest that for people with LRTI [late radiation tissue injury] affecting the head, neck, anus, and rectum, [HBOT] is associated with improved outcome. HBOT also appears to reduce the chance of ORN [osteoradionecrosis] following tooth extraction in an irradiated field. There was no such evidence of any important clinical effect on neurological tissues. The application of HBOT to selected patients and tissues may be justified.”
Stroke

Current evidence is insufficient to permit conclusions about whether HBOT improves health outcomes in the treatment of stroke or stroke-related functional limitations. The following is a summary of the available evidence:

Acute Stroke

- Systematic Reviews
  - In a 2005 Cochrane systematic review, Bennett evaluated HBOT for acute stroke.[61] The investigators identified 6 RCTs with a total of 283 participants that compared HBOT to sham HBOT or no treatment. The authors were only able to pool study findings for 1 outcome, the mortality rate at 3-6 months. A pooled analysis of 3 trials found no significant benefit of HBOT compared to the control for this outcome. Based on the available evidence, acute ischemic stroke is considered investigational
  - In a 2005 systematic review, Carson concluded that current evidence did not demonstrate any benefit with the use of HBOT for the treatment of stroke.[62] The authors noted it was undetermined whether there were any benefits with HBOT that would outweigh potential harms, and further study was required.
  - In a 2014 update of a Cochrane systematic review, Bennett evaluated HBOT for acute ischemic stroke. The investigators identified 11 RCTs with a total of 705 participants that compared HBOT with sham HBOT or no treatment. The authors were only able to pool study findings for 1 outcome; mortality at 3 to 6 months. A pooled analysis of data from 4 trials with a total of 106 participants did not find a significant benefit of HBOT compared with a control condition for this outcome (RR=0.97; 95% CI, 0.34 to 2.75).

- Randomized Controlled Trials (RCTs)
  No RCTs have been published since the 2005 systematic reviews.

Stroke-related motor dysfunction

- Randomized Controlled Trials (RCTs)

  In 2013, Efrati published an RCT evaluating HBOT for treatment of neurologic deficiencies associated with a history of stroke.[63] The study included 74 patients with at least one motor dysfunction who had an ischemic or hemorrhagic stroke 6-36 months prior to study participation. Participants were randomly assigned to receive 2 months of HBOT (40 daily sessions, 5 days per week, n=30) or delayed treatment (n=32). Patients were evaluated at baseline and 2 months. For patients in the delayed treatment control group, outcomes were evaluated at 4 months after crossing over and receiving HBOT. Twenty-nine of 32 patients (91%) in the delayed treatment group crossed over to the active intervention. Outcome measures included the National Institutes of Health Stroke Scale (NIHSS), which was measured by physicians blinded to treatment group, and several patient-reported quality-of-life and functional status measures.

  At 2 months’ follow-up, there was statistically significantly greater improvement in function in the HBOT group compared to the control group as measured by the NIHSS, quality-of-life scales and the ability to perform activities of daily living (ADLs). These
differences in outcome measures were accompanied by improvements in single-photon emission computed tomography (SPECT) imaging in the regions affected by stroke. For the delayed treatment control group, there was a statistically significant improvement in function after HBOT compared to before treatment. This RCT raises the possibility that HBOT may induce improvements in function and quality of life for post-stroke patients with motor deficits. However, the results are not definitive for a number of reasons. This RCT is small and enrolled a heterogeneous group of post-stroke patients. The study was not double-blind and the majority of outcome measures, except for the NIHSS, were patient reported and thus prone to the placebo effect. Also, there was a high total dropout rate of 20% at the 2-month follow-up point. Therefore, larger, double-blind studies with longer follow-up are needed to corroborate these results. Because of these limitations in the evidence, HBOT is considered investigational for treating motor dysfunction associated with stroke.

**Traumatic Brain Injury**

**Systematic Review**

The systematic review and pooled analysis by Hart (2019) evaluated HBOT for mild traumatic brain injury (mTBI) associated post-concussive symptoms (PCS) and posttraumatic stress disorder (PTSD). Data were aggregated from four Department of Defense (DoD) studies that included participant level data on 254 patients assigned to either HBOT or sham intervention. An additional three studies with summary-level participant data were summarized (N=135). The authors assessed changes from baseline to post-intervention on PCS, PTSD, and neuropsychological measures. The DoD data analyses indicated improvements with HBOT for PCS, measured by the Rivermead Total Score. Statistically significant improvements were seen for PTSD based on the PTSD Checklist Total Score, as well as for verbal memory based on CVLT-II Trial 1-5 Free Recall.

A 2016 meta-analysis by Wang (2016) assessed HBOT for TBI including eight studies with 519 participants that met the eligibility criteria. HBOT protocols varied across studies in the levels of oxygen and the length and frequency of treatments. The primary outcome was change in the Glasgow Coma Scale score. A pooled analysis of two studies found a significantly greater improvement in the mean Glasgow Coma Scale score in the HBOT group compared with control groups. Mortality (a secondary outcome) was reported in 3 of the 8 studies. Pooled analysis of these 3 studies found a significantly lower overall mortality rate in the HBOT group than in the control group.

A 2012 Cochrane systematic review addressed HBOT as adjunctive treatment for traumatic brain injury. The investigators identified 7 RCTs with a total of 571 participants comparing a standard intensive treatment regimen to the same treatment regimen with the addition of HBOT. The review did not include studies in which interventions occurred in a specialized acute care setting. The HBOT regimens varied among studies; for example, the total number of individual sessions varied from 3 to 30-40. No trial used sham treatment or blinded the staff members who were treating the patients, and only 1 had blinding of outcome assessment. Allocation concealment was inadequate in all of the studies. The primary outcomes of the review were mortality and functional outcomes. A pooled analysis of data from 4 trials that reported this outcome found a statistically significantly greater reduction in mortality when HBOT was added to a standard treatment regimen. However, when data from the 4 trials were pooled, the difference in the proportion of patients with an unfavorable functional outcome at final follow-up did not reach statistical significance. Unfavorable outcome was
commonly defined as a Glasgow Outcome Score (GOS) of 1, 2 or 3, which are described as ‘dead’, ‘vegetative state’ or ‘severely disabled’. Studies were generally small and were judged to have substantial risk of bias.

Randomized Controlled Trials

A 2012 sham-controlled double-blind trial evaluating HBOT was published after the 2012 Cochrane review. The study included 50 military service members, 48 of whom were male, with combat-related mild traumatic brain injury. Participants were randomized to 30 sessions of HBOT over 8 weeks (n=25) or a sham intervention (room air at 1.3 ATA) (n=25). The primary outcome measures were scores on the Immediate Post-Concussive Assessment and Cognitive Testing (ImPACT) and Post-Traumatic Disorder Check List–Military Version (PCL-M) instruments. Patients were evaluated after every 5 treatment sessions and at 6 weeks post-exposure. Forty-eight of 50 participants (96%) completed the study. There were no statistically significant differences on the ImPACT total mean score or the PCL-M composite score at any time point. While the sample size was relatively small, the study was powered to detect clinically significant differences among groups on the cognitive tests.

Several trials on mild traumatic brain injury in military populations have been published and these did not find significant benefits of HBOT compared with sham treatment. The first trial, published by Wolf in 2012, included 50 military service members, 48 of whom were male, with combat-related mild traumatic brain injury. Participants were randomized to 30 sessions of HBOT over 8 weeks (n=25) or a sham intervention (room air at 1.3 atmosphere, absolute [ata]) (n=25). The primary outcome measures were scores on the Immediate Post-Concussive Assessment and Cognitive Testing (ImPACT) and Post-Traumatic Disorder Check List–Military Version (PCL-M) instruments. Patients were evaluated after every 5 treatment sessions and at 6 weeks post-exposure. Forty-eight of 50 participants (96%) completed the study. There were no statistically significant differences on the ImPACT total mean score or the PCL-M composite score at any time point. For example, at the 6-week follow-up, mean composite PCL-M scores were 41.6 in the HBOT group and 40.6 in the sham-control group (p=0.28). While the sample size was relatively small, the study was powered to detect clinically significant differences among groups on the cognitive tests.

A 2014 double-blind sham-controlled trial 2014 RCT by Cifu included 61 male Marines who had a history of mild traumatic brain injury and postconcussive syndrome. To maintain blinding, all patients were pressured inside a hyperbaric chamber to 2.0 ata. They were randomized to breathe 1 of 3 oxygen [nitrogen gas mixes equivalent to: (1) 75% oxygen at 1.5 ata (n=21); (2) 100% oxygen at 2.0 ata (n=19); and (3) sham treatment with surface room air (n=21). Patients underwent 40 once daily 60-minute sessions. Outcomes were assessed 3 months after the last exposure. The primary outcome was a clinically meaningful improvement, defined as a 10% difference between groups in the score on the Rivermead Post-Concussion Questionnaire (RPQ)–16 (scale range, 50-84; higher values indicate more severe symptoms). At follow-up, there was no statistically significant difference among groups on RPQ-16 score (p=0.41). A variety of secondary outcomes were also assessed. None of these, including measures of attention, cognition, or depression, differed significantly among groups at follow-up.

Also in 2014, Miller evaluated HBOT in 72 military service members with continuing symptoms at least 4 months after mild traumatic brain injury. Patients were randomized to receive 40 daily HBO sessions at 1.5 ata, 40 sham sessions consisting of room air at 1.2 ata or standard care with no hyperbaric chamber sessions. The primary outcome was change in
the RPQ. A cutoff of 15% improvement was deemed clinically important, which translates to a change score of at least 2 points on the RPQ-3 subscale. The proportion of patients who met the prespecified change of at least 2 points on the RPQ-3 was 52% in the HBOT group, 33% in the sham group and 25% in the standard care-only group. The difference between rates in the HBOT and sham groups was not statistically significant (p=0.24). None of the secondary outcomes significantly favored the HBOT group. A criticism of this study, as well as the other military population studies, was that the response in the sham group was not due to a placebo effect but to an intervention effect of slightly increased atmospheric pressure (1.2 ata).43 Other researchers have noted that room air delivered at 1.2 ata would not be considered an acceptable therapeutic dose for any indication, and especially for a condition with persistent symptoms like postconcussive syndrome.

Conclusion

A systematic review of small trials with limitations found a mortality reduction with HBOT but no significant improvement in patient function among survivors of traumatic brain injury. Two double-blind, sham-controlled RCTs of HBO treatment in a military population with mild traumatic brain injury did not find a statistically significant benefit with HBOT. Thus, the evidence is insufficient that HBOT improves health outcomes in patients with traumatic brain injury, and this indication is considered investigational.

Wounds Unrelated to Diabetes

Acute Surgical and Traumatic Wounds

- Systematic Reviews
  - A 2013 updated Cochrane review analyzed randomized controlled trials comparing either HBOT with a different intervention, or two HBOT regimens for acute wounds (e.g., surgical wounds, lacerations, traumatic wounds, and animal bites).[68] The four studies that met inclusion criteria ranged in size from 10 to 135 subjects. Reported outcomes were mixed. Meta-analysis of pooled data was not possible due to differences among studies with respect to patient characteristics, interventions studied, and outcome measures. Also identified was a high risk of bias due to insufficient disclosure of randomization methods and selective reporting of outcome data. Findings of individual studies were mixed.

The primary outcome examined by Cochrane reviewers, wound healing was not reported in either of the 2 trials comparing HBOT with usual care[69, 70] or in the 1 trial comparing HBOT with dexamethasone or heparin.[71] Complete wound healing was reported in the 1 RCT comparing active HBOT with sham HBOT.[72] In this small study (n=36), there was a statistically higher rate of wound healing in the active HBOT group. The time point for outcome measurement in this study was unclear, but there was no statistically significant difference between groups in the meantime to wound healing. Adverse effects included 2 additional surgical procedures in 1 patient in the HBOT group compared with 8 in 6 patients in the sham group. The HBOT group had significantly fewer patients who developed necrotic tissue (1 and 8, respectively). There were no amputations in the HBOT group compared with 2 amputations in the sham group, but this difference did not reach statistical significance. The authors concluded that evidence remains insufficient to support the routine use of HBOT for acute surgical or traumatic wounds. They recommended
further evaluation in high quality RCTs that include outcomes measures of complete wound closure and accelerated wound closure.

- In 2014 Dauwe published a systematic review that included 8 studies with sample sizes ranging from 5 to 125 patients. Four studies were randomized, 3 were prospective non-RCTs, and 1 was a retrospective non-RCT. Data were not pooled due to the heterogeneity described below. The authors noted that 7 of the 8 studies reported achieving statistical significance in their primary end points, but the end points differed among studies (eg, graft survival, length of hospital stay, wound size). Moreover, the studies were heterogeneous in terms of treatment regimens, patient indications (eg, burns, face lifts), and study designs, making it difficult to draw conclusions about the effect of HBOT on acute wound treatment.

- Randomized Controlled Trials (RCTs)

No RCTs have been published since those included in the systematic reviews summarized above.

Chronic Non-diabetic Wounds

- Systematic Reviews

Several systematic reviews of RCTs have been published. An updated 2007 Cochrane review of randomized controlled trials (RCTs) on HBOT for chronic wounds was published by Kranke in 2012.[73] The authors identified 9 RCTs with a total of 471 participants that compared the effect of HBOT on chronic wound healing compared with an alternative treatment approach that did not use HBOT. Eight of the 9 trials included in the review evaluated HBOT in patients with diabetes. The remaining trial addressed HBOT for patients with venous ulcers; that study had only 16 participants and the comparator treatment was not specified. In a pooled analysis of data from 3 trials, a significantly higher proportion of ulcers had healed at the end of the treatment period (6 weeks) in the group receiving HBOT compared to the group not receiving HBOT (RR: 5.20; 95% CI: 1.25 to 21.7). Pooled analyses, however, did not find significant differences between groups in the proportion of ulcers healed in the HBOT versus non-HBO-treated groups at 6 months (2 trials) or 12 months (3 trials). There were insufficient data to conduct pooled analyses of studies evaluating HBOT for treating patients with chronic wounds who did not have diabetes.

In 2013, O'Reilly[74] published a systematic review of studies on HBOT for treatment of diabetic ulcers. The authors identified 6 RCTs and 6 non-RCTs that compared HBOT with standard wound care or sham therapy in patients with diabetes who had nonhealing lower-limb ulcers. Pooled analyses of observational studies found statistically significant benefits of HBOT on rates of major amputation, minor amputation and the proportion of wounds healed at the end of the study period. However, in pooled analyses of RCT data, the stronger study design, there were no statistically significant differences between groups on key outcomes. This included the rate of major amputation (RR=0.40; 95% CI, 0.07 to 2.23; p=0.29), minor amputation (RR=0.79; 95% CI, 0.19 to 3.30, p=0.75), and the proportion of unhealed wounds at the end of the study period (RR=0.54, 95% CI, 0.26 to 1.13, p=0.1).
Systematic reviews have had mixed findings on the impact of HBOT on diabetic ulcers. A Cochrane review found short-term, but not long-term benefit on wound healing, and a 2013 meta-analysis did not find significant benefits of HBOT on outcomes in RCTs, but did find an effect in non-RCTs. There is insufficient evidence on HBOT for treatment of chronic wounds in patients without diabetes.

- Randomized Controlled Trials (RCTs)

No RCTs have been published since those included in the systematic reviews summarized above.

- Conclusion

Published clinical trial data is insufficient to determine the effectiveness of HBOT for wounds that are not related to diabetes. The UHMS does not include these wounds in their list of indications for HBOT, noting the lack of available evidence. As shown in studies of adjunctive HBOT for treatment of severe diabetic lower extremity ulcers, this treatment is well suited to randomized, controlled comparative trials. In spite of this, only 1 small (n=16) randomized, controlled trial was found for non-diabetic wounds. This trial is too small and short-term to be reliable.

Other Indications

No data from well-designed randomized, controlled clinical trials were found that supported HBOT for any other investigational indication, including but not limited to refractory mycoses and acute peripheral arterial insufficiency.

Other indications

For the indications listed below, insufficient evidence to support the use of HBOT was identified. Since 2000, there have been no published controlled trials or large case series (i.e., >25 patients):

- bone grafts;
- carbon tetrachloride poisoning, acute;
- cerebrovascular disease, acute (thrombotic or embolic) or chronic;
- fracture healing;
- hydrogen sulfide poisoning;
- intra-abdominal and intracranial abscesses;
- lepromatous leprosy;
- meningitis;
- pseudomembranous colitis (antimicrobial agent-induced colitis);
- radiation myelitis;
- sickle cell crisis and/or hematuria;
- amyotrophic lateral sclerosis;
- retinopathy, adjunct to scleral buckling procedures in patients with sickle cell peripheral retinopathy and retinal detachment;
- pyoderma gangrenosum;
• tumor sensitization for cancer treatments, including but not limited to, radiotherapy or chemotherapy;

SUMMARY OF EVIDENCE

There is sufficient published evidence to determine that use of hyperbaric oxygen therapy (HBOT) in selected patients with nonhealing diabetic wounds of the lower extremities, acute traumatic ischemia, soft-tissue radiation necrosis (eg, radiation enteritis, cystitis, proctitis), osteoradionecrosis (ie, pre- and posttreatment) for patients undergoing dental surgery (non-implant-related) of an irradiated jaw, gas gangrene, idiopathic sudden sensorineural hearing loss, and profound anemia with exceptional blood loss when blood transfusion is impossible or must be delayed improves the net health outcome. There is insufficient evidence for patients all other indications included in the Rationale section that HBOT improves the net health outcome.

PRACTICE GUIDELINE SUMMARY

U.S. FOOD AND DRUG ADMINISTRATION (FDA)

In 2013, the FDA published a position statement with a warning that HBOT has not been proven safe and effective for uses not cleared by the agency.[1] This statement was developed due to numerous complaints from consumers and health care professionals that unproven claims made by some HBOT centers may mislead consumers and ultimately endanger their health. The statement included the following conditions for which patients may be unaware that safety and effectiveness of HBOT have not been established:

• AIDS/HIV
• Alzheimer's Disease
• Asthma
• Bell's Palsy
• Brain Injury
• Cerebral Palsy
• Depression
• Heart Disease
• Hepatitis
• Migraine
• Multiple Sclerosis
• Parkinson's Disease
• Spinal Cord Injury
• Sport's Injury
• Stroke

UNDERSEA AND HYPERBARIC MEDICAL SOCIETY (UHMS)

In 2015, the Undersea and Hyperbaric Medical Society (UHMS) published a guideline on the use of HBOT for treatment diabetic foot ulcers.[77, 78] Recommendations are as follows:

• Suggest against using HBOT in patients with Wagner Grade 2 or lower diabetic foot ulcers
• Suggest adding HBOT in patients with Wagner Grade 3 or higher diabetic foot ulcers that have now shown significant improvement after 30 days of standard of care therapy
• Suggest adding acute post-operative HBOT to the standard of care in patients with Wagner Grade 3 or higher diabetic foot ulcers who have just had foot surgery related to their diabetic ulcers.

• Appropriate Indications for HBOT[79]

In 2014, the UHMS updated their list of indications considered appropriate for hyperbaric oxygen therapy. These indications are as follows:

- Acute thermal burn injury
- Air or gas embolism
- Arterial insufficiencies (central retinal artery occlusion; enhancement of healing in selected problem wounds)
- Carbon monoxide poisoning and carbon monoxide poisoning complicated by cyanide poisoning
- Clostridial myositis and myonecrosis (gas gangrene)
- Compromised grafts and flaps
- Crush injury, compartment syndrome, and other acute traumatic ischemias
- Decompression sickness
- Delayed radiation injury (soft tissue and bony necrosis)
- Intracranial abscess
- Idiopathic sudden sensorineural hearing loss (ISSNHL) (patients with moderate to profound ISSNHL who present within 14 days of symptom onset)
- Necrotizing soft tissue infections
- Osteomyelitis (refractory)
- Severe anemia

• Autism Spectrum Disorder (ASD)[13]

The 2009 UHMS position paper included a critical appraisal of the available literature, in particular the 2009 Rossignol RCT[11] which was the only RCT available at that time. The paper concluded that “the UHMS cannot recommend the routine treatment of ASD with HBO2T outside appropriate comparative research protocols.”

• Chronic Brain Injury[80]

The most recent UHMS position statement on chronic brain injury (e.g., traumatic brain injury, cerebral palsy, stroke) is from 2003. The statement considered the evidence to be insufficient to support a recommendation for HBOT for the chronic sequelae of traumatic or non-traumatic brain injury, but noted that continued monitoring of data is warranted.

• Idiopathic Sudden Sensorineural Hearing Loss (ISSNHL)[81]

In October 2011, the UHMS Executive Board approved ISSNHL as an additional indication. According to treatment guidelines, patients with moderate to profound ISSNHL who present within 14 days of symptom onset should be considered for HBOT treatment.

• Multiple Sclerosis[49]

A 2010 UHMS position paper reported that most RCTs have failed to show clinical benefit for HBOT therapy for multiple sclerosis. “We conclude that, while there is some case for further investigation of possible therapeutic effects in selected sub-groups of patients (well-characterized and preferably early in the disease course) and for the response to prolonged
courses of HBOT, this case is not strong. At this time, the UHMS cannot recommend the routine treatment of MS with HBOT outside appropriate comparative research protocols.”

- Topical Oxygen for Chronic Wounds [82]

A 2005 UHMS position statement reported that, “to date, mechanisms of action whereby topical oxygen might be effective have not been defined or substantiated. Conversely, cellular toxicities due to extended courses of topical oxygen have been reported, although, again these data are not conclusive, and no mechanism for toxicity has been examined scientifically...The only randomized trial for topical oxygen in diabetic foot ulcers actually showed a tendency toward impaired wound healing in the topical oxygen group. Contentions that topical oxygen is superior to hyperbaric oxygen are not proven.” Therefore, the UHMS recommends against application of topical oxygen outside a clinical trial setting, noting that topical oxygen “should be subjected to the same intense scientific scrutiny to which systemic hyperbaric oxygen has been held.”

NATIONAL BOARD OF DIVING & HYPERBARIC MEDICAL TECHNOLOGY[7]

As noted above, the current position statement concluded that “the installation and provision of in-home hyperbaric oxygen therapy is inherently unsafe and cannot be condoned.” This position is based on concern for the safety and well-being of patients as well as those people in proximity to the HBOT delivery system because in-home provision of HBOT is likely to:

1. Bypass otherwise mandatory federal, state, and local codes related to design, construction, installation, and operation of these devices; and
2. Occur without adequate physician oversight and the operational support of appropriately qualified HBOT providers.

AMERICAN ACADEMY OF OTOLARYNGOLOGY-HEAD AND NECK SURGERY (AAO-HNS)[83]

In 2012, the AAO-HNS published an evidence-based clinical practice guideline on treatment of sudden hearing loss. The guideline includes a statement that HBOT may be considered a treatment option for patients who present within 3 months of a diagnosis of idiopathic sudden sensorineural hearing loss. The document states, “Although HBOT is not widely available in the United States and is not recognized by many U.S. clinicians as an intervention for ISSNHL, the panel felt that the level of evidence for hearing improvement, albeit modest and imprecise, was sufficient to promote greater awareness of HBOT as an intervention for [this condition]” The strength of the recommendation was rated “Option” defined in this case as based on systematic reviews of RCTs with a balance between benefit and harm.

SUMMARY

Systemic hyperbaric oxygen therapy (HBOT) has been studied for a wide variety of clinical indications. There is enough evidence to show that systemic HBOT is safe and effective for a variety of indications. There are guidelines based on research that recommend the use of systemic HBOT for a variety of indications. Therefore, the use of systemic HBOT may be considered medically necessary when policy criteria are met.

Due to insufficient positive health outcomes for certain patients with non-healing diabetic wounds of the lower extremities, the use of hyperbaric oxygen therapy is considered not
medically necessary when criteria for non-healing diabetic wounds of the lower extremities are not met.

There is not enough evidence to permit conclusions concerning the effects of systemic hyperbaric oxygen therapy (HBOT) on final health outcomes for any other indication. Therefore, the use of systemic hyperbaric oxygen therapy for all other indications is investigational.

There is not enough evidence to permit conclusions concerning the effects of topical hyperbaric and topical normobaric oxygen therapies on health outcomes. Therefore, the use of topical hyperbaric and topical normobaric oxygen therapies for any indication is investigational.

REFERENCES


These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


48. CAfDaTiH (CADTH). Hyperbaric Oxygen Therapy for Adults with Mental Illness: A Review of the Clinical Effectiveness. 2014. PMID:


**CODES**

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>99183</td>
<td>Physician or other qualified health care professional attendance and supervision of hyperbaric oxygen therapy, per session</td>
</tr>
<tr>
<td></td>
<td>99199</td>
<td>Unlisted special service, procedure or report</td>
</tr>
<tr>
<td>HCPCS</td>
<td>A4575</td>
<td>Topical hyperbaric oxygen chamber, disposable</td>
</tr>
<tr>
<td></td>
<td>E0446</td>
<td>Topical oxygen delivery system, not otherwise specified, includes all supplies and accessories</td>
</tr>
<tr>
<td></td>
<td>E1399</td>
<td>Durable medical equipment, miscellaneous</td>
</tr>
<tr>
<td></td>
<td>G0277</td>
<td>Hyperbaric oxygen under pressure, full body chamber, per 30 minute interval</td>
</tr>
</tbody>
</table>

**Date of Origin:** January 1996
Progenitor Cell Therapy for the Treatment of Damaged Myocardium Due to Ischemia

Effective: December 1, 2021

Next Review: October 2022
Last Review: October 2021

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Progenitor cell therapy describes the use of multipotent cells of various cell lineages (autologous or allogeneic) for tissue repair and/or regeneration. Progenitor cell therapy is being investigated for the treatment of damaged myocardium resulting from acute or chronic cardiac ischemia.

MEDICAL POLICY CRITERIA

I. Progenitor cell therapy, including but not limited to skeletal myoblasts or hematopoietic stem cells, is considered **investigational** as a treatment of damaged myocardium.

II. Infusion of growth factors (i.e., granulocyte colony stimulating factor [GCSF]) is considered **investigational** as a technique to increase the numbers of circulating hematopoietic stem cells as treatment of damaged myocardium.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Stem-cell Therapy for Peripheral Arterial Disease, Medicine, Policy No. 141
2. Orthopedic Applications of Stem Cell Therapy, Including Bone Substitutes Used with Autologous Bone Marrow, Medicine, Policy No. 142

BACKGROUND

Ischemia is the most common cause of cardiovascular disease and myocardial damage in the developed world. Despite impressive advances in treatment, ischemic heart disease is still associated with high morbidity and mortality. Current treatments for ischemic heart disease seek to revascularize occluded arteries, optimize pump function, and prevent future myocardial damage. However, current treatments are not able to reverse existing damage to heart muscle.[1, 2] Treatment with progenitor cells (i.e., stem cells) offers potential benefits beyond those of standard medical care, including the potential for repair and/or regeneration of damaged myocardium.

Various types of autologous cell transplantation have been researched as a technique to either stimulate regeneration of the myocardium or modify ventricular remodeling after infarct. The ideal donor cell is uncertain, and there are scientific as well as ethical concerns involved in choosing the ideal source of donor cells.[1] The range of potential sources of donor cells includes embryonic stem cells, adult stem cell, fetal myocytes, and adult blood progenitor cells. The potential sources of embryonic and adult donor cells include skeletal myoblasts, bone marrow cells, circulating blood-derived progenitor cells, endometrial mesenchymal stem cells (MSCs), adult testis pluripotent stem cells, mesothelial cells, adipose-derived stromal cells, embryonic cells, induced pluripotent stem cells, and bone marrow MSCs, all of which are able to differentiate into cardiomyocytes and vascular endothelial cells.

The mechanism of benefit following treatment with progenitor cells is not entirely understood.[2, 3] Differentiation of progenitor cells into mature myocytes and engraftment of progenitor cells into areas of damaged myocardium has been suggested in animal studies using tagged progenitor cells.[3, 4] However, there is controversy concerning whether injected progenitor cells actually engraft and differentiate into mature myocytes in humans to a degree that might result in clinical benefit.[2]

Other mechanisms of benefit have been hypothesized. Progenitor cells may improve perfusion to areas of ischemic myocardium.[5] Basic science research also suggests that injected stem cells secrete cytokines with antiapoptotic and pro-angiogenesis properties.[5, 6] Clinical benefit may result if these paracrine factors are successful at limiting cell death from ischemia or stimulating recovery. For example, myocardial protection can occur through modulation of inflammatory and fibrogenic process. Alternatively, paracrine factors might affect intrinsic repair mechanisms of the heart through neovascularization, cardiac metabolism and contractility, increase in cardiomyocyte proliferation, or activation of resident stem and progenitor cells. The relative importance of these proposed paracrine actions will depend on the age of the infarct, e.g., cytoprotective effects with acute ischemia versus cell proliferation with chronic ischemia. Investigation of the specific factors that are induced by administration of progenitor cells is ongoing.[3, 5, 7]

There is a variety of potential delivery mechanisms for donor cells, encompassing a wide range of invasiveness. Donor cells can be delivered via thoracotomy and direct injection into areas of damaged myocardium.[4, 8] Injection of progenitor cells into the coronary circulation can also be done using percutaneous, catheter-based techniques. Finally, progenitor cells can be delivered intravenously via a peripheral vein. With this approach, the cells must be able to target damaged myocardium and concentrate at the site of myocardial damage.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Adverse effects of treatment with progenitor cells include the risk of the delivery procedure (e.g., thoracotomy, percutaneous catheter-based, etc.) and the risks of the donor cells themselves. Donor progenitor cells can differentiate into fibroblasts rather than myocytes. This may create a substrate for malignant ventricular arrhythmias. There is also a theoretical risk that tumors, such as teratomas, can arise from progenitor cells, but the actual risk of this occurring in humans is not known at present.

**REGULATORY STATUS**

U.S. Food and Drug Administration (FDA) approval is not required in situations in which autologous cells are processed on site with existing laboratory procedures and injected with existing catheter devices. However, there are several products that require FDA approval.

- **MyoCell®** consists of patient autologous skeletal myoblasts that are expanded ex vivo and supplied as a cell suspension in a buffered salt solution for injection into the area of damaged myocardium. MyoCell SDF-1 (BioHeart Inc.) is similar to MyoCell®, but before injection, myoblast cells are genetically modified to release excess stromal-derived factor (SDF)-1. Increased SDF-1 levels at the site of myocardial damage may accelerate recruitment of native stem cells to increase tissue repair and neovascularization. For products, the myoblast isolation and expansion occurs at a single reference laboratory (BioHeart) and are, therefore, subject to FDA approval. Currently neither product is FDA-cleared. In addition, implantation may require the use of a unique catheter delivery system (MyoCath™) that is FDA-cleared.

- An allogeneic human mesenchymal stem cell (hMSC) product (Prochymal®) is being developed by Osiris Therapeutics, Inc. for treatment of acute myocardial infarction. Prochymal (also referred to as Provaceel) is a highly purified preparation of ex vivo cultured adult hMSCs isolated from the bone marrow of healthy young adult donors. Prochymal has been granted “fast track” status by the FDA for Crohn’s disease and graft-versus-host disease (GvHD), and has orphan drug status for GvHD from the FDA and the European Medicines Agency. Prochymal is being studied in Phase II trials for the treatment of acute myocardial infarction, pulmonary disease, and type 1 diabetes.

- **MultiStem®** (Athersys) is an allogeneic bone marrow-derived adherent adult stem-cell product. MultiStem has received orphan drug status from the FDA for GVHD and has received authorization from the FDA for a Phase II trial for treatment of acute myocardial infarction with an adventitial delivery system.

- **Ixmyelocel-T** is a patient-autologous, multicellular treatment containing selectively expanded mesenchymal cells, monocytes and alternatively activated macrophages from bone marrow (Varicel Corporation). Ixmyelocel-T received orphan drug status from the FDA in 2007 for the treatment of cardiomyopathy.

- **CardiAMP™** Cell Therapy system consists of a proprietary assay to identify patients with a high probability to respond to autologous cell therapy, a proprietary cell processing system to isolate process and concentrate the stem cells from a bone marrow harvest at the point of care, and a proprietary delivery system to percutaneously inject the autologous cells into the myocardium. BioCardia has received an investigational device exemption from FDA to perform a trial of CardiAMP™.
Autologous progenitor cell transplantation for the treatment of damaged myocardium is a rapidly evolving field, with a number of areas of substantial uncertainty.[1-3, 14]

- The mechanism of benefit is not well understood.
- Patient selection criteria are still evolving, and the current studies have been performed in highly selected populations.
- There is a lack of standardization in treatment protocols, with uncertainty in cell type and in the optimal methods for harvesting of donor cells, the timing of the transplantation, and the optimal delivery mode (directly into myocardium, intracoronary artery or sinus, or intravenously).
- Strategies to enhance cell engraftment and prolong cell survival are lacking.

The most clinically relevant outcome of any treatment of acute or chronic ischemic myocardial damage is improvement of symptoms, exercise tolerance, and quality of life, and reduction of future myocardial damage and mortality. Evaluating the safety and efficacy of progenitor cell therapy requires randomized comparisons with conventional medical treatments. These comparisons are necessary to determine whether any benefits of progenitor cell therapy outweigh any risks and whether the therapy offers advantages over conventional medical treatment.

ACUTE ISCHEMIA

Systematic Reviews

Fisher (2016) published a trial sequential analyses of two Cochrane reviews to address limitations associated with meta-analyses. The trial sequential analyses were conducted on two clinical outcomes using cell therapy, all-cause mortality and hospitalization for heart failure as well as left ventricular ejection fraction. The results of this analysis suggested that there is evidence of reduced risk of mortality and hospitalization in heart failure, but insufficient to determine if there was a treatment effect in acute ischemia. The cell therapy did not improve left ventricular ejection fraction by more than a mean difference of 4% in patients.

A 2012 Cochrane review included 33 RCTs (39 comparisons with 1,765 participants) on bone marrow-derived stem-cell (BMC) therapy for acute MI (AMI).[15] Twenty-five trials compared stem/progenitor cell therapy with no intervention, and 14 trials compared the active intervention with placebo. There was a high degree of statistical and clinical heterogeneity in the included trials, including variability in the cell dose, delivery and composition. Overall, stem-cell therapy was found to improve left-ventricular ejection fraction (LVEF) in both the short-term (<12 months, weighted mean difference of 2.9 percentage points, 95% confidence interval [CI], 2.0 to 3.7, I²=73%) and long-term (12 to 61 months, weighted mean difference of 3.8 percentage points, 95% CI 2.6 to 4.9, I²=72%). Stem-cell treatment reduced left-ventricular end systolic and end-diastolic volumes at certain times and reduced infarct size in long-term follow-up. There were positive correlations between mononuclear cell dose infused and the effect on LVEF and between the timing of stem-cell treatment and the effect on LVEF. Although the quality of evidence on LVEF was rated as high, the clinical significance of the change in LVEF is unclear. The quality of evidence on health outcomes was rated as moderate. Stem/progenitor cell treatment was not associated with statistically significant changes in the incidence of mortality or morbidity (re-infarction, arrhythmias, hospital re-admission, restenosis, and target vessel revascularization), although the studies may have been underpowered to
detect differences in clinical outcomes. Due to variability in outcomes measured, it was not possible to combine data on health-related quality of life or performance status.

Fisher (2015) published an updated Cochrane review assessing the safety and efficacy of stem-cell therapy for AMI.[16] Literature was searched through March 2015, and 41 RCTs with a total of 2,732 participants (1,564 cell therapy and 1,168 controls) were included.[16-24] There was a low degree of statistical heterogeneity and low risk of bias in the included trials, but substantial clinical heterogeneity within and between trials. At long-term follow-up (≥12 months) moderate quality evidence indicated that stem cell treatment was not associated with any changes in risk in all-cause mortality (6.3% vs 6.9%, relative risk [RR], 0.93, 95% CI 0.58 to 1.50), cardiovascular mortality (8.3% vs 7.2%, RR 1.04, 95% CI 0.54 to 1.99) or reinfarction/re-hospitalization (9.2% vs 14.0%, RR, 0.63, 95% CI 0.36 to 1.10). Similar results were reported for short-term follow-up. Stem cell therapy had no effect on morbidity or quality of life/performance, and the differences in mean LVEF between treatment groups, while reaching statistical significance in the majority of trials, was too low to be clinically relevant. While there remains insufficient evidence for a significant beneficial effect of stem cell therapy for AMI patients, the included RTCs may have been underpowered to detect differences in clinical outcomes.

Delewi (2014) published a systematic review of bone marrow cell therapy in patients with ST-elevation myocardial infarction (STEMI) that included 16 RCTs (n=1,641).[25] A meta-analysis of placebo-controlled RCTs that reported LVEF found statistically significant increases in LVEF with bone marrow stem-cell infusion compared with placebo (< six months, mean difference of 2.6 percentage points, 95% CI 1.8 to 3.3, p<0.001, I²=84%). Statistically significant reductions in LV end diastolic volumes were reported. Based on these findings, the authors concluded that intracoronary bone marrow cell infusion “is associated with improvement of LV function and remodeling in patients after STEMI.” Limitations of the meta-analysis included substantial statistical heterogeneity (I²≥55%).

De Jong (2014) conducted a meta-analysis of major adverse cardiac and cerebrovascular events based on literature through August 2013.[26] The analysis included 22 RCTs (n=1,513), 13 of which (n=1,300) were also included in the Delewi (2014) meta-analysis. Analysis of placebo-controlled RCTs that reported LVEF found statistically significant increases in LVEF with bone marrow stem-cell infusion compared with placebo (<18 months, mean difference of 2.1 percentage points, 95% CI 0.7 to 3.5, p<0.004, I²=80%). With median follow-up of six months, there was no difference between bone marrow cell infusion and placebo in all-cause mortality, cardiac mortality, restenosis rate, thrombosis, target vessel revascularization, stroke, recurrent AMI, or implantable cardioverter defibrillator implantations. Infusion with bone marrow progenitor cells, but not bone marrow mononuclear cells, led to a statistically significant reduction in the rate of rehospitalization for heart failure (odds ratio vs placebo, 0.14, 95% CI 0.04 to 0.52, p=0.003). Based on these findings, the authors concluded that, although safe, intracoronary infusion of bone marrow stem cells does not improve clinical outcome and clinical efficacy “needs to be defined in clinical trials.” Limitations of the meta-analysis included substantial statistical between-study heterogeneity (I²≥55%).

Lipinski (2007) published a quantitative meta-analysis of studies that estimated the magnitude of benefit of progenitor cell treatment on LV function and infarct size.[27] This analysis included 10 controlled trials with a total of 698 patients. Results for the primary endpoint, change in LVEF, showed a statistically significant greater improvement of 3.0% (95% CI 1.9 to 4.1%, p<0.00001) for the progenitor cell group. There was also a statistically significant greater
improvement in infarct size for the progenitor cell group with an incremental improvement of -5.6% over the control group (95% CI -8.7 to -2.5, p<0.001).

A 2008 BlueCross BlueShield Association Technology Evaluation Center (TEC) Assessment systematically reviewed RCTs of progenitor cell therapy versus standard medical care for treatment of either acute or chronic myocardial ischemia. The TEC Assessment focused on the impact of progenitor cell therapy on clinical outcomes, but also included data on physiologic outcomes such as change in LVEF. For acute ischemia, the TEC Assessment reviewed a total of 10 publications from six unique studies enrolling a total of 556 patients. These trials had similar inclusion criteria, enrolling patients with acute ST-segment elevation MI treated successfully with percutaneous coronary intervention (PCI) and stenting, with evidence of residual myocardial dysfunction in the region of the acute infarct. Progenitor cell therapy was delivered via an additional PCI procedure within one week of the acute event.

The REPAIR-AMI trial was the largest trial in this review, and had the largest number of clinical outcomes reported. This was a double-blind trial that employed a sham placebo control infusion of the patients’ own serum. This trial enrolled 204 patients with acute ST-segment elevation MI meeting strict inclusion criteria from 17 centers in Germany and Switzerland. At 12 months of follow-up, there were statistically significant decreases in the progenitor cell group for myocardial infarction (MI, 0 vs 6, p<0.03) and revascularization (22 vs 37, p<0.03) as well as for the composite outcome of death, MI, and revascularization (24 vs 42, p<0.009). The other trials had a very few number of clinical events, precluding meaningful analysis of clinical outcomes. The primary evidence from these other trials consists of physiologic outcomes measures such as change in LVEF and change in infarct size.

The primary endpoint in all six trials was change in LVEF. In each trial, there was a greater increase in the LVEF for the progenitor cell group compared with the control group. In four of the six studies, this difference reached statistical significance, while in two studies there was a nonsignificant increase in favor of the treatment group. The magnitude of the incremental improvement in LVEF was not large in most cases, with five of the six studies reporting an incremental change of 1.0% to 6.0%, and the final study reporting a larger incremental change of 18%.

At least four meta-analyses of BMC treatment for AMI were also found, each examining between six and 13 randomized, controlled trials, have been published since the 2008 TEC Assessment. All four meta-analyses concluded that there was a modest improvement in LVEF for patients treated with progenitor cells. The mean estimated improvement in ejection fraction over control ranged from 2.9 to 6.1%. The studies also concluded that myocardial perfusion and/or infarct size was improved in the progenitor cell treatment group, although different outcome parameters were used. All four of the meta-analyses concluded that there were no demonstrable differences in clinical outcomes for patients treated with progenitor cells.

Gyöngyösi (2015) conducted an individual patient data meta-analysis of 12 RCTs (n=1,252) on autologous intracoronary cell therapy after AMI, including the REPAIR-AMI trial discussed above, using a collaborative, multinational database, ACCRUE (meta-Analysis of Cell-based CaRdiac study, NCT01098591). All patients had STEMI treated with PCI. Mean (standard deviation [SD]) baseline LVEF was approximately 46% (12%). Most studies used bone marrow mononuclear cells and administered cell therapies within two weeks after AMI. Median follow-up duration was six months. Eight trials had low risk of bias, and four single-blind (assessor)
trials had medium-low risk of bias. Adjusted (for cardiovascular risk factors) random effects meta-analyses showed no effect of cell therapy on the primary end point, MACCE (major adverse cardiac and cerebrovascular events, a composite of all-cause death, AMI recurrence, coronary target vessel revascularization, and stroke) (186 events, 14.0% cell therapy vs 16.3% control, hazard ratio [HR], 0.86, 95% CI 0.63 to 1.18, I²=0%); death (21 events, 1.4% cell therapy vs 2.1% control); or a composite of clinical hard end points (death, AMI recurrence, and stroke, 45 events; 2.9% cell therapy vs 4.7% control). Compared with controls, changes in LVEF (mean difference, 0.96%, 95% CI −0.2 to 2.1), end-diastolic volume (mean difference, 1.2 mL, 95% CI -3.4 to 5.8), or end-systolic volume (mean difference, 3.6 mL, 95% CI -3.4 to 4.1) were not observed. The study was limited by variation in the time from AMI to cell delivery (median, 6.5 days) and in imaging modality for assessing cardiac function (magnetic resonance imaging [MRI], single-proton emission computed tomography [SPECT], angiography, echocardiography).

Section Summary

Reported study outcomes have ranged from modest improvement to no improvement with cell therapy compared with placebo in patients with acute ischemia. The current evidence to date should be viewed as preliminary rather than definitive. Most studies reported secondary outcomes such as LVEF and revascularization; minimal data was included for the primary outcomes of recurrent MI or mortality rates. All of the trials had one or more methodologic limitations. The most common limitations were lack of double-blinding and failure to account for all randomized patients in the analysis. The REPAIR-AMI trial was the highest methodologic quality, and was double-blinded. However, this trial excluded 17 of 204 randomized patients from the analysis, and thus was not considered to meet the criteria for a high-quality trial. While the evidence for a beneficial impact on physiologic outcomes, particularly LVEF, is fairly strong, the magnitude of effect does not appear to be large. As a result, it is not certain whether the improvement in LVEF translates to meaningful improvements in clinical outcomes, but further adequately powered trials are still needed to prove the efficacy of this intervention.

CHRONIC ISCHEMIC HEART DISEASE (IHD)

Systematic Reviews and Technology Assessments

Fisher (2016)[44] published an update to a 2014 Cochrane review with meta-analysis of autologous stem-cell therapy for chronic ischemic heart disease and congestive heart failure.[45] The review included 38 RCTs (n=1,907). The overall quality of the evidence was considered low because selected studies were small (only three included >100 participants) and the number of events was low, leading to a risk of small-study bias and spuriously inflated effect sizes. Results of the 2016 Cochrane review are shown in Table 1. While reviewers were unable to detect evidence of publication bias using funnel plots, they noted that, of 28 identified ongoing trials, 11 trials with 787 participants were recorded as having been completed or were due to have been completed in advance of the search date, but had no publications. Therefore, publication bias cannot be ruled out. Similar results were reported in 2014 meta-analyses conducted by Xu (2014)[46] and by Xiao (2014)[47].
Table 1. Cochrane Review Results of Stem Cell Therapy for Chronic Ischemic Heart Disease[44]

<table>
<thead>
<tr>
<th>Variables</th>
<th>Short-Term&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Long-Term&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Long-Term&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Long-Term&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Short-Term&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Short-Term&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mortality</td>
<td>Mortality</td>
<td>Rehospitalization</td>
<td>MACE</td>
<td>NYHA Classification</td>
<td>LVEF (%)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>N</td>
<td>1,637</td>
<td>1,010</td>
<td>495</td>
<td>201</td>
<td>658</td>
<td>352</td>
</tr>
<tr>
<td>PE (95% CI), p value</td>
<td>0.48 (0.26 to 0.87), 0.02</td>
<td>0.68 (0.25 to 0.58), &lt;0.001</td>
<td>0.62 (0.36 to 1.04), 0.07</td>
<td>0.68 (0.41 to 1.12), 0.13</td>
<td>-0.42 (-0.84 to -0.00), 0.05</td>
<td>3.01 (-0.05 to 6.07), 0.054</td>
</tr>
<tr>
<td>I² (p)</td>
<td>0% (0.76)</td>
<td>0% (0.97)</td>
<td>0% (0.70)</td>
<td>0% (0.80)</td>
<td>97% (&lt;0.001)</td>
<td>59% (0.01)</td>
</tr>
</tbody>
</table>

CI: confidence interval

Fisher (2016) also reported on the results of a sequential trial analysis using cumulative data obtained from two previous Cochrane reviews with updated results to March 2015.[48] The intent of their analysis was to obtain estimates of sample sizes required for a meta-analysis to detect a significant treatment effect while controlling for random errors due to repeat testing. Twenty-two trials that included all-cause mortality were selected. Six trials reported no deaths, while the remaining 16 trials reported 25 (5.6%) deaths in 444 patients who received progenitor cells compared with 50 (15.9%) deaths in 315 patients who did not. Meta-analysis of the pooled data revealed a significant reduction in mortality associated with cell therapy in patients with heart failure (RR=0.42, 95% CI 0.27 to 0.64, p<0.001).

The 2008 TEC Assessment, described above, included a total of six trials randomizing 231 patients for treatment of chronic ischemic heart disease. Three of these trials randomized a total of 125 patients to progenitor cell therapy versus standard medical care.[49-51] The other three trials randomized a total of 106 patients undergoing coronary artery bypass grafting (CABG) to CABG plus progenitor cell treatment versus CABG alone.[52-54] Four trials employed bone-marrow-derived progenitor cells as the donor cell source, one trial used circulating progenitor cells (CPC), and the final trial included both a CPC treatment group and a bone-marrow-derived treatment group.[49] The primary physiologic measurement reported in these trials was change in LVEF. In all six trials there was a greater improvement in LVEF for the treatment group compared with the control group, and in four of six trials, this difference reached statistical significance. For the three trials of progenitor cell treatment versus standard medical care, the range of incremental improvement in LVEF was 2.7% to 6.0%. For the trials of progenitor cell treatment plus CABG versus CABG alone, the range of improvement in LVEF was 2.5% to 10.1%. Only one trial reported comparative analysis of data on the change in size of ischemic myocardium. This trial reported that there was no difference in size of ischemic myocardium between treatment groups.[53]

There are limited data from this group of studies on clinical outcomes, with only two studies reporting any clinical outcomes.[49, 54] Both trials reported on change in New York Heart Association (NYHA) class between groups. Assmus also reported an improvement in mean NYHA class of 0.25 (0 to 4 scale) for the bone-marrow treatment group and an improvement of 0.23 for the CPC group, compared with a worsening of 0.18 for the standard medical therapy group (p<0.01).[49] Adverse cardiac events were reported to be extremely small in number with no differences between groups. Patel reported a greater improvement in mean NYHA class for patients in the CABG plus progenitor cell group compared to CABG alone (2.7 vs 0.7, p value not reported), but no statistical testing for this outcome was reported.[54]

Recent systematic reviews of smaller size have been published that include several new RCTs.[55-57] Xu (2014)[46] published a meta-analysis of 19 RCTs (n=886) using similar study
inclusion criteria to the Cochrane review with additional RCTs. Statistically significant improvement of LVEF was detected, as was a significant decrease in all-cause death (RR=0.49, 95% CI 0.29 to 0.84, p=0.01). Xiao (2014) \cite{47} included 20 RCTs that assessed stem cell therapy safety and efficacy in two subgroups of CIHD patients: those with revascularization and without revascularization. Bone marrow cell (BMC) transplantation significantly improved LVEF in patients both with and without revascularization, and patients without revascularization also had other measures of cardiac function significantly improve after BMC transplantation. In both studies the increases in cardiac function, although statistically significant, are too low to be considered clinically relevant. Both studies concluded that additional research in larger studies are required to confirm the efficacy of efficacy of BMC transplantation in CIHD patients.

**Randomized Controlled Trials**

Two phase 3 RCTs with more than 100 participants were identified. Bartunek (2017) reported on the results of a well-conducted double-blind trial in which 271 patients with NYHA class II or greater symptomatic heart failure (LVEF ≤35%) were randomized to bone marrow−derived mesenchymal cardiopoietic cells (n=120) or sham (n=151).\cite{58} The primary outcome was Finkelstein–Schoenfeld hierarchical composite (all-cause mortality, worsening heart failure, Minnesota Living with Heart Failure Questionnaire score, six-minute walk distance, left ventricular end-systolic volume, and ejection fraction) at 39 weeks. Sixteen patients who died and three who withdrew consent after randomization were not included in analysis. In addition, 19 patients whose cell product did not meet release criteria were excluded from analysis in the cardiopoietic cell group. The probability that the treatment group had a better outcome on the composite primary outcome was 0.54 (a value >0.5 favors active treatment, 95% CI 0.47 to 0.61, p=0.27). Exploratory subgroup analysis reported treatment benefit in patients, with baseline left ventricular end-diastolic volumes of 200 to 370 mL (60% of patients) (0.61, 95% CI 0.52 to 0.70, p=0.015). There was no statistical difference in serious adverse events between treatment arms. One (0.9%) cardiopoietic cell patient and nine (5.4%) sham patients experienced aborted or sudden cardiac death.

Pokushalov (2010) reported on the results of an RCT of intramyocardial injections of autologous bone marrow mononuclear cells (n=55) compared with optimal medical management (n=54) in patients who had chronic, ischemic heart failure.\cite{59} The trial appears to have been conducted in Russia; dates of study conduct were not reported. Power calculations were not reported, and it is not clear if the trial was registered. Comparative treatment effects were not calculated for many outcomes. The RCT reported statistically significantly improvements in mortality rates at 12 months for cell therapy (11%) vs medical therapy (39%) favoring medical therapy (p<0.001)

**Nonrandomized Studies**

The STAR-Heart trial evaluated stem cell therapy for chronic heart failure due to ischemic cardiomyopathy. This nonrandomized open-label study, reported by Strauer (2010), evaluated 391 patients with chronic heart failure.\cite{60} In this trial, 191 patients received intracoronary BMC therapy, and 200 patients who did not accept the treatment agreed to undergo follow-up testing served as controls. Mean time between percutaneous coronary intervention for infarction and admission to the tertiary clinic was 8.5 years. For BMC therapy, mononuclear cells were isolated and identified (included CD34-positive cells, AC133-positive cells, CD45-/CD14-negative cells). Cells were infused directly into the infarct-related artery. At up to five
years after intracoronary BMC therapy, there was a significant improvement in hemodynamics (LVEF, cardiac index), exercise capacity (NYHA classification), oxygen uptake, and left ventricular contractility compared with controls. There also was a significant decrease in long-term mortality in the BMC-treated patients (0.75% per year) compared with the control group (3.68% per year, p<0.01). However, the trial was limited by the potential for selection bias (patient self-selection into treatment groups). For example, there was a 7% difference in baseline ejection fraction rates between groups, suggesting that the groups were not comparable on important clinical characteristics at baseline. Additionally, lack of blinding raises the possibility of bias in patient-reported outcomes such as NYHA class.

Section Summary

For chronic ischemic heart disease, too few primary clinical outcome events (e.g., mortality rates) have been reported across studies to permit meaningful analysis. Other clinical outcomes such as NYHA class are confined to very small numbers of patients and lack sufficient methodologic rigor to permit conclusions. One well-conducted, phase 3 trial failed to demonstrate superiority for cell therapy for the primary outcome that included death, worsening heart failure, and other multiple events. The nonrandomized STAR-Heart trial showed a mortality benefit as well as a favorable hemodynamic effect but the lack of randomization limits interpretation due to concerns about selection bias and differences in known and unknown prognostic variables at baseline between arms. Overall, this evidence has suggested that progenitor cell treatment may be a promising intervention, but robust data on clinical outcomes are lacking. High-quality RCTs, powered to detect differences in clinical outcomes, are needed.

REFRACTORY ANGINA

Stem-cell therapy is also being investigated in patients with intractable angina who are not candidates for evascularization.

Systematic Reviews

A meta-analysis by Khan (2016) included six RCTs studying cell-based therapy in patients with refractory angina.61 The pooled outcomes of these trials were indices of angina (anginal episodes, Canadian Cardiovascular Society angina class, exercise tolerance, and antianginal medications, myocardial perfusion, and clinical endpoints. The authors created a composite end point, major adverse cardiac events, by combining myocardial infarction, cardiac-related hospitalization, and mortality. The analysis indicated that cell therapy led to improvements in many outcomes, compared with placebo, including anginal episodes (mean difference [MD] -7.81, 95% CI -15.22 to -0.41) Canadian Cardiovascular Society class (MD, -0.58, 95% CI -1.00 to -0.16), use of antianginal medications (standardized MD, -0.59, 95% CI -1.03 to -0.14), myocardial perfusion (standardized MD, -0.49, 95% CI -0.76 to -0.21), exercise tolerance (standardized MD, 0.331, 95% CI 0.08 to 0.55), risk of major adverse cardiac events (odds ratio, 0.49, 95% CI 0.25 to 0.98), and arrhythmias (odds ratio, 0.25, 95% CI 0.06 to 0.98). The authors suggest that these results require confirmation in larger, phase III RCTs.

The 2014 Cochrane review, described above, reported six studies that included patients with intractable or refractory angina.45 Five studies measured angina frequency. Combined data showed a significant difference (p=0.0002) in the short-term (<12 months follow-up) in favor of the stem cell groups compared to standard treatment without stem cells. The impact of stem cell therapy on mortality in patients with intractable/refractory angina is unclear because
participants included in the meta-analysis also had varying severity of IHD and heart failure. The authors ranked the level of evidence for this indication to be low quality and recommended further study in larger clinical trials to confirm present findings.

Li (2013) published a meta-analysis that included five RCTs (n=381) for stem cell therapy in patients with refractory angina.\[62\] Compared with controls, patients who received stem cells had a significant improvement in exercise tolerance (p=0.005), reduction in angina frequency (p=0.02), and lower risk of MI (p=0.04). No difference was found for risk of death (p=0.13). The authors concluded that the currently available findings require confirmation in larger studies with long-term follow-up.

Randomized Controlled Trials

Povsic (2016) reported on the industry-sponsored Efficacy and Safety of Targeted Intramyocardial Delivery of Auto CD34+ Stem Cells (RENEW) trial.\[63\] This three-arm multicenter trial compared outcomes from the intramyocardial administration of autologous CD34-positive cells using exercise capacity at 3, 6, or 12 months. Patients underwent cell mobilization with G-CSF for four days followed by apheresis. The peripheral cell product was shipped to a central processing facility (Progenitor Cell Therapy) for selection of CD34-positive cells. The trial was terminated after enrollment of 112 of a planned 444 patients before data analysis due to strategic considerations. The progenitor cell group had greater exercise capacity than the standard therapy group but was no better than the double-blinded placebo group, consistent with a placebo effect. Additionally, with only 122 participants, the trial was not adequately powered to detect a between-group difference.

Section Summary

Evidence on stem cell therapy for refractory angina includes early-phase trials, as well as a phase 3 pivotal trial terminated early and insufficiently powered to evaluate clinical outcomes. Additional larger trials are needed to determine whether progenitor cell therapy improves health outcomes in patients with refractory angina.

TREATMENT WITH GRANULOCYTE COLONY STIMULATING FACTOR (G-CSF)

Systematic Review

Moazzami (2013) published a Cochrane review of G-CSF for AMI.\[64\] Literature was searched in November 2010, and seven small, placebo-controlled RCTs (n=354) were included. Overall risk of bias was considered low. All-cause mortality did not differ between groups (RR 0.6, 95% CI 0.2 to 2.8, p=0.55, I^2=0%). Similarly, change in LVEF, LV end systolic volume, and LV end diastolic volume did not differ between groups. Evidence was insufficient to draw conclusions about the safety of the procedure. The study indicated a lack of evidence for benefit of G-CSF therapy in patients with AMI.

Randomized Controlled Trials

The following RCTs were published after the 2013 Cochrane summarized above:

Brenner (2016) evaluated G-CSF and Sitagliptin compared with placebo in 174 patients with AMI who had successful revascularization.\[65\] Both diabetic and nondiabetic patients were included. The primary endpoint of the trial was the hierarchically combined global left and right ventricular ejection fraction changes from baseline to six months follow-up, determined
by MRI. There were no significant differences between groups for this endpoint, and they had a similar risk of major cardiac adverse events.

Achilli (2010, 2014) published six-month[66] and three-year[67] results of their multicenter, placebo-controlled RCT, STEM-AMI. Sixty consecutive patients with first anterior STEMI, who underwent primary PCI within 12 hours after symptom onset and had LVEF of 45% or less were enrolled. Patients were randomized 1:1 to G-CSF 5 mg/kg body weight or placebo. Standard STEMI care was provided to all patients. Among cardiac MRI outcomes (LVEF, LV end systolic volume, LV end diastolic volume) at six months and three years, only LV end diastolic volume at three years was statistically significantly improved in the G-CSF group compared with placebo. At three years, there was no statistical difference in clinical outcomes, including death, reinfarction, target vessel restenosis or revascularization, heart failure, and stroke. The study was likely underpowered to detect statistically significant differences in most of these parameters.

Hibbert (2014) randomized 86 patients with LVEF less than 45% after anterior-wall MI to receive either G-CSF or placebo.[68] Eighty patients completed six-month follow-up. While both groups had improved LV function, the improvement was lower in the G-CSF group than in the placebo group. Similar rates in both groups were reported for target vessel revascularization. Both groups had one or more major adverse cardiac events in eight (19%) patients. The authors cautioned that careful monitoring for safety is warranted in future studies of G-CSF in this population.

**Section Summary**

The small number of trials that use G-CSF as a treatment for acute ischemia generally did not report an improvement in physiologic or clinical outcomes. The 2013 Cochrane review of seven placebo-controlled trials reported a lack of evidence for benefit. This evidence is not supportive of the use of G-CSF in the treatment of acute ischemia.

**PRACTICE GUIDELINE SUMMARY**

There are no clinical practice guidelines that address the use of progenitor cell therapy for the treatment of damaged myocardium due to ischemia.

**SUMMARY**

There is not enough research to determine whether progenitor cell therapy can improve health outcomes for patients with ischemic heart disease. No clinical guidelines based on research recommend progenitor cell therapy for patients with ischemic heart disease. Therefore, progenitor cell therapy is considered investigational for the treatment of ischemic heart disease.

**REFERENCES**


28. TEC Assessment 2008. "Progenitor cell therapy for treatment of myocardial damage due to ischemia." BlueCross BlueShield Association Technology Evaluation Center,


34. GP Meyer, KC Wollert, J Lotz, et al. Intracoronary bone marrow cell transfer after myocardial infarction: eighteen months' follow-up data from the randomized, controlled BOOST (BOne marrOw transfer to enhance ST-elevation infarct regeneration) trial. *Circulation.* 2006;113(10):1287-94. PMID: 16520413


---

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


**CODES**

**NOTE:** There are no specific codes for this procedure, either describing the laboratory component of processing the harvested autologous cells or for the implantation procedure. In some situations, the implantation may be an added component of a scheduled coronary artery bypass graft (CABG); in other situations, the implantation may be performed as a unique indication for a cardiac catheterization procedure.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>33999</td>
<td>Unlisted procedure, cardiac surgery</td>
</tr>
<tr>
<td></td>
<td>38205</td>
<td>Blood-derived hematopoietic progenitor cell harvesting for transplantation, per collection; allogeneic</td>
</tr>
<tr>
<td></td>
<td>38206</td>
<td>Blood-derived hematopoietic progenitor cell harvesting for transplantation, per collection; autologous</td>
</tr>
<tr>
<td></td>
<td>38240</td>
<td>Hematopoietic progenitor cell (HPC); allogeneic transplantation per donor</td>
</tr>
<tr>
<td></td>
<td>38241</td>
<td>Hematopoietic progenitor cell (HPC); autologous transplantation</td>
</tr>
</tbody>
</table>

**Date of Origin:** August 2004
In Vivo Analysis of Colorectal Lesions

Effective: February 1, 2022

Next Review: October 2022
Last Review: December 2021

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Several adjunct techniques of in vivo analysis of polyps are being researched for the purpose of improving the analysis of lesions and detection of changes in the colon. Use of these devices is proposed to increase the rate of polyp detection and/or to distinguish premalignant from benign lesions for removal.

MEDICAL POLICY CRITERIA

In vivo analysis of colorectal lesions, including but not limited to polyps, is considered investigational.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Confocal Laser Endomicroscopy, Medicine, Policy No. 151

BACKGROUND

During a colonoscopy or sigmoidoscopy as a screening test for colorectal cancer, the physician must often decide which polyp should be removed for histologic diagnosis. While
hyperplastic polyps are considered benign without malignant potential, adenomatous polyps are thought to represent one of the earliest stages in the progression to a malignancy. Identification of these premalignant lesions is considered one of the cornerstones of colorectal cancer prevention. The physician must thus balance the time and potential morbidity of removing all polyps, many of which will be benign, versus removal of those polyps most likely to be adenomatous.

Several techniques of in vivo analysis of polyps are being researched for the purpose of improving the analysis of lesions and detection of changes in walls of colon. These methods are intended to be used as an adjunct to colonoscopy. Some of these methods include autofluorescence, narrow band imaging (NBI), multi-band imaging, chromoendoscopy, third eye retroscope and fiberoptic analysis. It is proposed that these technologies may allow for in vivo analysis of the polyps, possibly avoiding unnecessary biopsies and increasing detection of difficult to visualize lesions (e.g., flat lesions).

The first system developed was based on the observation that benign and malignant tissues emit different patterns and wavelengths of fluorescence after exposure to a laser light. This system consists of an optical fiber, emitting a laser that is directed against three different regions of the same polyp. The subsequent fluorescent signal is collected, measured, and analyzed by a proprietary software system, which classifies a polyp as "suspicious" (i.e., adenomatous) or "not suspicious" (i.e., hyperplastic). There are several different types of spectroscopy-based in vivo techniques that rely on autofluorescence, emitting light at different frequencies in an attempt to distinguish between hyperplastic and adenomatous lesions.

Narrow band imaging (NBI) is another new technique that allows visualization of the mucosal surface and capillary vessels and thus may assist in the differentiation of abnormal from normal mucosa during colonoscopy. Two NBI systems are available. The NBI color chip system is used in the United States; in this system a single filter with a two-band pass characteristic is used to generate central wavelengths at 415 nm (blue) and 540 nm (green and red). The NBI red-green-blue sequential illumination system uses narrow spectra of red, green, and blue light and a video endoscopic system with a frame sequential lighting method. The light source unit consists of a xenon lamp and a rotation disk with three optical filters. The rotation disk and monochrome charge-coupled device are synchronized and sequentially generate images in three optical filter bands. By use of all three band images, a single color endoscopic image is synthesized by the video processor. NBI has limited penetration into the mucosal surface and has enhanced visualization of capillary vessels and their fine structure on the surface layer of colonic tissue.

Chromoendoscopy, also known as chromoscopy and chromocolonoscopy, refers to the application of topical stains or dyes during endoscopy to enhance tissue differentiation or characterization and facilitate identification of mucosal abnormalities. Chromoendoscopy may be particularly useful for detecting flat or depressed lesions. Standard colonoscopy uses white light to view the colon. In chromoendoscopy, stains are applied, resulting in color highlighting of areas of surface morphology of epithelial tissue. The dyes or stains are applied via a spray catheter that is inserted down the working channel of the endoscope. Chromoendoscopy can be used in the whole colon (pancolonic chromoendoscopy) on an untargeted basis or can be directed to a specific lesion or lesions (targeted chromoendoscopy). Chromoendoscopy differs from endoscopic tattooing in that the former uses transient stains, whereas tattooing involves the use of a long-lasting pigment for future localization of lesions.
Virtual chromoendoscopy (also called electronic chromoendoscopy) involves imaging enhancements with endoscopy systems that could be an alternative to dye spraying. One system is the Fujinon® Intelligent Color Enhancement (FICE®) feature (Fujinon Inc.). This technology uses postprocessing computer algorithms to modify the light reflected from the mucosa from conventional white light to various other wavelengths.

REGULATORY STATUS

Auto-fluorescence

In 2000, the Optical Biopsy™ System (SpectraScience™, Inc.) was approved by the Food and Drug Administration (FDA). The FDA-labeled indication for the Optical Biopsy™ System reads as follows:[1]

"The SpectraScience™ Optical Biopsy™ System is indicated for use as an adjunct to lower gastrointestinal endoscopy. The device is intended for the evaluation of polyps less than 1 cm in diameter that the physician has not already elected to remove. The device is only to be used in deciding whether such polyps should be removed (which includes submission for histological examination)."

NBI

NBI received FDA clearance through the 510K process in 2005. This clearance (K051645) added NBI with the EVIS EXERA 160A System (Olympus Medical Systems Corp.) to existing endoscopic equipment. FDA indications are for endoscopic diagnosis, treatment, and video observation. In addition, in 2012, the EVIS EXERA III System, which has dual focus (DF) capabilities received FDA approval.[2]

Chromoendoscopy

In August of 2016, the Fuse Colonoscope with FuseBox Processor was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process.[3] This system is indicated for use within the lower digestive tract for adult patients. This system includes Lumos and is intended to be used as an optional adjunct following white light endoscopy and is not intended to replace histopathological sampling as a means of diagnosis.

In August 2014, the Fujifilm EPX-4440HD Digital Video Processor with FICE and Light Source was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process. In October of 2015, the PMA was extended to include and additional digital video processor, EPX-4440. FDA documents state that FICE can be used to supplement white-light endoscopy but is not intended to replace histopathologic sampling as a means of diagnosis. In January 2017, the Fujifilm Processor VP-7000 and Light source BL-7000 was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process with the EPX-4440HD as a predicate device.[4] FDA documents state “BLI (Blue Light Imaging), LCI (Linked Color Imaging) and FICE (Flexible spectral-Imaging Color Enhancement) are adjunctive tools for gastrointestinal endoscopic examination which can be used to supplement Fujifilm white light endoscopy. BLI, LCI and FICE are not intended to replace histopathological sampling as a means of diagnosis.”

In April 2003, the i-scan™ (Pentax), used for virtual chromoendoscopy, was cleared for marketing by FDA through the 510(k) process.[5] This is a digital image enhancement technology and is part of the Pentax EPK-i5010 and EPK-i7010 Video Processors. The i-scan
has several modes that digitally enhance images in real-time during endoscopy. FDA documents state that i-scan is intended as an adjunct following white-light endoscopy and is not intended to replace histopathologic analysis.

No dye or stain product has been specifically approved by FDA for use in chromoendoscopy.

EVIDENCE SUMMARY

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition. The comparator for the techniques discussed in this review is standard definition white light endoscopy (SD-WLE) or high-definition WLE (HD-WLE). The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

MULTIPLE TECHNIQUES

Systematic Reviews

El-Dallal (2020) reported results of a meta-analysis comparing virtual chromoendoscopy, dye-spraying chromoendoscopy, and HD-WLE. Eleven randomized controlled trials met inclusion criteria. The quality of evidence was moderate in the HD-WLE studies and low to moderate in the DCE studies. In the per-patient analysis of the 1,328 patients, there were no statistically significant differences between virtual chromoendoscopy and dye-spraying chromoendoscopy (risk ratio [RR] 0.77; 95% CI 0.55 to 1.08) or between virtual chromoendoscopy and HD-WLE (RR 0.72; 95% CI 0.45 to 1.15). In the per-dysplasia analysis, there were no statistically significant differences between virtual chromoendoscopy and dye-spraying chromoendoscopy (RR 0.72; 95% CI 0.47 to 1.11), but virtual chromoendoscopy was inferior to HD-WLE (RR 0.62; 95% CI 0.44 to 0.88).

Facciorusso (2019) performed a systematic review of RCTs comparing the efficacy of a variety of devices for the detection of adenomas. A total of 74 two-arm trials assessing add-on devices, enhanced imaging techniques, new scopes, and low-cost optimization of existing resources were included. Moderate increases in adenoma detection rate were found for low-cost optimization of existing resources (odds ratio [OR], 1.29; 95% confidence interval [CI] 1.17 to 1.43), enhanced imaging techniques (OR,1.21; 95% CI 1.09 to 1.35), and add-on devices (OR,1.18; 95% CI 1.07 to 1.29). Of those, no specific technology was superior to others for detection of advanced adenomas, polyp detection rate, or mean number of adenomas per patient, indicating that low-cost optimization of existing resources was as effective as enhanced endoscopic imaging.

Bessissow (2018) performed a systematic review and meta-analysis of RCTs that compared dysplasia detection techniques in patients with ulcerative colitis. Eight parallel group RCTs including 924 patients met inclusion criteria. Patients were adults with long-standing ulcerative colitis (UC) undergoing surveillance colonoscopy with SD-WLE, HD-WLE, narrow band imaging (NBI), or dye-based chromoendoscopy. The evidence was rated as low-to very low-quality using GRADE. The meta-analysis supported chromoendoscopy over SD-WLE (odds
ratio [OR], 2.37; 95% credible interval [CrI], 0.81 to 6.94) for any dysplasia detection with low-quality evidence, whereas very low-quality evidence supports using HD-WLE or NBI over SD-WLE (HD-WLE [vs SD-WLE]: OR, 1.21; 95% CrI, 0.30 to 4.85; NBI: OR, 1.68; 95% CrI, 0.54 to 5.22).

Lord (2018) performed a systematic review of the diagnostic accuracy of several techniques of colonic lesion characterization. A total of 22 studies assessing techniques for in-vivo optical characterization of lesions in patients with colonic IBD during colonoscopy, including 1,491 patients, met inclusion criteria. Techniques examined were virtual chromoendoscopy (VCE), dye-based chromoendoscopy (DBC), magnification endoscopy and confocal laser endomicroscopy (CLE). The quality of included studies was rated and there was mixed quality for all three domains of risk of bias (patient selection, index test, and reference standard). Pooled sensitivities of CLE, magnification endoscopy, DBC, and VCE were 91% (95% CI 94 to 98%), 90% (95% CI 77 to 96%), 67% (95% CI 44 to 84%) and 86% (95% CI 62 to 95%), respectively. Pooled specificities of magnification endoscopy, VCE, and DBC were 87% (95% CI 81 to 91%), 87% (95% CI 72 to 95%), and 86% (95% CI 72 to 94%), respectively, and the area under the SROC curve for CLE was 0.98 (95% CI 0.97-0.99). The authors concluded that real-time CLE is a highly accurate technology while acknowledging that this study is limited by the fact that most CLE studies were performed by single expert users within tertiary centers.

In 2013, Wanders assessed the sensitivity, specificity, and real-time negative predictive value of NBI, image-enhanced endoscopy (i-scan), Fujinon intelligent chromoendoscopy (FICE), CLE, and autofluorescence imaging for differentiating neoplastic from non-neoplastic colon lesions. A total of 91 studies were included in the analysis (NBI=56, i-scan=10, FICE=14, CLE=11 and autofluorescence imaging=11). The authors reported the following for each modality:

- "For NBI, overall sensitivity was 91.0% (95% CI 88.6 to 93.0), specificity 85.6% (81.3 to 89.0), and real-time negative predictive value 82.5% (75.4 to 87.9).
- For i-scan, overall sensitivity was 89.3% (83.3 to 93.3), specificity 88.2% (80.3 to 93.2), and real-time negative predictive value 86.5% (78.0 to 92.1).
- For FICE, overall sensitivity was 91.8% (87.1 to 94.9), specificity 83.5% (77.2 to 88.3), and real-time negative predictive value 83.7% (77.5 to 88.4).
- For autofluorescence imaging, overall sensitivity was 86.7% (79.5 to 91.6), specificity 65.9% (50.9 to 78.2), and real-time negative predictive value 81.5% (54.0 to 94.3).
- For CLE, overall sensitivity was 93.3% (88.4 to 96.2), specificity 89.9% (81.8 to 94.6), and real-time negative predictive value 94.8% (86.6 to 98.1)."

The authors did not recommend autofluorescence imaging as a reliable optical diagnostic option due to low specificity rates. This study did not assess whether any of these optical imaging modalities improved patient management or overall health outcomes.

**Randomized Controlled Trials**

Iacucci (2018) performed a randomized non-inferiority trial to determine detection rates of neoplastic lesions in IBD patients with longstanding colitis. A total of 270 patients with inactive disease were enrolled and divided evenly to be assessed by high definition (HD), dye spraying chromoendoscopy (DCE), or VCE using i-scan image enhanced colonoscopy. Neoplastic lesions were classified by the Paris classification and Kudo pit pattern followed by histological classification using the Vienna classification. VCE was determined to have non-
inferior neoplastic lesion detection rates compared to DCE. HD rates of detection of all neoplastic lesions were non-inferior to DCE and VCE. Kudo pit pattern and location at the right colon were found to predict neoplastic lesions. The authors concluded that HD-WLE alone was sufficient for detection of dysplasia, adenocarcinoma, or all neoplastic lesions.

**AUTO-FLUORESCENCE IMAGING**

**Nonrandomized Studies**

In 2013, Inomata conducted a prospective nonrandomized trial to evaluate colorectal lesions using a new auto-fluorescence imaging (AFI) system.[12] A total of 88 patients with 163 lesions greater than 5 mm were evaluated using the novel AFI system which assessed the green/red (G/R) ratio for each lesion using a computer-assisted color analysis system that permits real-time color analysis during endoscopic procedures. Authors reported significant differences in the G/R ratios of hyperplastic polyps, adenoma/intramucosal cancer/submucosal (SM) superficial cancer, and SM deep cancer (p<0.0001). The mean ± SD G/R ratios were 0.984 ± 0.118 in hyperplastic polyps and 0.827 ± 0.081 in neoplastic lesions. When a cut-off value of >0.89 was applied to non-neoplastic lesions, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were 83.9%, 82.6%, 53.1%, 95.6% and 82.8%, respectively. When a cut-off value of <0.77 was applied to identify SM deep cancers, the sensitivity, specificity, PPV, NPV, and accuracy were 80.0%, 84.4%, 29.6%, 98.1% and 84.1%, respectively. Additional studies are needed to validate these cut-off values and to assess the impact of AFI upon improved health outcomes.

The FDA approval for the SpectraScience™ Optical Biopsy™ System was based on a prospective, nonrandomized phase II study involving 101 subjects from five sites. The data from this trial have not been published in a peer-reviewed journal but are available as an FDA summary of safety and effectiveness.[1] Patients who participated in the study had undergone a prior lower GI endoscopic procedure with at least one polyp identified. They were then referred for an additional colonoscopy exam, in which fiberoptic analysis of the polyps was performed. At the time of the colonoscopy, the physicians documented whether or not the polyp was considered hyperplastic or adenomatous, and whether or not they would remove the polyp. The fiberoptic probe was then applied to three different portions of the polyp and a segment of normal adjacent mucosa. The physician did not know the results of the analysis and thus the test did not affect patient treatment. The effectiveness of the analysis was then calculated as its ability to correctly identify adenomatous polyps (sensitivity) and to correctly identify hyperplastic polyps (specificity), either alone or in conjunction with the physician assessment. The sensitivity and specificity of the physician assessment alone was 82.7% and 50%, respectively, compared to a combined sensitivity and specificity of 96.3% and 33%, respectively. In other words, fiberoptic analysis identified additional adenomatous polyps that the physician had classified as hyperplastic and presumably would not have removed based on visual assessment alone. This increase in sensitivity comes at the price of a decrease in specificity, as more hyperplastic polyps will undergo biopsy. However, according to the FDA, the risk of taking biopsies of additional hyperplastic polyps is minimal.

The clinical significance of these results and their effect on patient management is difficult to interpret from the data presented. It is not clear how the physician decided to select additional polyps for fiberoptic analysis (it is not entirely clear whether all polyps were analyzed and then underwent biopsy), or whether the same results could be obtained by simply randomly taking a biopsy of a subset of polyps that were considered hyperplastic on visual assessment. While
adenomatous polyps are considered premalignant lesions, the evolution to cancer is a slow process requiring seven to eight years, and thus the immediate removal of all adenomatous polyps is not required. In addition, the finding of an adenomatous polyp serves as a marker that the patient should undergo more frequent endoscopic exams. It is well known that the current practice of visual inspection of polyps will certainly miss some adenomatous polyps, but this lack of sensitivity is considered acceptable if at least one adenomatous polyp is identified and the patient undergoes more frequent screening.

Few studies have been published on the SpectraScience™ Optical Biopsy™ System since 2002. A feasibility study of fiberoptic analysis of normal, adenomatous, and cancerous tissue in 11 patients was published by Mayinger in 2003.[13] No additional literature on the Optical Biopsy™ System was found, but a report in 2006 detailed the results of spectral scattering to different colonic lesions in a small series of 45 patients.[14]

NARROW BAND IMAGING (NBI)

The following evidence review for the diagnostic utility of NBI will focus on RCTs comparing NBI with white light and standard colonoscopy techniques.

Systematic Reviews

Feuerstein (2019) performed a systematic review and meta-analysis of RCTs and non-RCTs that assessed the efficacy of NBI versus white light endoscopy.[15] Six RCTs and four non-RCTs met inclusion criteria. The reported detection rates were 17% and 11%, respectively, for chromoendoscopy and white light endoscopy, respectively (relative risk 1.50; 95% CI, 1.08 to 2.10). The quality of evidence from RCTs was moderate. In data from non-RCTs, chromoendoscopy was more effective than white light endoscopy (16% versus 6%; RR, 3.41; 95% CI 2.13 to 5.47). The quality of evidence from non-RCTs was very low.

Atkinson (2019) performed a systematic review and meta-analysis of RCTs that assessed the adenoma detection rate in NBI versus white light endoscopy.[16] Studies of patients with inflammatory bowel disease or with familial or genetic syndromes were excluded. A total of 11 trials met inclusion criteria and data from 4491 patients were analyzed. A risk of bias assessment was performed, and little evidence of publication bias was found. The detection rate was similar overall, with an unadjusted OR for detection of adenoma by white light endoscopy vs NBI of 1.14 (95% CI 1.01 to 1.29; p=0.04). However, in cases when bowel preparation was considered best, NBI outperformed WLE (adequate preparation OR, 1.07, 95% CI 0.92 to 1.24, p=0.38; vs best preparation OR, 1.30 95% CI 1.04 to 1.62, p=0.02). Additionally, second-generation, but not first-generation, NBI had a better detection rate than white light endoscopy (second-generation NBI OR, 1.28; 95% CI 1.05 to 1.56, p=0.02).

Sabbagh (2011) conducted a meta-analysis of studies (regardless of indication) evaluating NBI compared to colonoscopy and did not find any significant differences in the mean number of polyps (five RCT, 2479 participants), the mean number of adenomas (eight RCTs, 3517 participants), and the rate of patients with at least one adenoma (eight RCTs, 3512 participants).[17] However, individual studies included in the analysis were noted to have heterogeneous populations and indications, as well as diverse findings. Overall, the authors concluded that NBI did not improve detection of colorectal polyps when compared with conventional colonoscopy.
Additional reviews assessing the ability of NBI to differentiate between neoplastic and non-neoplastic polyps have been published; however, these studies are limited due to their inclusion of nonrandomized studies and lack of analysis regarding the impact of NBI upon patient management of overall health outcomes.[18]

Randomized Controlled Trials

Data from several randomized trials of NBI versus white-light colonoscopy (WLE) failed to show any advantage in total detection rate for NBI.[17, 19-23] Published randomized trials differ from the conventional approach to the assessment of diagnostic tests. In these trials patients were randomized to one test or the other (i.e., they received only one test). In general, when comparing diagnostic tests, each patient would receive both tests and the test results would be compared, which more recent trials have done.

Jung (2021) reported results of a randomized study comparing NBI and WLE to detect remnant tissue following removal of suspicious sessile-serrated adenoma.[24] A total of 145 lesions were removed from 138 patients. There were no statistically significant differences in histologic diagnostic rate (89.9% (62/69) vs. 85.5% (65/76); p>0.05), detection of remnant tissue (12.9% (8/62) vs. 15.4% (10/65); p>0.05), the proportion of SSA in remnant tissue (11.3% (7/62) vs. 12.3% (8/65); p>0.05), or the proportion of incomplete resection (6.5 (4/62) vs. 10.8 (7/65); p>0.05) between the NBI and WLE inspection groups, respectively.

Riu Pons (2020) conducted a randomized cross-over trial to compare NBI with HD-WLE in 41 patients with prior detection of at least one serrated polyp ≥10 mm or ≥ 3 serrated polyps larger than 5 mm, both proximal to the sigmoid colon.[25] All patients received tandem same-day colonoscopies with both techniques, performed by one of five experienced endoscopists, with the order being randomized 1:1 to NBI-HD-WLE or HD-WLE-NBI. All tandem colonoscopies were performed by the same endoscopist. No differences were reported in serrated lesion detection rate (47.4% for NBI versus 51.9% for HD-WLE; OR 0.84, 95% CI 0.37 to 1.91) or polyp miss rate (21.3% for NBI versus 26.1% for HD-WLE; OR 0.77, 95% CI 0.43 to 1.38).

Kim (2019) randomized 117 patients to NBI using the new 290 system (290-NBI) or HDWL colonoscopy.[26] All patients were then inspected with the technology not used initially, such that each patient was inspected with both NBI and HDWL with the order randomized. While the adenoma or polyp detection rates were not different between the two groups (polyp miss rates for 290-NBI and HDWL were 20.6% and 33.9%, respectively; p=0.068), the non-adenomatous polyp miss rate for 290-NBI was significantly lower than that of HDWL (11.5% vs. 52.2%, p=0.002). In addition, for polyps on the left side of the colon, flat-type polyps, and non-adenomatous polyps miss rates were significantly lower for 290-NBI than HDWL.

In a 2017 RCT, Min reported on 152 patients (142 were included in the analysis) that underwent crossover colonoscopies with white light endoscopy and linked color imaging (LCI), which uses narrow-band short-wavelength light and WL, randomized for order.[27] The sensitivities in the white light and LCI groups were significantly different, at 73% and 91%, respectively. Negative predictive value was not reported.

In a 2016 RCT, Klare randomized 380 patients to the NBI arm or the high-definition white light arm.[28] Accuracy was 73.7% and 79.2%, sensitivity was 82.4% and 79.8%, and negative predictive value was 75.5% and 73.4% in the NBI and white light arms, respectively. These values were not significantly different between arms.
In a randomized controlled trial reported by Gross (2011), 100 patients undergoing routine screening and surveillance were randomized to receive tandem colonoscopies with standard definition white light (SDWL) and image-enhanced (HD-NBI) colonoscopy. The main outcome measurement was the per-polyp false-negative (“miss”) rate. Secondary outcomes were adenoma miss rate, and per-patient polyp and adenoma miss rates. Polyp and adenoma miss rates for SDWL colonoscopy were 57% (60/105) and 49% (19/39); those for image-enhanced colonoscopy were 31% (22/72) and 27% (9/33) (p=0.005 and p=0.036 for polyps and adenomas, respectively). Image-enhanced and SDWL approaches had similar per-patient miss rates for polyps (6/35 vs. 9/32, p=0.27) and adenomas (4/22 vs. 8/20, p=0.11). The authors concluded that utilization of multiple recent improvements in image-enhanced colonoscopy was associated with a reduced miss rate for all polyps and for adenomatous polyps. It is not known which individual feature or combination of image-enhancement features led to the improvement.

Kakol (2013) evaluated the usefulness of NBI for detection of missed polyps after colonoscopy comparing white light (WL) to NBI. After initial colonoscopy 253 patients were randomized to a second colonoscopy with either NBI or WL. Authors found no significant difference between missed polyps or adenomas between groups.

In 2014, Wallace published results an RCT which compared NBI to standard colonoscopy and found no differences between groups. A total of 522 patients were randomized and 927 total polyps were analyzed. No differences were observed in adenoma detection rate or diagnostic accuracy, regardless of polyp size.

Several randomized trials addressed both total detection rate and differentiation of neoplastic from nonneoplastic lesions.

Pohl conducted a randomized multicenter trial in 2009 of virtual chromoendoscopy with the “Fujinon intelligent colour enhancement” system (FICE or NBI) versus standard colonoscopy with targeted indigocarmine chromoscopy. This German trial included 764 patients in the final analysis and reported that FICE/NBI was not superior to control for overall adenoma detection rates; it was comparable on the differentiation of neoplastic and non-neoplastic lesions. The sensitivity of FICE/NBI was 92.7% versus 90.4% for the control.

Additional RCTs were identified; however, these studies contained several methodological flaws in that they only reported on the accuracy of the NBI system in the in vivo evaluation of colonic polyps. In addition, none of the studies evaluated the impact of this technology on outcomes including whether or not there would be an improvement in the selection of polyps for removal during colonoscopy. Furthermore, subsequent RCTs demonstrate no differences in polyp detection rate of NBI compared to WL.

**CHROMOENDOSCOPY**

**Systematic Reviews**

Azizi (2018) performed a systematic review comparing white light endoscopy and chromoendoscopy for identifying dysplastic or cancerous lesions in patients with ulcerative colitis without primary sclerosing (PSC) or Crohn’s disease (CD). Studies were included if they reported on colonoscopy detection rates of dysplasia and cancers in UC without involvement of PSC or CD. Ten studies met inclusion criteria; most were of moderate quality. Publication bias was not assessed due to the low number of publications per incidence.
outcome. A meta-analysis of the five studies reporting overall pick-up rate of dysplastic/cancerous lesions on WLE random biopsies calculated showed a pooled rate of 5.6%. Only one study reported on the use of chroomeoscopy for ulcerative colitis patients without PSC. The reported pick-up rate of dysplastic lesions in this study was 7%.

In 2016, Brown updated their 2010 Cochrane review that compared chroomeoscopy and conventional colonoscopy for the detection of colorectal lesions in individuals at increased risk of colorectal neoplasia due to family history, previous polyp detection, or previous CRC resection.[38, 39] The review excluded studies of individuals with IBD or a known polyposis syndrome. Seven RCTs (2,727 participants) were included, five of which were used for a meta-analysis. All of these studies were published prior to 2012. The review found that chromoscopy was likely to yield more people with at least one neoplastic lesion (odds ratio (OR) 1.53, 95% CI 1.31 to 1.79; seven trials; 2,727 participants), and significantly more people with three or more neoplastic lesions were also detected, but only when studies that used high-definition colonoscopy in the control group were excluded (OR 4.63, 95% CI 1.99 to 10.80; two trials; 519 participants). None of the included studies reported any adverse events related to the use of the contrast dye. However, all the trials had some methodological drawbacks, and all were graded as low quality. In addition, some of the included studies were underpowered and significant heterogeneity was present between the included studies (variability of the colonoscopes used in the studies and differences in dye-spraying technique). There are also differences in the study inclusion criteria between the included studies).

Randomized Controlled Trials

The following randomized controlled trials were not included in the above systematic reviews.

Wan (2021) conducted a prospective, multicenter randomized controlled study on patients with longstanding (at least six years) ulcerative colitis.[40] The study compared chroomeoscopy with targeted biopsies to white-light endoscopy with targeted biopsies and random biopsies. In the full-analysis data set, a total of 122 patients with 447 colonoscopies were analyzed, and the randomized groups were as follows: chroomeoscopy (n=39), white-light endoscopy-targeted (n=43), and white-light endoscopy-random (n=40). The primary outcome of the study was the number of colonoscopies that diagnosed dysplasia in each group. The median follow-up period during the study was 55 months; white-light endoscopy-random and chroomeoscopy treated patients had more colonoscopies than white-light endoscopy-targeted treated patients (8.0% vs. 1.9%, p=.013; 9.3% vs. 1.9%, p=0.004, respectively). There was no significant difference found between the white-light endoscopy-random and chroomeoscopy groups. In a sub-group analysis in the second half of the follow-up period (37 to 69 months), chroomeoscopy had more colonoscopies that diagnosed dysplasia than white-light endoscopy-targeted (13.3% vs. 1.6%, p=0.015) and had results that indicated a trend for increasing dysplasia detection rates compared to white-light endoscopy-random (13.3% vs. 4.9%, p=0.107).

Alexandersson (2020) conducted a single-center, prospective study on 305 patients with longstanding (at least eight years) ulcerative colitis or Crohn colitis.[41] The study compared high-definition chroomeoscopy with high-definition white-light endoscopy. Patients were randomized into each group: chroomeoscopy (n=152) and white-light endoscopy (n=153). The primary outcome was the number of patients with dysplastic lesions. Dysplastic lesions were detected in 17 patients in the chroomeoscopy group (11%) and in seven patients in the white-light endoscopy group (5%), which was statistically significant (p=0.032). The total
The number of macroscopic lesions detected in the chromoendoscopy group versus the white-light endoscopy group (n= 89 vs. 41, respectively) was statistically significant (p<0.001), and the total number of macroscopic lesions containing dysplasia was higher in the chromoendoscopy group (n=24; p=0.029). The study found that chromoendoscopy was superior to white-light endoscopy in the detection of dysplastic lesions during colonoscopy; however, the study was limited to a single-center institution in Sweden and the expertise of the endoscopists was not detailed.

Yang (2019) performed a randomized controlled trial comparing HD-WLE with random biopsy versus high-definition chromoendoscopy with targeted biopsy in 210 patients with longstanding ulcerative colitis.[42] The difference in detection rates of colitis-associated dysplasia were not statistically significant between groups (20.6% for chromoendoscopy vs. 12.0% for HD-WLE; p=0.093). The median length of colonoscopy withdrawal was not significantly different between groups (17.6 vs 16.5 minutes; p=0.212) but the difference in total number of biopsies was statistically significant, with 34 in the HD-WLE group and nine in the chromoendoscopy group (p<0.001).

Haanstra (2019) reported results of a multicenter RCT of patients with Lynch syndrome who were undergoing regular surveillance by colonoscopy.[43] A total of 246 patients were randomly assigned (1:1) to conventional WLE (n=123) or colonoscopy with CE in the proximal colon (n=123). Patients were stratified for previous colorectal adenomas and enrolling center. The primary outcome was the proportion of patients with the detection of at least one neoplastic lesion at baseline and after two years. Detection rates were not significantly different between groups at either baseline (27% for WLE versus 30% for CE; OR, 1.23; 95% CI 0.69 to 2.2; p=0.56) or two years (26% for the original WLE group versus 28% for the CE group (OR, 1.1; p=0.81).

Rondonotti (2019) compared blue-light imaging (BLI) chromoendoscopy with HDWL endoscopy for the characterization of polyps in patients undergoing colonoscopy.[44] A total of 358 consecutive patients undergoing outpatient colonoscopy who had at least one polyp less than 10mm were randomized to BLI or HDWL for polyp characterization. The number of polyps characterized with high confidence was not significantly different between groups (p=0.887), though the overall accuracy was, in favor of BLI (92% versus 84%, p=0.011).

Vleugels (2018) randomized patients undergoing dysplasia surveillance for longstanding ulcerative colitis at five centers in the Netherlands and the UK to receive autofluorescence imaging or chromoendoscopy.[45] Patients were eligible if they were age 18 years or older and were undergoing dysplasia surveillance after a diagnosis of extensive colitis at least eight years before the study start or left-sided colitis at least 15 years before the study start. Each group contained 105 patients. Primary outcomes were the proportion of patients in whom at least one dysplastic lesion was detected and the mean number of dysplastic lesions per patient. Dysplasia was detected in 12% and 19% of patients in the autofluorescence and chromoendoscopy groups, respectively. The mean number of detected dysplastic lesions per patient was 0.13 (SD 0.37) and 0.37 (SD 1.02) for autofluorescence and chromoendoscopy, respectively. Two and three adverse events were reported in the autofluorescence and chromoendoscopy groups, respectively. Autofluorescence imaging did not meet criteria for proceeding to a large non-inferiority trial.

VIRTUAL CHROMOENDOSCOPY

Systematic Reviews
Aziz (2019) performed a systematic review of RCTs comparing “distal attachments” (endocap, endocuff, and endoring) or “electronic chromoendoscopy” (narrow-band imaging, iScan, blue-light imaging, autofluorescence imaging, and linked-color imaging) with high definition white light endoscopy for the detection of serrated adenomas. A total of 17 studies including 13,631 patients met inclusion criteria. There was no statistically significant improvement in serrated adenoma detection rate identified using distal attachments (RR 1.21; p=0.45) or electronic chromoendoscopy (RR 1.29; p=0.09).

A meta-analysis by Omata published in 2014 compared the rate of polyp detection by virtual chromoendoscopy (i.e., FICE or i-scan) with white-light colonoscopy. The review included patients of all risk levels and was limited to RCTs. Five trials on FICE/i-scan met eligibility criteria and the analysis did not find a significantly higher detection rate with virtual chromoendoscopy. The pooled relative risk of adenoma/neoplasia detected by virtual chromoendoscopy versus conventional chromoendoscopy was 1.09 (95% CI 0.97 to 1.23; p>0.05).

Randomized Controlled Trials

Kandiah (2021) published a multicenter RCT comparing the performance of high-definition white light versus high-definition virtual chromoendoscopy in patients in the United Kingdom with longstanding (at least 8 years) ulcerative or Crohn colitis. Patients were randomized, prior to starting surveillance colonoscopy, to either white light (n=92) or virtual chromoendoscopy (n=92) for a total of 184 patients included in the final analysis. The primary outcome was the difference in neoplasia detection rate between the two arms. Twenty-five neoplastic lesions were found in 14 patients in the virtual chromoendoscopy arm; 27 lesions were found in 22 patients in the white light arm. Compared to the virtual chromoendoscopy arm, neoplasia detection rate was higher in the white light arm (23.4% vs. 14.9%), but this was not statistically significant (p=0.14). The mean number of biopsies taken per patient was 35.9 in each arm of the study, and the difference in the mean number of neoplasia per patient was not statistically significant between the two arms (p=0.75).

Kidambi (2018) randomized 740 patients undergoing screening and surveillance for colorectal neoplasia to receive colonoscopies with i-scan or with standard high-definition white-light. Endoscopists were permitted to switch between i-scan and high-definition white-light imaging to confirm polyps. Polyps were collected and analyzed by histology. The primary outcome was adenoma detection rate (ADR, proportion of subjects with at least one adenoma of any size). Intention to treat and per-protocol analyses were performed. ADR was significantly higher in the i-scan group for both the intent to treat and per-protocol analyses, with values of 47.2% and 47.6% in the i-scan group and 37.7% and 37.2% in the standard group, respectively. However, there was inconsistency across endoscopists. Secondary analyses showed that increased ADR was associated with improved detection of diminutive flat adenomas in the right colon. The groups had significantly different rates of neoplasia detection (i-scan, 56.4%; standard, 46.1%; p=0.005), but not detection of sessile serrated polyps.

Nonrandomized Studies

In 2016, Albrecht assessed the sensitivity, specificity, and positive and negative predictive values of i-scan. A total of 298 images of colonic lesions were assessed by endoscopists after undergoing a dedicated training. The sensitivity was 94.2% and the specificity was 90.9%. The positive predictive value was 87.5% and the negative predictive value was 95.9%. The intraobserver agreement was 0.9301.
In 2014, a large study using modified back-to-back designs in patients undergoing screening colonoscopy was conducted by Chung in South Korea, and included 1650 adults at average risk of CRC, who were randomly divided across three groups. During the colonoscopy, the endoscope was fully inserted and each of three colonic segments (ascending, transverse, descending) was inspected twice during withdrawal. Participants received first withdrawal with narrow-band imaging (NBI), virtual chromoendoscopy using FICE, or white-light colonoscopy (n=550 each group). White light was used in all groups for the second inspection. Ninety-one patients (5.5%) were excluded from analysis due to inadequate bowel preparation. For the primary outcome of adenoma detection rate, no statistically significant difference was found among the three groups. The percentage of patients with at least one adenoma was 24.5% in the NBI group, 23.6% in the FICE group, and 25.3% in the white-light group (p=0.75).

Moreover, the mean number of adenomas per patient was 0.35 in the NBI group, 0.36 in the FICE group, and 0.37 in the white-light group (p=0.59). The adenoma miss rate, defined as an adenoma identified only during the second inspection, was 22.9% in the NBI group, 26.0% in the FICE group, and 20.8% in the white-light-only group; a difference that was not statistically significant (p=0.30). The mean size of the missed adenomas was 3.6 mm, which was smaller than the mean size of adenomas found during the first withdrawal, which was 4.4 mm.

A study using a modified back-to-back colonoscopy design was published in 2012 by Kiriyama in Japan. The study included 102 consecutive patients with increased risk of colon cancer who received virtual chromoendoscopy using FICE and white-light colonoscopy in random order. Patients were eligible for study inclusion if they had been referred for a colonoscopy following sigmoidoscopy or for postoperative surveillance after anterior resection. Those with known IBD, bleeding, and polyposis syndrome were excluded; the right-sided colon was examined in the remaining patients. All lesions identified on either examination were removed, and specimens were sent for evaluation. Two patients were excluded from the analysis because insertion was not possible, leaving 100 patients in the analysis. A total of 110 lesions were detected. Of these, 65 lesions were detected using FICE and 45 with white light; the difference in the number of detected lesions did not differ significantly between groups. Most of the lesions detected were neoplastic; of these, 59 (91%) were found using FICE and 38 (84%) were found with white-light colonoscopy. The miss rate was defined as the proportion of total lesions in that grouping that were detected on the second examination. The miss rate for all polyps with FICE (12/39 lesions [31%]) was significantly less than that with white light (28/61 lesions [46%]) (p=0.03). Twenty-six of 59 (44%) neoplastic lesions detected by FICE and 14 of 38 (37%) of neoplastic lesions detected by white-light colonoscopy were at least 5 mm in size. For neoplastic lesions larger than 5 mm, there was no statistically significant difference between the FICE and white-light examinations in terms of the number of lesions detected.

In 2010, Cha evaluated South Korean patients at increased risk of CRC due to a personal history of polyps or gastrointestinal symptoms. A total of 135 patients underwent colonoscopy, and seven were excluded due to poor bowel preparation or diagnosis of colon cancer or intestinal disease. Thus, 128 patients were randomized to white-light colonoscopy (n=65) or virtual chromoendoscopy with FICE (n=63). The overall percentage of adenomas and the overall number of polyps did not differ significantly between groups. A total of 31 patients (49.2%) in the FICE group and 23 (35.4%) in the white-light group were found to have one or more adenomas (p=0.12). The mean number of adenomas identified per patient was also similar between groups: 1.39 in the FICE group and 1.96 in the white-light group (p=0.46). The number of adenomas less than 5 mm in size (the primary study outcome) differed significantly between groups. A total of 28 (44.4%) of patients in the FICE group and 14 (21.5%) in the white-light group (p=0.006) were found to have adenomas between 0 and 5
mm. All adenomas identified were low grade and no complications were reported in either group.

A 2010 study by Chung included 359 asymptomatic patients receiving screening colonoscopies.[53] All received back-to-back examinations with white-light colonoscopy or FICE in random order (n=181 received white light first, n=178 received FICE first). In the initial colonoscopy, a total of 60 (33.7%) of patients in the FICE group and 55 (30.4%) in the white-light group were found to have at least one adenoma; the difference between groups was not statistically significant (p=0.74). The adenoma miss rate was 6.6% in the FICE group and 8.3% in the white-light group; the difference in miss rates was not statistically significant (p=0.59). All of the missed adenomas were low grade and nonpedunculated. All but one (which was 6 mm) were 5 mm or less in size. In both Chung studies, virtual chroendoendoscopy was not found to improve the rate of adenoma detection compared with white-light endoscopy and did not identify more large adenomas.

A 2009 industry-supported multicenter RCT by Pohl in Germany compared FICE and targeted standard chroendoendoscopy using indigo carmine stain.[54] The study enrolled 871 patients presenting for screening (57%) or diagnostic (43%) colonoscopy. All patients were examined using high-resolution zoom endoscopes. Patients in the group receiving standard chroendoendoscopy underwent withdrawal using white-light colonoscopy. Indigo carmine was applied using a spray catheter through the working channel of the colonoscope for further assessment of any lesions that were identified. In the FICE group, withdrawal was performed using FICE at the preset for examining colorectal mucosa. Data were available for analysis on a total of 764 patients (368 in the FICE group, 396 in the standard chroendoendoscopy group); 107 patients were excluded for poor bowel preparation, incomplete colonoscopy, or incomplete documentation. A total of 131 (35.6%) patients in the FICE group and 140 (35.4%) patients in the standard chroendoendoscopy group had at least one adenoma; the difference between groups was not statistically significant (p=1.0). The number of small adenomas (here defined as no more than 10 mm) did not differ significantly between groups (p=0.41). The proportion of large adenomas greater than 10 mm identified in the two groups was not reported. The proportion of patients with carcinoma was small in both groups and did not differ significantly; 12 (3.3%) in the FICE group and 12 (3.0%) in the standard chroendoendoscopy group (p=0.85).

**PRACTICE GUIDELINE SUMMARY**

**NATIONAL COMPREHENSIVE CANCER NETWORK**

The National Comprehensive Cancer Network (NCCN) guidelines for colorectal cancer screening (v2.2021) recommend surveillance for individuals with a personal history of ulcerative colitis or Crohn’s colitis eight years after onset of symptoms using chroendoendoscopy with targeted biopsy or high-definition white light endoscopy.[55] Chromoendoscopy is also recommended following incomplete endoscopic resection and invisible dysplasia confirmed by a GI pathologist, although this latter recommendation is consensus-based.

**U.S. MULTI-SOCIETY TASK FORCE ON COLORECTAL CANCER**

This consensus-based guideline on colonoscopy surveillance after screening and polypectomy, published in 2012, stated that chroendoendoscopy and narrow-band imaging may enable endoscopists to accurately determine if lesions are neoplastic, and if there is a need to remove them and send specimens to pathology. The guideline noted that, at this point, these technologies have not been studied in surveillance cohorts and therefore do not have an
impact on surveillance interval.[56] The task force published evidence based recommendations for colorectal cancer screening in 2017.[57] These recommendations do not include in vivo analysis of colorectal polyps.

AMERICAN GASTROENTEROLOGICAL ASSOCIATION

In 2008, the American Gastroenterological Association (AGA) published a technology assessment of image-enhanced endoscopy, which mentions optical and electronic devices potentially playing a role in colon screening in the future, but currently, more data are needed.[58] In a 2010 position statement regarding diagnosis of colorectal neoplasia in patients with inflammatory bowel disease, the AGA stated, “Additional studies are needed to evaluate the efficiency of other imaging methods, such as narrow band imaging and confocal endomicroscopy, in detecting dysplasia.”[59]

AMERICAN COLLEGE OF GASTROENTEROLOGY

In 2018, the American College of Gastroenterology (ACG) published an evidence based clinical guideline on the management of Crohn’s Disease in adults.[60] The guideline makes the following statements regarding adjunct colonoscopy technologies:

- In patients at particularly high risk for colorectal neoplasia (e.g., personal history of dysplasia, primary sclerosing cholangitis), chromoendoscopy should be used during colonoscopy, as it may increase the diagnostic yield for detection of colorectal dysplasia, especially compared with standard-definition white light endoscopy (conditional recommendation, low level of evidence).
- For patients undergoing surveillance colonoscopy there is insufficient evidence to recommend universal chromoendoscopy for IBD colorectal neoplasia surveillance if the endoscopist has access to high-definition white light endoscopy (conditional recommendation, moderate level of evidence)
- Narrow-band imaging should not be used during colorectal neoplasia surveillance examinations for Crohn’s disease (conditional recommendation, very low level of evidence)

In 2019, the ACG published evidence-based clinical guidelines on the management of Ulcerative Colitis in adults.[61] The guidelines make the following statements regarding adjunct colonoscopy technologies:

- When using standard-definition colonoscopes in patients with UC undergoing surveillance, we recommend dye spray chromoendoscopy with methylene blue or indigo carmine to identify dysplasia (strong recommendation, low quality of evidence).
- When using high-definition colonoscopes in patients with UC undergoing surveillance, we suggest white-light endoscopy with narrow-band imaging or dye spray chromoendoscopy with methylene blue or indigo carmine to identify dysplasia (conditional recommendation, low quality of evidence).

SUMMARY

More research is needed to know whether in vivo assessment of colorectal lesions (including polyps) using various imaging systems as adjuncts to colonoscopy improves health
There is not enough research to show whether there would be an improvement in the selection of polyps for removal during colonoscopy. Therefore, in vivo analysis of colorectal lesions using any system is considered investigational.

REFERENCES


17. LC Sabbagh, L Reveiz, D Aponte, S de Aguiar. Narrow-band imaging does not improve detection of colorectal polyps when compared to conventional colonoscopy: a randomized controlled trial and meta-analysis of published studies. *BMC gastroenterology.* 2011;11:100. PMID: 21943365


22. DK Rex, CC Helbig. High yields of small and flat adenomas with high-definition colonoscopes using either white light or narrow band imaging. *Gastroenterology.* 2007;133(1):42-7. PMID: 17631129


These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>44799</td>
<td>Unlisted procedure, small intestine</td>
</tr>
<tr>
<td></td>
<td>45399</td>
<td>Unlisted procedure, colon</td>
</tr>
<tr>
<td></td>
<td>88375</td>
<td>Optical endomicroscopic image(s), interpretation and report, real-time or referred, each endoscopic session</td>
</tr>
</tbody>
</table>

**Codes**

**Date of Origin:** June 2002

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Medical Policy Manual

Radioembolization, Transarterial Embolization (TAE), and Transarterial Chemoembolization (TACE)

Effective: October 1, 2021

Next Review: July 2022
Last Review: August 2021

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Radioembolization, transarterial embolization (TAE), and transarterial chemoembolization (TACE) involve delivery of small radioactive, chemotherapeutic, or inert beads for treatment of various conditions.

MEDICAL POLICY CRITERIA

I. Radioembolization may be considered medically necessary for any of the following:
   A. Locations other than the liver; or
   B. Primary or metastatic liver tumors, when any of the following are met:
      1. Unresectable primary liver tumors (hepatocellular carcinoma [HCC]); or
      2. As a bridge to transplantation in primary HCC; or
      3. Unresectable hepatic metastases from neuroendocrine or colorectal tumors, or melanoma when any of the following are met:
         a. Neuroendocrine tumors (carcinoid and noncarcinoid) when both of the following criteria (i. and ii.) are met:
i. The disease is liver-dominant and diffuse (defined as tumor tissue spread throughout the affected organ) and symptomatic; and

ii. Systemic therapy has failed to control symptoms, or the patient is not a candidate for systemic therapy.

b. Colorectal tumors, including but not limited to adenocarcinoma when both of the following criteria (i. and ii.) are met:

i. The disease is liver-dominant, progressive, and diffuse (diffuse is defined as tumor tissue spread throughout the affected organ); and

ii. The patient is refractory to or not a candidate for chemotherapy.

c. Melanoma (ocular/uveal or cutaneous) when the disease is liver-dominant, progressive, and diffuse.

4. Unresectable primary intrahepatic cholangiocarcinoma.

II. Transarterial embolization (TAE) with non-radioactive agents may be considered medically necessary for any indication.

III. Transarterial chemoembolization (TACE) may be considered medically necessary for any indication.

IV. Radioembolization for the treatment of primary and metastatic tumors of the liver is considered investigational for all other scenarios not meeting the policy criteria above.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

POLICY GUIDELINES

Neuroendocrine tumors are rare, slow-growing, hormone-secreting tumors that may occur in numerous locations in the body.[1] Neuroendocrine tumors include the following:

- Carcinoid Tumors
- Islet Cell Tumors (also known as Pancreatic Endocrine Tumors)
- Neuroendocrine Unknown Primary
- Adrenal Gland Tumors
- Pheochromocytoma/paraganglioma
- Poorly Differentiated (High Grade or Anaplastic)/Small Cell
- Multiple Endocrine Neoplasia, Type 1 (also known as MEN-1 syndrome or Wermer's syndrome)
- Multiple Endocrine Neoplasia, Type 2 a or b (also known as pheochromocytoma and amyloid producing medullary thyroid carcinoma, PTC syndrome, or Sipple syndrome)

Neuroendocrine tumors may also be referred to by their location (e.g., pulmonary neuroendocrine tumors; gastroenteropancreatic neuroendocrine tumors)

Some appendiceal carcinoids, also called adeno carcinoids, goblet cell carcinoids or crypt cell carcinoids, have mixed histology, including elements of adenocarcinoma. While these biphasic tumors have both neuroendocrine and adenocarcinoma components, the National Comprehensive Cancer Network (NCCN) recommends they be managed according to colon cancer guidelines.
LIST OF INFORMATION NEEDED FOR REVIEW

REQUIRED DOCUMENTATION:

The information below must be submitted for review to determine whether policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

For requests pertaining to primary or metastatic liver tumors:

- Description of the planned therapy including the approach and the embolization agent to be used
- Specific description of the disease including the following:
  - Tumor type (primary vs. metastatic)
  - Extent and location of disease including whether the tumor is liver-dominant, progressive, and diffuse, and the presence or absence of extra-hepatic disease
  - For neuroendocrine metastases, description of the presence or absence of tumor-related symptoms
- Rationale for the determination that the patient is not a surgical candidate or the tumor is unresectable
- Prior treatments, if any, and tumor response
- Rationale for the determination that the patient is not a candidate for initial or continued systemic therapy
- For treatment of hepatocellular carcinoma, specify if whether treatment is proposed as a bridge to transplantation

CROSS REFERENCES

1. Charged-Particle (Proton) Radiotherapy, Medicine, Policy No. 49
2. Radiofrequency Ablation (RFA) of Tumors Other than Liver, Surgery, Policy No. 92
3. Cryosurgical Ablation of Miscellaneous Solid Tumors, Surgery, Policy No. 132
4. Magnetic Resonance (MR) Guided Focused Ultrasound (MRgFUS) and High Intensity Focused Ultrasound (HIFU) Ablation, Surgery, Policy No. 139
5. Microwave Tumor Ablation, Surgery, Policy No. 189
6. Ablation of Primary and Metastatic Liver Tumors, Surgery, Policy No. 204

BACKGROUND

TRANSARTERIAL EMBOLIZATION

According to the National Cancer Institute, transarterial embolization is defined as:

A procedure in which the blood supply to a tumor or an abnormal area of tissue is blocked. During transarterial embolization, a small incision (cut) is made in the inner thigh and a catheter (thin, flexible tube) is inserted and guided into an artery near the tumor or abnormal tissue. Once the catheter is in place, small particles made of tiny gelatin sponges or beads are injected. This blocks the artery and stops the flow of blood to the tumor or abnormal area of tissue. Transarterial embolization is used to treat some types of liver cancer, kidney cancer, and neuroendocrine tumors. It may also be used to treat uterine fibroids, aneurysms, and other conditions. Also called arterial embolization and TAE.

Types of transarterial embolization include bland embolization, chemoembolization, and
radioembolization. This policy is predominantly focused on information and evidence regarding radioembolization, which is also a form of radiation therapy.

Transarterial embolization (TAE) with non-radioactive (bland embolization) agents and transarterial chemoembolization (TACE) are also used to treat some types of cancer and other conditions, including uterine artery embolization for the treatment of fibroids. These techniques may be considered medically necessary.

RADIOEMBOLIZATION

Radioembolization, formerly referred to as selective internal radiation therapy or “SIRT”, is the intra-arterial delivery of small beads (microspheres) impregnated with yttrium-90 via the bloodstream. This technique is used to treat cancer – most commonly cancer in the liver, which is the focus of this policy. In treating cancer in the liver, the microspheres, which become permanently embedded, are delivered to tumor preferentially to normal liver, as the hepatic circulation is uniquely organized, whereby tumors greater than 0.5 cm rely on the hepatic artery for blood supply while normal liver is primarily perfused via the portal vein. Yttrium-90 is a pure beta-emitter with a relatively limited effective range and short half-life that helps focus the radiation and minimize its spread. Radioembolization is generally reserved for patients with adequate functional status (ECOG 0-2), adequate liver function and reserve, Child Pugh score A or B, and liver-dominant metastases. Candidates for radioembolization are initially examined by hepatic angiogram to identify and map the hepatic arterial system, and at that time, a mixture of albumin particles is delivered via the hepatic artery to simulate microspheres. After, single-photon emission CT gamma imaging is used to detect possible shunting of the albumin particles into gastrointestinal or pulmonary vasculature.

Hepatic tumors can arise either as primary liver cancer or by metastasis to the liver from other organs. Potentially curative local treatments include surgical resection with tumor-free margins, liver transplantation, ablative techniques, and external-beam radiation therapies. Unfortunately, most hepatic tumors are unresectable at diagnosis, due either to their anatomic location, size and number of lesions, concurrent nonmalignant liver disease, or insufficient hepatic reserve.

The use of external beam radiotherapy, 3-D or more advanced radiotherapy approaches such as intensity-modulated radiotherapy (IMRT) may be of limited use in patients with diffuse, multiple lesions due to the low tolerance of normal liver to radiation compared to the higher doses of radiation needed to kill the tumor.

Various nonsurgical and non-external irradiation based ablative techniques have been investigated that seek to cure or palliate unresectable hepatic tumors by improving locoregional control. These techniques rely on extreme temperature changes, particle and wave physics (microwave or laser ablation), or arterial embolization therapy including chemoembolization, bland embolization, or radioembolization.

UNRESECTABLE PRIMARY LIVER CANCER [HEPATOCELLULAR CARCINOMA (HCC)]

The majority of patients with HCC present with unresectable disease and treatment options are limited secondary to the chemoresistance of HCC and the intolerance of normal liver parenchyma to tumoricidal radiation doses.

Other Treatment Options
• Radioembolization. In general, radioembolization is used for unresectable HCC that is greater than 3 cm.
• Transarterial chemoembolization (TACE) therapy. Results of two randomized controlled trials have shown a survival benefit using TACE versus supportive care in patients with unresectable HCC.[3, 4]
• Transarterial embolization (TAE). In one study, patients were randomly assigned to TACE, TAE, or supportive care. One-year survival rates for TACE, TAE, and supportive care were 82%, 75%, and 63%, respectively and two-year survival rates were 63%, 50%, and 27%, respectively.
• Targeted therapies. A 2007 multicenter, randomized, double-blind placebo controlled Phase III trial that enrolled 602 patients with advanced HCC randomly assigned patients to receive sorafenib versus placebo.[5] Overall survival (OS) was significantly longer in the sorafenib group compared with placebo (10.7 versus 7.9 months, respectively; hazard ratio for sorafenib: 0.69; p<0.001).

UNRESECTABLE METASTATIC COLORECTAL CARCINOMA

The role of local (liver-directed) therapy (including radioembolization, chemoembolization, and conformal radiation therapy) for complete tumor removal or destruction is widely accepted in clinical practice. Incomplete “debulking” of unresectable metastatic disease in the liver remains controversial.[6]

Fifty to sixty percent of patients with colorectal cancer develop metastases, either synchronously or metachronously. Emphasis on treating patients with potentially curable disease is on complete destruction or removal of all tumor tissue. The majority of patients diagnosed with metastatic colorectal disease are initially classified as having unresectable disease.

Other Treatment Options

• In patients with metastatic disease limited to the liver, preoperative chemotherapy is sometimes used in an attempt to downsize the metastases in order to convert the metastatic lesions to a resectable status (conversion chemotherapy).
• In patients with unresectable disease that cannot be converted to resectable disease, the primary treatment goal is palliative, with survival benefit shown with both second and third-line systemic chemotherapy.[7]
• Advances in chemotherapy have doubled the median survival in this population from less than one year to more than two years.
• Palliative chemotherapy by combined systemic and hepatic artery infusion therapy (HAI) may increase disease-free intervals for patients with unresectable hepatic metastases from colorectal cancer.
• Ablation techniques (see Cross References)
• Radiation therapy (see Cross References).

UNRESECTABLE METASTATIC NEUROENDOCRINE TUMORS

Neuroendocrine tumors are an uncommon, heterogeneous group of mostly slow-growing, hormone-secreting malignancies, with an average patient age of 60 years. Primary neuroendocrine tumors vary in location, but most are either carcinoids (which most commonly arise in the midgut) or pancreatic islet cells. Carcinoid tumors, particularly if they metastasize
to the liver, can result in excessive vasoactive amine secretion including serotonin and are commonly associated with the carcinoid syndrome (diarrhea, flush, bronchoconstriction, and right valvular heart failure).

Although they are considered to be indolent tumors, at the time of diagnosis, up to 75% of patients have liver metastases. The five-year survival rates with metastases to the liver are less than 20%. Less than 10% of patients are eligible for resection as most patients have diffuse, multiple lesions.

Conventional therapy is largely considered to be palliative supportive care to control, eradicate, or debulk hepatic metastases, often to palliate carcinoid syndrome or local pain from liver capsular stretching.

Other Treatment Options

- Medical treatment includes somatostatin analogs, like octreotide or lanreotide, or systemic chemotherapy. Although patients often achieve symptom relief with octreotide, the disease eventually becomes refractory, with a median duration of symptom relief of approximately 13 months, with no known effect on survival. Systemic chemotherapy for these tumors has shown modest response rates of limited duration, is better for pancreatic neuroendocrine tumors compared to carcinoids, and is frequently associated with significant toxicity.\[8\]
- Radiofrequency or cryosurgical tumor ablation (see Cross References)
- Transarterial chemoembolization (TACE) therapy. Chemoembolization has shown response rates of nearly 80%, but the effect is of short duration and a survival benefit has not been demonstrated.\[8\]
- TAE with non-radioactive agents
- Radiation therapy (see Cross References)

UNRESECTABLE INTRAHEPATIC CHOLANGIOCARCINOMA

Cholangiocarcinomas are tumors that arise from the epithelium of the bile duct and are separated into intrahepatic and extrahepatic types. Intrahepatic cholangiocarcinomas appear in the hepatic parenchyma and are also known as peripheral cholangiocarcinomas.\[9\] Resection is the only treatment with the potential for cure and five-year survival rates have been in the range of 20% to 43%.

Other Treatment Options

Patients with unresectable disease may select among fluoropyrimidine-based or gemcitabine-based chemotherapy, fluoropyrimidine chemoradiation, or best supportive care.

MISCELLANEOUS METASTATIC TUMORS

Small case reports have been published on the use of radioembolization in many other types of cancer with metastases, including breast, head, and neck (including parotid gland), pancreaticobiliary, anal, thymic, thyroid, endometrial, lung, kidney, gastric, small bowel, esophageal, ovarian, cervical, prostatic, bladder, and for melanoma, sarcoma and lymphoma.\[10\]

REGULATORY STATUS
Currently, two commercial forms of yttrium-90 microspheres are available: a glass sphere, TheraSphere® (MDS Nordion, Inc. used under license by BTG International) and a resin sphere, SIR-Spheres® (Sirtex Medical Limited). Noncommercial forms are mostly used outside the U.S. While the commercial products use the same radioisotope (yttrium-90) and have the same target dose (100 Gy), they differ in microsphere size profile, base material (i.e., resin vs. glass), and size of commercially available doses. These physical characteristics of the active and inactive ingredients affect the flow of microspheres during injection, their retention at the tumor site, spread outside the therapeutic target region, and dosimetry calculations.

Note also that the U.S. Food and Drug Administration (FDA) granted premarket approval of SIR-Spheres® for use in combination with 5-flourouridine (5-FUDR) chemotherapy by HAI to treat unresectable hepatic metastases from colorectal cancer. In contrast, TheraSphere® was approved by humanitarian device exemption (HDE) for use as monotherapy to treat unresectable hepatocellular carcinoma (HCC). In January 2007, this HDE was expanded to include patients with hepatocellular carcinoma who have partial or branch portal vein thrombosis. For these reasons, results obtained with one product do not necessarily apply to other commercial (or noncommercial) products.

**EVIDENCE SUMMARY**

This evidence review does not include summaries for transarterial embolization (TAE) with non-radioactive or transarterial chemoembolization (TACE), which may be considered medically necessary.

The principal health outcomes associated with treatment of malignancies are typically measured in units of survival past treatment: disease-free survival (DFS), a period of time following treatment where the disease is undetectable; progression-free survival (PFS), the duration of time after treatment before the advancement or progression of disease; and overall survival (OS), the period of time the patient remains alive following treatment.

In order to understand the impact of radioembolization (RE) on these outcomes, well-designed randomized controlled trials (RCTs) are needed that compare this therapy with standard medical and/or surgical treatment of tumors in the liver.

**RADIOEMBOLIZATION FOR UNRESECTABLE PRIMARY LIVER CANCER [HEPATOCELLULAR CARCINOMA (HCC)]**

The following literature review on RE for unresectable HCC focused on systematic literature reviews and comparative studies (randomized and nonrandomized).

**Systematic Reviews**

Various meta-analyses have been performed to compare the effects of TACE, drug-eluting bead (DEB) plus TACE (DEB-TACE), and RE in patients with unresectable HCC, each of which performed slightly different analyses (e.g., pairwise vs indirect comparisons and assessment of different outcomes or comparator groups). Results of these meta-analyses are summarized below.

Abdel-Rahman (2020) conducted a meta-analysis of RCTs comparing RE alone or combined with other systemic or locoregional treatments to placebo, no treatment, or other similar interventions in patients with unresectable HCC.[11] Six RCTs (total n=1340) were identified, all of which were assessed by authors as being at high risk of bias. The authors reported the
certainty of evidence as low to very low. Meta-analysis was able to be performed using data from more than one RCT for few comparisons. Based on meta-analysis of two RCTS, disease control rate was not significantly different between RE and sorafenib (relative risk [RR], 0.94; 95% confidence interval [CI] 0.84 to 1.05), though RE was associated with less hand-foot skin reactions (RR, 0.02; 95% CI 0.00 to 0.06), skin rash (RR, 0.11; 95% CI 0.04 to 0.34), diarrhea (RR, 0.11, 95% CI 0.04 to 0.34), and hypertension (RR, 0.10; 95% CI 0.01 to 0.86). Based on meta-analysis of three RCTs, the risk of serious adverse events did not differ between RE and TACE (RR, 1.47; 95% CI 0.66 to 3.25). Meta-analysis could not be performed for other comparisons; thus, results of other included trials are described individually in the section below on RCTs.[12, 13]

Venerito (2020) performed a meta-analysis to assess the noninferiority of SIRT as monotherapy or followed by sorafenib versus sorafenib monotherapy on OS.[14] A noninferiority margin of 1.08 in terms of hazard ratio (HR) was prespecified. Three RCTs were included (total n=1,243), and meta-analysis demonstrated SIRT with or without sorafenib was noninferior to sorafenib monotherapy in OS (median, 10.2 and 9.2 months; HR, 0.91; 95% CI 0.78 to 1.05). Treatment-related severe adverse events were reported in 28.9% vs 43.3% of patients treated with SIRT and sorafenib monotherapy, respectively (p<0.01).

Yang (2020) conducted a meta-analysis of RCTs to compare effects of DEB-TACE, TACE, and RE on the primary outcome of overall survival.[15] Compared with TACE, RE was associated with similar one-year OS (RR, 0.91; 95% CI 0.79 to 1.05), but a better OS than TACE at two years (RR, 0.87; 95% CI 0.80 to 0.95) and three years (RR, 0.90; 95% CI 0.85 to 0.96). Overall survival was not significantly different between RE and DEB-TACE at one year (RR, 0.83, 95% CI 0.68 to 1.02), but DEB-TACE was associated with better OS at two years than RE (RR, 0.40; 95% CI 0.19 to 0.84). However, pooled HRs indicated that RE was superior to TACE in overall survival (HR, 0.84; 95% CI 0.70 to 1.00) and that DEB-TACE was superior to RE in overall survival (HR, 0.59; 95% CI 0.38 to 0.91).

Tao (2017) reported on a network meta-analysis comparing nine minimally invasive surgeries for treatment of unresectable hepatocellular carcinoma (HCC).[16] The interventions included were transarterial chemoembolization (TACE), TACE plus sorafenib, sorafenib, TACE plus high-intensity focused ultrasound, TACE plus percutaneous ethanol injection, drug-eluting bead (DEB) plus TACE (DEB-TACE), yttrium-90 radioembolization (90Y RE), TACE plus external-beam radiation therapy (EBRT), and ethanol ablation. The network included 17 studies with 2,669 patients and four studies with 230 patients including 90Y RE. In a pairwise meta-analysis, patients treated with 90Y RE were more likely to achieve complete remission than those who received TACE (odds ratio [OR], 4.5; 95% CI 1.3 to 15.1). However, in the network meta-analysis, there was no significant difference between the corresponding eight treatments and TACE with respect to complete remission, partial response, stable disease, and objective response rate. The treatments were ranked for several outcomes using surface under the cumulative ranking curves (SUCRA). TACE plus EBRT had the highest SUCRA ranking in complete remission (77%), partial response (89%), progressive disease (95%), and objective response rate (81%).

Ludwig (2017) conducted an indirect meta-analysis of studies that indirectly compared DEB-TACE with 90Y RE for HCC.[17] Fourteen studies (total N=2065 patients) comparing DEB-TACE or 90Y RE with conventional TACE for primary HCC treatment were included. The pooled estimate of median survival was 23 months for DEB-TACE and 15 months for RE. The estimated one-year survival was significantly higher for DEB-TACE (79%) than for RE (55%;
OR=0.57; 95% CI 0.36 to 0.92; p=0.02). Survival did not differ statistically significantly at two or three years but did favor DEB-TACE. At two years, survival was 61% for DEB-TACE and 34% or RE (OR=0.65; 95% CI 0.29 to 1.44; p=0.29) and at three years survival was 56% and 21% (OR=0.71; 95% CI 0.21 to 2.55; p=0.62), respectively.

Two systematic reviews published in 2016 compared RE with TACE for the treatment of unresectable HCC. Lobo (2016) selected five retrospective observational studies (total n=533 patients).[18] Survival at one year did not differ statistically between RE (42%) and TACE (46%; RR, 0.93; 95% CI 0.81 to 1.08; p=0.33). At two years, the survival rate was higher for RE (27% vs 18%; RR=1.36; 95 CI 1.05 to 1.76; p=0.02), but there was no statistically significant difference in survival rates at three, four, or five years. Postprocedural complications were also similar in the two groups. Facciorusso (2016) included 10 studies (total n=1,557 patients), two of which were randomized controlled trials (RCTs).[17] The OR for survival was not statistically significant at one year (OR=1.0; 95% CI 0.8 to 1.3; p=0.93) but favored RE in years two (OR=1.4; 95% CI 1.1 to 1.90; p=0.01) and three (OR=1.5; 1.0 to 2.1; p=0.04).

Vente (2009) conducted a meta-analysis evaluating tumor response and survival in patients who received yttrium-90 glass or resin microsphere radioembolization for the treatment HCC or metastases from colorectal cancer (CRC).[19] (See below under unresectable metastatic CRC section for the data from the meta-analysis as pertains to that disease.) Included studies were from 1986 onward and presented tumor response measured by CT scans and data on median survival times. To allow comparability of results with regard to tumor response, the category of “any response” was introduced, and included complete response, partial response, and stable disease. Overall tumor response could only be assessed as any response because response categories were not uniformly defined in the analyzed studies.

In 14 articles, clinical data were presented on tumor response and survival for 425 patients with HCC who had received yttrium-90 radioembolization. Treatment with resin microspheres was associated with a significantly higher proportion of any response than glass microsphere treatment (0.89 vs. 0.78, respectively; p=0.02). Median survival was reported in seven studies in which survival time was defined as survival from microsphere treatment or from diagnosis or recurrence of HCC. Median survival from microsphere treatment varied between 7.1 and 21.0 months, and median survival from diagnosis or recurrence was 9.4 to 24.0 months.

The authors of the meta-analysis concluded that yttrium-90 radioembolization is associated with high response rates, both in salvage and first-line settings, but that the true impact on survival will only become known after publication of several ongoing and/or to-be-initiated Phase III studies, as well as the results of trials in which yttrium-90 radioembolization and modern chemotherapy agents are combined with novel biologic agents.

In May 2013 a comparative effectiveness review of local therapies (i.e., ablation, embolization, and radiotherapy) for patients with unresectable HCC was conducted by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center for the Agency for Healthcare Research and Quality (AHRQ).[20] The review sought to report on overall survival and quality of life outcomes and adverse events. Transplant candidates were excluded from this review. Three prospective case series and one retrospective case series with a total of 187 participants met inclusion criteria for review. There were no randomized controlled trials and no comparative trials that met inclusion criteria. Therefore, the strength of evidence was rated as insufficient to evaluate the outcomes of interest. One study reported a one-year survival rate of 75%; three studies reported a median survival range of 11 to 15
months. Quality of life, local recurrence, and disease progression were not reported in any of the included studies. Adverse events were rare and no liver failure or hepatic abscess was reported. The authors recommended studies that compare various embolization techniques including radioembolization.

**Randomized Controlled Trials**

In 2014, Kolligs reported results of a small pilot randomized controlled trial (RCT) comparing RE with TACE for the treatment of unresectable HCC, the SIR-TACE study. The study included 28 subjects with unresectable HCC, preserved liver function, and an Eastern Cooperative Oncology Group [ECOG] Performance Status of 2 or less, with no vascular invasion or extrahepatic spread, who had five or fewer liver lesions or a single lesion of 10 cm or less. Patients were randomized to RE (n=13) or TACE (n=15). Over posttreatment follow up, PR rates were 13.3% for TACE and 30.8% for RE, with rates of disease control (CR, SD, PR) of 73.3% for TACE and 76.9% for RE. Median progression-free survival (PFS) was 3.6 months for TACE and 3.7 months for RE.

In 2014, Pitton reported results from a small RCT comparing RE with TACE with drug eluting beads TACE (DEB-TACE) for the treatment of unresectable HCC. The study included 24 patients, 12 randomized to each group. No deaths occurred within 30 days of the procedure for either group. There were no statistically significant differences between the groups in terms of PFS (180 days for RE vs 216 for TACE; p=0.619) and overall survival (OS; 592 days for RE vs 788 for TACE; p=0.927).

**Nonrandomized Comparison Studies**

A propensity score matching analysis reported by Martelletti (2021) compared patient outcomes between transarterial radioembolization (TARE) and sorafenib. HCC patients (total n=65) were treated with TARE (n=41) or sorafenib (n=24). Downstaging to curative-intent surgery occurred in 10 of 41 TARE patients and one of 24 sorafenib patients. In the non-downstaged patients, median survival was 20.3 in the TARE patients and 9.1 months in the sorafenib patients (p=0.0001), and one-, two-, and three-year OS rates were 64.5%, 42.6% and 37.3%, respectively, in the TARE patients and 39.1%, 13.0% and 0%, respectively, in the sorafenib patients. Propensity score and Bayesian model averaging analyses indicated that there was an improvement in overall survival in the TARE group compared with sorafenib treatment.

Bekki (2021) reported a comparative study of portal vein embolization versus radiation lobectomy before resection of hepatocellular carcinoma in chronic liver disease patients. A total of 73 patients were treated with portal vein embolization and 22 with radioembolization. Additional procedures were required for tumor control in 47% of portal vein embolization patients and 27% of radioembolization patients. The degree of hypertrophy was 63% for RE and 36% for portal vein embolization (p<0.01). Resectability rate was 85% for portal vein embolization and 64% for RE (p=0.03). For 18% of patients not pursuing surgery follow RE, the reason was complete tumor control.

Facciorusso (2020) performed a retrospective analysis that compared patients with HCC treated with RE plus sorafenib (n=45) with propensity score-matched patients treated with sorafenib alone (n=90). No significant differences were identified in median OS (10 vs. 10 months; p=0.711), median PFS (six versus seven months; p=0.992), and objective response rate (45.5% versus 42.8%; p=1).
Padia (2017) reported on a single-center, retrospective study (2010-2015) comparing segmental RE with segmental chemoembolization in 101 patients with localized, unresectable HCC not amenable to ablation. Patients receiving chemoembolization had poorer ECOG Performance Status ratings and Child-Pugh class while those receiving RE had larger and more infiltrative tumors. Overall complete remission was 84% with RE and 58% with chemoembolization (p=0.001). Median PFS was 564 days and 271 days (p=0.002) and median OS was 1198 days and 1043 days (p=0.35), respectively, for the RE group and the chemotherapy group.

In 2016, Soydala reported a retrospective study comparing outcomes of patients receiving RE and TACE for HCC. Each group included 40 patients. RE patients had a mean survival of 39 months versus 31 months for TACE (p=0.014). There was no significant difference in chronic complications and recurrence of disease.

In 2016, Oladeru reported a retrospective study based on SEER registry data comparing survival outcomes of patients receiving RE and external beam radiation of HCC. A total of 189 patients with unresectable HCC (77 receiving RE, 112 external beam radiotherapy) receiving treatment between 2004 and 2011 were evaluated. Median OS for RE was 12 months versus 14 months for external beam radiotherapy. Median disease-specific survival was identical for both groups at 14 months. After adjustment for differences between patients, multivariable survival analysis showed no association of treatment and OS or disease-specific survival.

In 2015, El Fouly reported results of a nonrandomized study comparing yttrium-90 RE with TACE among 86 patients with intermediate stage, nonresectable HCC. Sixty-three patients at one institution were treated with TACE, while 53 patients at a second institution were treated with RE. Median OS in for TACE and RE was not significantly different between groups (18 months for TACE vs 16.4 months for RE); similarly median time to progression (TTP) was not significantly different between groups (6.8 months for TACE vs 13.3 months for RE). TACE patients had higher numbers of treatment sessions, hospital times, and rates of adverse events. Also in 2015, Gramenzi conducted a retrospective cohort study to compare RE with yttrium-90 with sorafenib for intermediate- or advanced-stage HCC. Patients with HCC refractory to other therapies and no metastases or systemic chemotherapy were included, 74 of whom were treated with sorafenib and 63 treated with RE. Median OS between groups was similar (14.4 months for sorafenib-treated patients vs. 13.2 months for RE-treated patients). After propensity-score matching of 32 subjects in each group, there were no significant differences in median OS or one-, two-, and three-year survival rates between groups.

**RADIOEMBOLIZATION AS A BRIDGE TO LIVER TRANSPLANTATION FOR PRIMARY HCC**

**Systematic Reviews**

Kulik (2018) published a systematic review of 18 comparative studies and 31 noncomparative studies that included patients with unresectable HCC who needed a liver transplant and received transplant alone or some type of bridging therapy as well. Of the 18 comparative studies, two studies (n=257 patients) reported on the incidence of dropout from transplantation wait-lists, and patients receiving bridging therapy. This group had reduced risk of dropout due to disease progression compared with those receiving transplantation alone (RR=0.32). Between-group differences were not statistically significant for mortality (five comparative studies; n=531 patients) or recurrence rate (10 comparative studies; n=889 patients).
Subgroup analysis was conducted for types of bridging therapy: for all-cause mortality after transplantation, the RR was 1.124 with TAE compared with transplantation alone (one cohort). For disease recurrence, the RR for this bridging therapy type was 2.374 compared with transplantation alone. No RCTs were identified, and most of the selected studies had a high risk of bias on patient selection, adequate follow-up, and funding source when reported.

**Randomized Controlled Trials**

Salem (2016) reported on results of a phase 2 RCT comparing conventional TACE and TheraSphere radioembolization (Y90) for treatment of unresectable, unreatable HCC.[14] Twenty-four patients were assigned to Y90 and 21 patients to conventional TACE; the ultimate goal of treatment for these patients was liver transplantation. The primary outcome was time to progression using intention-to-treat analysis. Median follow-up was 17 months. In the conventional TACE group, there were seven transplants at a median of nine months (range, 3 to 17 months). In the Y90 group, there were 13 transplants at a median of nine months (range, 4 to 15 months). Median time to progression exceeded 26 months in the Y90 group and 6.8 months in the conventional TACE group (hazard ratio, 0.12; 95% CI 0.03 to 0.56; p=0.007). Median survival was 19 months in Y90 and 18 months in conventional TACE (p=0.99). Adverse events were similar between groups, with the exception of more diarrhea (21% vs 0%) and hypoalbuminemia (58% vs 4%) in the conventional TACE group. A limitation of the OS analysis was the censoring of the survival outcome at liver transplantation given that transplantation is related to the treatment effect.

In 2014, Kulik reported results of a pilot RCT of yttrium-90 RE with or without sorafenib for patients with HCC awaiting liver transplantation.[28] The study randomized 23 subjects; after accounting for losses due to self-withdrawal from the study, failure to confirm HCC, and death, the modified intention-to-treat (ITT) population included 10 subjects randomized to RE alone and 10 randomized to RE with sorafenib. Overall, 17 of 20 patients underwent liver transplantation, with no difference in median time-to-transplant between groups. However, the addition of sorafenib was associated with increased peritransplant biliary complications, and acute rejection.

**Nonrandomized Studies**

Salem (2021) reported the results of the multicenter, single-arm, retrospective LEGACY trial investigating yttrium-90 radioembolization with TheraSphere for the treatment of solitary, unresectable HCC.[29] The aim of the study was to evaluate the objective response rate (ORR) and the duration of response (DOR) based on modified RECIST criteria as evaluated by blinded, independent, central review. Eligibility criteria included: solitary HCC ≤8 cm, Child-Pugh A cirrhosis, and ECOG performance status 0 to 1. Of 162 enrolled patients, 60.5% were ECOG 0 and RE served as neoadjuvant therapy for transplantation or resection in 21% and 6.8% of patients, respectively. Median follow-up duration was 29.9 months. ORR (best response) was 88.3% (95% CI, 82.4 to 92.4) with 62.2% (95% CI, 54.1 to 69.8) exhibiting a DOR ≥6 months. Three-year OS was 86.6% for all patients and 92.8% for neoadjuvant patients resected or transplanted. This study supported FDA premarket approval of TheraSphere for use in HCC.[30]

Pellegrinelli (2021) reported on an eight-year single-center experience utilizing RE for the treatment of patients with unresectable HCC (n=44), metastatic colorectal cancer (n=20), and intrahepatic cholangiocarcinoma (n=6).[31] Treatment with prior chemotherapy was reported in 48.6% of all patients, and RE-related grade 3 or higher adverse events impacted 17.1% of
patients. Patients were treated with RE as bridge to transplant (4.3%), for downstaging prior to surgical resection (15.7%), as ablative therapy (1.4%), and for palliative treatment (78.6%). Median follow-up was 32.1 months, during which disease progression occurred in 63 (90%) of all patients. Among patients with HCC at study end, complete and partial responses were achieved in 1 and 2 patients, respectively. Median OS was 16.1 months (range, 1.0 to 72.5 months) with no significant differences in survival among disease groups.

Gabr (2020) performed a retrospective review that reported on long-term outcomes of liver transplantation for patients with HCC who were bridged or downstaged with RE.[32] From 2004 to 2018, 207 patients underwent transplant after RE. Median OS from transplant was 12.5 years, with median time to liver transplantation of 7.5 months (interquartile range, 4.4 to 10.3). Overall, 169 patients were bridged and 38 were downstaged to liver transplant. Overall survival rates at 3, 5, and 10 years were 84%, 77%, and 60%, respectively.

Zori (2020) performed a retrospective cohort analysis that compared patients with HCC undergoing bridging locoregional therapy with RE (n=28) or TACE (n=37) prior to liver transplant.[33] Three-year survival was not significantly different with RE vs TACE (92.9% vs. 75.7%; p=0.052). However, microvascular invasion occurred in 3.6% versus 27% of patients treated with RE versus TACE (p=0.013).

In a 2013 retrospective review, Tohme reported on 20 consecutive HCC patients on liver transplant waiting lists who received radioembolization as bridge therapy.[34] When radioembolization began, Milan criteria (extent of disease) for liver transplantation were met by 14 patients and sustained until transplantation. Of the six patients who did not meet Milan criteria initially, radioembolization was able to downstage two patients to meet Milan criteria. Complete or partial radiologic response to radioembolization on modified Response Evaluation Criteria In Solid Tumors (RECIST) occurred in nine patients. Additionally, on pathologic examination, five patients who met Milan criteria had complete tumor necrosis with no evidence of viable tumor.

In 2014, Ramanathan reported on multimodality therapy, including radioembolization, for 715 HCC patients of which 231 were intended for transplant.[35] In the intention-to-treat with transplantation arm, 60.2% were able to receive a transplant. Survival rates posttransplant were 97.1% and 72.5% at one and five years, respectively. Tumor recurrence rates were 2.4%, 6.2%, and 11.6% at one, three, and five years, respectively. Since this study included multimodality therapy, it is not possible to isolate the effect of radioembolization.

Lewandowski (2009) compared RE with chemoembolization in the efficacy of downstaging 86 patients with HCC from stage T3 to T2 (potentially making patients liver transplant candidates).[36] Patients were treated with either radioembolization using yttrium-90 microspheres (n=43) or TACE (n=43). Median tumor size was similar between the two treatment groups (5.7 and 5.6 cm, for TACE vs. radioembolization, respectively.) Partial response rates were 61% versus 37% for radioembolization vs. TACE, respectively, with downstaging from T3 to T2 in 58% of patients treated with RE versus 31% with TACE (p<0.05).

RADIOEMBOLIZATION FOR UNRESECTABLE METASTATIC COLORECTAL CARCINOMA (CRC)

Systematic Reviews
A 2010 technology assessment\cite{7}, a 2009 Cochrane review\cite{37}, and a 2009 systematic review with meta-analysis\cite{19} all concluded that data from large Phase III trials were needed in order to fully understand the impact of radioembolization on survival in patient with CRC metastases in the liver.

Two additional systematic reviews were published in 2013:

Rosenbaum considered radioembolization, either as monotherapy or concomitant with chemotherapy, to be an emerging treatment for CRC liver metastases, with a limited amount of data from heterogeneous studies.\cite{38} This review evaluated 13 articles on radioembolization as monotherapy and 13 studies on radioembolization combined with chemotherapy for chemoresistant, unresectable CRC liver metastasis. Heterogeneity between studies prohibited pooling of data. This heterogeneity included varying patient inclusion criteria such as the amount of intrahepatic and extrahepatic tumor burden, patient performance status, previous systemic treatments, and protocols for assessing tumor response. CR, PR, and stable disease (SD) rates ranged from 29% to 90% with radioembolization alone and from 59% to 100% for radioembolization with chemotherapy. At 12 months, survival ranged from 37% to 59% with radioembolization alone and from 43% to 74% for radioembolization combined with chemotherapy. As with prior reviews, the authors concluded that additional data is needed from high-quality randomized trials.

In contrast to the prior systematic reviews, Saxena considered the evidence sufficient to recommend increased utilization of radioembolization as salvage treatment for CRC liver metastases.\cite{39} The review evaluated a total of 979 patients in 20 studies including two RCTs\cite{40, 41}. The majority of patients had previously undergone at least three lines of chemotherapy (range 2 to 5). After radioembolization, the average reported CRs and PRs from 16 studies was 0% (range, 0% to 6%) and 31% (range, 0% to 73%), respectively. The median time to intrahepatic progress was nine months (range 6 to 16 months) and the median survival time was 12 months (range 8.3 to 36 months). The mean rate of acute toxicity was 40.5% (range 11% to 100%); most cases were mild and did not require intervention. Despite concluding that radioembolization was safe and effective, the authors noted the need for continued evaluation of clinical outcomes.

**Randomized Controlled Trial**

A phase 3 RCT by van Hazel (2016) of 530 patients compared patients receiving modified FOLFOX chemotherapy and FOLFOX chemotherapy plus SIRT in patients with previously untreated liver-dominant metastatic disease.\cite{42} Bevacizumab was allowed as additional treatment at the discretion of the treating physician. About 40% of patients had extrahepatic metastases at randomization. About 28% of patients had more than 25% liver involvement of metastases. The primary end point was overall (any site) progression-free survival (PFS). Secondary end points included liver-specific outcomes such as PFS in the liver, tumor response rate, and liver resection rate. The primary end point of PFS at any site showed no difference between groups (10.2 months vs 10.6 months control vs RE; hazard ratio, 0.93; p=0.43). Secondary liver-specific end points of median PFS in the liver and objective response rate in the liver were improved in the RE group (liver PFS, 12.6 months vs 20.5 months control versus RE; liver response rate, 68.8% vs 78.7% control vs RE). Wasan (2017) analyzed OS from this study in combination with two other studies of chemotherapy with and without RE.\cite{43} Overall, 549 patients were randomly assigned to FOLFOX alone and 554 patients were
assigned FOLFOX plus SIRT. Overall survival was not significantly different between groups (HR, 1.04; 95% CI 0.90 to 1.19).

**Nonrandomized Studies**

Since the systematic reviews were published, a number of additional nonrandomized studies have reported outcomes of RE for patients with CRC liver metastases who failed or were not candidates for chemotherapy.[44-47] The majority of these were noncomparative studies which precluded conclusions on the survival benefit of RE compared to other treatments. There was a wide range of clinical response to RE; although the rate of complete response was low, partial response averaged 35% and stable disease was reported in 32 to 71% of patients. The few studies that compared RE to best supportive care reported a statistically significant survival benefit with RE. The rates of Grade 3 to 4 toxicities ranged from 0% to 39% and included absolute lymphocyte, alkaline phosphatase, bilirubin, and albumin. Factors associated with poorer prognosis included large tumor volume, poor radiological response to treatment, and the number of prior chemotherapy treatments.

The most recent comparative study, published by Mokkarala (2019), performed a propensity score-matched retrospective analysis of patients with colorectal metastases treated with DEB-TACE (n=47) or RE (n=155).[48] Extra-hepatic metastasis was more frequent with DEB-TACE (68.1% vs. 47.7%; p=0.014), as was occurrence of ≥10 liver lesions (42.2% vs. 68.8%; p=0.001). Toxicity was not significantly different between DEB-TACE and RE (27% vs. 9.1%, respectively; p=0.057). Treatment with DEB-TACE was not a prognostic factor for survival (HR, 0.94; 95% CI 0.54 to 1.65).

A 2021 study by Haber evaluated the addition of radioembolization to systemic therapy in the salvage setting for hepatic metastases from colorectal cancer.[49] Twenty-one patients who underwent radioembolization plus systemic therapy were matched with a cohort of 173 patients who received systemic chemotherapy alone in the salvage setting, defined as progression on at least two different regimens of systemic chemotherapy. The difference in median survival from the date of primary diagnosis between groups was not statistically significant (38 [95% CI 26 to 50] vs. 25 [95% CI 15 to 35] months for radioembolization with systemic therapy vs. systemic therapy alone; p=0.17). When measured from the date of hepatic metastases, median survival was 31 (95% CI 23.8 to 38.2) for those treated with radioembolization with systemic therapy compared to 20 months (95% CI 10.2 to 29.8) for those treated with systemic therapy alone (p=0.03).

**RADIOEMBOLIZATION FOR MELANOMA METASTASES IN THE LIVER**

Many studies of metastatic melanoma focus on patients with uveal melanoma in whom the liver is the most common site of metastatic disease.

**Systematic Reviews**

Rowcroft (2020) planned to perform a meta-analysis of studies of patients with liver-only metastases of uveal melanoma treated with systemic therapy, isolated hepatic perfusion, hepatic artery infusion, TACE, SIRT, and immunoembolization.[50] However, due to heterogeneity in available data, meta-analysis was not performed. The authors descriptively reported that six non-comparative retrospective cohort studies (n=150; range, 8 to 71) evaluated the use of SIRT, which reported median OS ranged from 9 to 24 months.
Randomized Controlled Trials

No randomized controlled trials were identified for radioembolization of melanoma metastases in the liver.

Nonrandomized Comparative Studies

Gonsalves (2019) performed a prospective study of patients with liver metastases of uveal melanoma treated with RE.\[^{51}\] Among patients who were treatment-naive, complete response, partial response, or stable disease was achieved in 20 of 23 patients (87.0%; 95% CI 66.4%, 97.2%), median progression-free survival from liver metastasis was 8.1 months (95% CI 6.4, 11.8), and median OS was 18.5 months (95% CI 11.3 to 23.5). Among patients who progressed after immunoembolization, complete response, partial response, or stable disease was achieved in 14 of 24 patients (58.3%; 95% CI 36.3%, 77.9%), median PFS from liver metastasis was 5.2 months (95% CI 3.7 to 9.8), and median OS was 19.2 months (95% CI 11.5 to 24.0).

In 2014, Xing conducted a retrospective observational study to compare outcomes for patients with unresectable melanoma (both uveal and cutaneous) liver metastases refractory to standard chemotherapy treated with either yttrium-90 RE (n=28) or best supportive care (n=30).\[^{52}\] The groups were similar at baseline in terms of Child-Pugh class, ECOG performance status scores, age, sex, and race. However, patients treated with RE had significantly larger tumor size at baseline than those treated with best supportive care (mean, 7.28 cm vs 4.19 cm; p=0.02). Median OS from diagnosis of melanoma liver metastases was longer in RE-treated subjects (19.9 mo. vs 4.8 mo.; p=0.000), as was the median OS from diagnosis of the primary melanoma (119.9 months vs 26.1 months; p<0.001). Pre- and post-treatment imaging studies were available for 24/28 (85.7%) of those treated with RE. Of those, no patients had a CR; five patients (17.9%) had PR, nine patients (32.1%) had SD, and 10 patients (35.7%) had PD. Two patients receiving RE had major (grade 5) clinical toxicities (ascites and hepatic encephalopathy and eventual mortality). Significant factors for longer OS were <10 metastatic liver lesions, absence of extrahepatic metastases, and Child-Pugh class A. Although this study was retrospective and included small sample sizes, it included relatively long-term follow-up and provided comparison between RE and best supportive care.

Nonrandomized Non-comparative Studies

In 2014, Eldredge-Hindy retrospectively evaluated outcomes for the use of yttrium-90 RE in 71 patients with biopsy-confirmed uveal melanoma liver metastases.\[^{53}\] The median time from the diagnosis of liver metastases to RE was 9.8 months (95% CI 7.4 to 12.2 months), and 82% of patients had received prior liver-directed therapies. Sixty-one patients (86%) had CT or magnetic resonance imaging (MRI) evaluation of treatment response at three months post-RE. Of those, five patients (8%) had a PR, 32 patients (52%) had SD, and 24 patients (39%) had DP. Median OS RE was 12.3 months (range, 1.9 to 49.3 months).

Small studies (n=8 to 32) have reported on use of RE in patients with hepatic metastases from melanoma.\[^{54-60}\] Five of the studies included only patients with ocular melanoma, and two included patients with ocular, cutaneous, or other site melanoma. Three studies excluded those patients with poor performance status. Median age was in the 50s for four studies and 61 in one study. One article did not describe any previous treatment and one described it incompletely. Four studies reported tumor response data, by RECIST criteria.
• Treatment response. Among 32 patients in the study by Gonsalves, one patient had a CR (3%), one had a PR, 18 patients had SD (56%) and 12 patients had PD (38%). In the study of 13 patients published by Klingenstein, none had a CR, eight had a PR (62%), two had SD (15%) and three had PD (23%). Nine of 11 patients in the article by Kennedy provided response data: one had CR, six had PR, one had SD and one had PD. Of the eight patients in the Schelhorn study, four (50%) had SD and four (50%) had PD. Memon reported PD and SD in 13 (81%) patients and PD in three (19%) patients. Ponti reported disease control at six months post-RE in 52% of patients.

• Survival. Median survival in Gonsalves, Klingenstein, Schelhorn, Ponti, and Kennedy were 10.0 months, 19 months, 20 months, 18 months, and not yet reached, respectively.

• Toxicity. Gonsalves reported four patients (12.5%) with grade 3 to 4 liver toxicity and Ponti reported grade 3 to 4 biologic and clinical toxicities in 24% of patients. Klingenstein observed one patient with marked hepatomegaly. Kennedy described one grade 3 gastric ulcer. Memon reported Grade 3 toxicity in two (12%) (absolute lymphocyte toxicity) and one (7%) (aspartate aminotransferase toxicity) patients; and grade 4 bilirubin toxicity in one patient. One study[57] (n=12) did not include any toxicity data.

RADIOEMBOLIZATION FOR UNRESECTABLE METASTATIC NEUROENDOCRINE TUMORS

Systematic Reviews

Ngo (2021) conducted a meta-analysis of six retrospective cohort studies with a total of 643 patients treated with TACE (n=422) or RE (n=221) for neuroendocrine liver metastases.[61] Patients treated with TACE exhibited significantly improved OS (OR 1.92; 95% CI 1.14 to 3.22; p=0.014) compared to those treated with RE. No significant differences in hepatic progression-free survival (p=0.96) or overall tumor response (p=0.99) were observed. Although the overall proportion of patients with unresectable disease is unclear, the history of resection or ablation in the two groups was not significantly different (OR 1.20; 95% CI 0.71 to 2.02; p=0.49). Patients receiving RE were more likely to have received prior systemic chemotherapy (OR 0.48; 95% CI 0.27 to 0.83; p=0.009) and octreotide therapy (OR, 0.50; 95% CI 0.30 to 0.84; p=0.009).

Frilling (2019) reported results from a case series of 24 patients that were then included in a meta-analysis of patients treated with SIRT for neuroendocrine liver metastases.[62] Overall, 26 additional studies were included in the meta-analyses, which reported a fixed effects weighted averages for objective response rate of 51% (95% CI 47% to 54%) and disease control rate (complete response, partial response, or stable disease) of 88% (95% CI 85% to 90%).

A 2012 systematic review evaluated the safety and efficacy of chemoembolization, bland embolization, and radioembolization in patients with unresectable metastatic neuroendocrine tumors (mNET) in the liver.[63] A total of 37 studies with 1575 total patients were reviewed for response to treatment, survival outcome, and toxicity. The authors reported that each of these therapies were found to be safe and effective, and recommended additional prospective trials to compare relative efficacy and toxicity.

In 2014, a meta-analysis of 12 studies that met inclusion criteria reported complete and partial responses of 50% for radioembolization of metastatic neuroendocrine tumors (mNET) in the liver.[64] Weighted average disease control was 86%. It was noted that patients with pancreatic...
mNET was marginally associated with poorer response (p=0.03). The authors concluded that the meta-analysis confirmed the effectiveness of radioembolization of hepatic mNET.

**Randomized Controlled Trials**

No randomized controlled trials were found for radioembolization of metastatic neuroendocrine tumors in the liver.

**Nonrandomized Comparative Studies**

Egger (2020) performed a retrospective cohort analysis comparing patients with neuroendocrine liver metastases treated with RE (n=51) or TACE (n=197). Between RE and TACE, there were no differences in overall morbidity (13.7% vs. 22.6%, respectively; p=0.17), grade 3/4 complication (5.9% vs. 9.2%; p=0.58), 90-day mortality (9.8% vs 5.2%; p=0.21), median OS (35.9 months vs. 50.1 months; p=0.3), or progression-free survival (15.9 vs. 19.9 months; p=0.37). However, disease control rate was greater for TACE compared with RE (96% vs. 83%, p<0.01).

Engelman retrospectively compared locoregional therapies including transarterial, liver-directed therapies including RE, hepatic artery embolization (HAE), and hepatic artery chemoembolization (HACE) in 42 patients treated for metastatic neuroendocrine tumors. Treatment decisions were at the discretion of the referring physician and interventional radiologist, but the decision to proceed with therapy was typically based on progression of symptoms nonresponsive to octreotide therapy or rapid progression of liver tumor burden on imaging. Seventeen patients had HACE, 13 had HAE, and 12 had RE. Among the 27 patients with symptoms from their liver metastases, there were no statistically significant differences in symptom improvement at three months after first liver-directed therapy across treatment modalities (6/13 for HACE; 4/8 for HAE; 5/6 for RE; p=0.265). There were no differences between treatment modalities in radiographic response at six months postprocedure (p=0.134), TTP (p=0.968), or OS (p=0.30).

**Nonrandomized Non-Comparative Studies**

In 2015 Peker reported on 30 patients with unresectable hepatic mNET who received resin-based RE. Post-treatment response was assessed by imaging using the Response Evaluation Criteria in Solid Tumors (RECIST) guidelines. Mean follow-up was 23 months. Median OS was 39 months (range 12.6-65.4 months) with 1- and 2-year survival rates of 71% and 45%, respectively. PR was 43%, CR 3%, SD 37%, and PD 17%. The following were not significant prognostic factors: extrahepatic disease, radiographic response, age, and primary NET site.

In 2010, Cao reported the outcomes of 58 patients with unresectable neuroendocrine liver metastases from two different hospitals treated with yttrium-90 microspheres (SIR-Spheres) from 2003 to 2008. Data were examined retrospectively from a database. Response was assessed with radiographic evidence before and after radioembolization and measured by RECIST guidelines. Patients typically had a CT scan within three months of treatment and every three to six months until disease progression or death. Systemic chemotherapy was routinely given at one institution but not the other. Mean patient age at the time of radioembolization was 61 (range: 29 to 84 years), and 67% of patients were men. Primary tumor site was variable and included small bowel, pancreas, colon, thyroid, lung, and unknown. Thirty-one patients underwent surgical resection of their primary tumor, which was
classified as low-grade in 15, intermediate-grade in seven, and high-grade in seven. Forty-three percent of patients had extrahepatic metastatic disease at study entry. Prior therapies before radioembolization included liver resection in 19 patients, TAE or TACE in six, ablation or percutaneous ethanol injection in 10, previous chemotherapy in 20, concurrent chemotherapy in 34, and post-radioembolization chemotherapy in five patients. Median follow-up was 21 months (range 1-61 months). Fifty-one patients were evaluable, and six achieved a complete response, 14 a partial response, 14 had stable disease, and 17 had disease progression. Overall survival (OS) rates at one, two, and three years were 86, 58, and 47%, respectively. Median survival was 36 months (range: 1 to 61 months). Prognostic factors for survival included extent of tumor involvement of the liver, radiographic response to treatment, presence of extrahepatic disease at the time of radioembolization, histological grade of tumor, and whether patients were responders (versus nonresponders) to radioembolization. Factors that were not significant prognostic features included age, sex, ECOG status, and previous therapy.

King reported outcomes in patients treated in a single-institution prospective study.[8] Thirty-four patients with unresectable neuroendocrine liver metastases were given radioactive microspheres [SIR-Spheres] and concomitant seven-day systemic infusion of 5-FU, between 2003 and 2005. Mean patient age was 61 years (range: 32-79 years), and 65% were men. Mean follow-up was 35.2 +/- 3.2 months. The mean interval from diagnosis of hepatic metastases and treatment with SIR therapy was 36.6 +/- 6.7 months. Primary tumor sites were variable and included bronchus (n=1), thyroid (n=2), gastrointestinal (n=15), pancreas (n=8), and unknown (n=8). Subjective changes from baseline hormone symptoms were reported every three months. Twenty-four patients (71%) had, at baseline assessment, symptoms of carcinoid syndrome, including diarrhea, flushing, or rash. At three months, 18 of 33 patients (55%) reported improvement of symptoms, as did 16 of 32 (50%) at six months. Radiologic tumor response was observed in 50% of patients and included six CR (18%), and 11 PR (32%). Mean OS was 29.4 +/- 3.4 months.

RADIOEMBOLIZATION FOR INTRAHEPATIC CHOLANGIOCARCINOMA

Systematic Reviews

Mosconi (2021) published a systematic review and meta-analysis of TACE and TARE for unresectable intrahepatic cholangiocarcinoma.[69] Of the 31 total articles included, 13 were on TACE (906 patients) and 18 were on TARE (789 patients). There was moderate heterogeneity between groups for clinical and tumor characteristics. The median survival after treatment was 14.2 months for TACE and 13.5 months after TARE. The survival difference effect size (d) at one, two, and three years was 0.112, 0.028, and 0.049, respectively. Clinical adverse events occurred in 58.5% and 43.0% of TACE and TARE patients, respectively (d=0.032).

In 2015, Al-Adra reported results from a systematic review of studies reporting outcomes for RE for ICC.[70] The review included 12 publications, seven of which were published in abstract form only. Of the peer reviewed manuscripts, three were described as prospective cohort studies.[71-73] The overall weighted median survival was 15.5 months (range 7 to 22.2 months), based on 11 included studies. A weighted mean PR was seen in 28% of patients and stable disease was seen in 54% at three months posttreatment.

In 2015, Boehm conducted a meta-analysis to compare hepatic artery-based therapies including hepatic arterial infusion (HAI), TACE, DEB-TACE, and yttrium-90 RE for unresectable ICC.[74] Twenty studies met inclusion criteria, five of which evaluated yttrium-90
RE. Median OS across studies was 22.8 months for HAI, 13.9 months for RE, 12.4 months for TACE, and 12.3 months for DEB-TACE. CR or PR occurred in 56.9% of patients treated with HAI, compared with 27.4% of those treated with RE and 17.3% of those treated with TACE. While HAI showed the highest median OS, it also had the highest rate of grade III and IV toxicity.

**Randomized Controlled Trials**

No randomized controlled trials were found for radioembolization of ICC.

**Nonrandomized Studies**

Additional nonrandomized studies not included in the above systematic reviews are included here.

Riby (2020) retrospectively reviewed data from 169 consecutive patients with ICC in a single-center database. Patients were treated with upfront surgery (n=137) or downstaging with chemotherapy alone (n=13) or chemotherapy plus radioembolization (n=19). Chemotherapy regimens included gemcitabine and oxaliplatin, FOLFIRINOX, and cisplatin-based treatments. Overall survival of patients resected following downstaging treatment was not significantly different compared with patients that were resected upfront. Delivery of radioembolization as a downstaging treatment was associated with a significant benefit in multivariable analysis (HR, 0.34; 95% CI 0.14 to 0.84; p=0.019).

Edeline (2019) published results from the phase 2, MISPHEC trial (Yttrium-90 Microspheres in Cholangiocarcinoma), which included 41 patients with unresectable ICC treated in the first-line setting with cisplatin, gemcitabine, and RE in French centers with experience with glass microspheres. Fifteen (37%) patients underwent more than one RE treatment. The response rate at three months according to RECIST version 1.1 criteria was 39% (90% CI 26% to 53%) according to local review, with a disease control rate of 98%. After a median follow-up of 36 months, median PFS was 14 months (95% CI, 8 to 17 months) and median OS was 22 months (95% CI, 14 to 52 months). Of 41 patients, 29 (71%) experienced grade 3 and 4 toxic events, including neutropenia (51%), thrombocytopenia (24%), asthenia (225), anemia (20%), and abdominal pain (12%). Fourteen patients experienced hepatic failure, including five nonreversible cases in patients with cirrhosis who had received whole-liver RE. Nine patients (22%) were downstaged to surgical intervention, with eight cases achieving an R0 surgical resection. A follow-up phase 3 trial randomizing patients with unresectable ICC to chemotherapy alone or RE followed by chemotherapy in the first-line setting is currently underway.

Jia (2017) retrospectively reviewed all 24 patients who underwent Y90 RE for unresectable and failed first-line chemotherapy for ICC at a single institution. Mean follow-up was 11 months (range, 3 to 36 months). Median OS from time of diagnosis was 24 months (range, 18 to 30 months) and from the RE procedure was nine months (range, 6 to 12 months). Survival rates at 6, 12, and 30 months was 70%, 33%, and 20%, respectively.

Rayar (2015) reported successful downstaging after RE in eight patients with unresectable ICCs. Initial unresectability was due to involvement of hepatic veins or portal veins of the future liver remnant. After RE there was significant decrease in tumor volume and all patients were subsequently able to undergo successful resection. At median follow-up of 15.6 months (range 4 to 40.7 months) after medical treatment and 7.2 months (range 0.13 to 36.4 months) after...
surgery, five patients were still alive, one of which was alive at 40 months after medical treatment. Two patients had tumor recurrence.

**RADIOEMBOLIZATION FOR METASTATIC BREAST TUMORS**

**Systematic Reviews**

Aarts (2021) published a systematic review and meta-analysis of intra-arterial therapies for breast cancer metastatic to the liver.[78] Twenty-six studies (1,266 patients), 11 on TARE, 10 on transarterial chemoembolization (TACE) and four on chemo-infusion met inclusion criteria. One study was a retrospective comparative study of TARE and TACE. According to the meta-analysis, pooled response rates were 49% for TARE (95% CI 32 to 67%), 34% for TACE (95% CI 22 to 50%) and 19% for chemo-infusion (95% CI 14 to 25%) and pooled median survival was 9.2 months (range 6.1 to 35.4 months) for TARE, 17.8 months (range 4.6 to 47.0) for TACE and 7.9 months (range 7.0 to 14.2) for chemo-infusion. Missing survival rates at specific time points (one- and two-year OS) and large heterogeneity prevented comparisons of OS.

A systematic review by Smitz (2013) included six studies with a total of 198 patients with breast cancer metastases in the liver.[79] Five studies reported tumor response. Overall disease control (complete response, partial response, and stable disease) at two to four months post-treatment ranged from 78% to 96%. Median survival was reported in four studies and ranged from 10.8 to 20.9 months. Adverse effects included gastric ulceration in 10 patients (5%) and treatment-related mortality in three patients (2%). The authors concluded that these studies showed safety and effectiveness of treatment and strongly encouraged comparative studies, in particular, combining radioembolization with systemic therapy.

**Nonrandomized Studies**

Ridouani (2021) published the results of a retrospective study reviewing all breast cancer patients undergoing RE of liver metastases from 2011 to 2019 at a single center.[80] RE was performed with glass (66%) or resin (34%) microspheres based on operator preference. Imaging response assessments were available for 60/64 patients, of which 46 (77%, 95% CI 64% to 86%) achieved an objective response (OR), demonstrating a 30% or greater reduction in metabolic activity. Patients with OR had a high median dose deliver to the tumor (167 Gy) compared to patients not achieving an OR (54 Gy; p<.001). Eight patients developed grade 3 or higher treatment-related hepatotoxicity.

Davisson (2020) retrospectively reviewed 24 patients with chemotherapy-refractory hepatic metastases from breast cancer who underwent RE from 2013 to 2018.[81] Extrahepatic metastases were reported in 18 and 20 continued to receive concurrent chemotherapy and/or immunotherapy. Median OS was 35.4 months from first RE. Radioembolization within six months of hepatic metastasis diagnosis and estrogen receptor-positive status were identified as positive predictors of overall survival.

**Table 1. Retrospective Case Series of Radioembolization for Liver Metastases in Breast Cancer**

<table>
<thead>
<tr>
<th>Study (Year)</th>
<th>Populations</th>
<th>Outcomes</th>
</tr>
</thead>
</table>
| Pieper (2016)[82] | 44 women with unresectable liver-dominant breast metastases who had failed 2+ lines of chemotherapy who | ORR: 29%  
Disease control rate: 71%  
Median TTP: 101 d  
Median survival: 184 d |
<table>
<thead>
<tr>
<th>Study</th>
<th>Patients</th>
<th>Disease</th>
<th>Treatment</th>
<th>Toxicity</th>
<th>Survival</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gordon (2014)[83]</td>
<td>75 women with stable extrahepatic disease who had hepatic tumor progression after systemic chemotherapy treated with yttrium-90 RE at a single center</td>
<td>30-day mortality: 4%</td>
<td></td>
<td>Grade 2 toxicity: 1 (cholecystitis) Grade 3 toxicity: 1 (duodenal ulceration)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saxen (2014)[84]</td>
<td>40 women with unresectable, chemoresistant breast cancer–related liver metastases treated from 2006-2012 at a single institution who had received at least one line of systemic chemotherapy</td>
<td>Grade 1 or 2 clinical toxicity: 40% Of 38 women with ≥1 mo follow-up: CR: 5% PR: 26% SD: 39% PD: 29%</td>
<td>Median survival: 13.6 mo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cianni (2013)[85]</td>
<td>52 women with chemotherapy-refractory breast cancer and inoperable liver metastases; chemotherapy administered previously to all patients, surgery in 17.3%, TACE in 3.8%, and RFA in 3.8%</td>
<td>CR: 0% PR: 56% SD: 35% PD: 10%</td>
<td>Median OS: 11.5 mo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haug (2012)[72]</td>
<td>58 women with chemotherapy-refractory breast cancer and unresectable hepatic metastases</td>
<td>Mean follow-up: 27.5 wk CR: 0% PR: 25.6% SD: 62.8% PD: 11.6%</td>
<td>Median survival: 47 wk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jakobs (2008)[32]</td>
<td>30 (29 women, 1 man) patients who underwent RE with resin microspheres in a single-session, whole-liver treatment for breast cancer metastases and had failed prior polychemotherapy regimens</td>
<td>For 23 patients with follow-up data, after median follow-up of 4 mo: PR: 61% SD: 35% PD: 4% One death due to treatment-related hepatic toxicity after median follow-up of 14.2 mo</td>
<td>Median OS: 11.7 mo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangash (2007)[33]</td>
<td>27 women with progressive liver metastases from breast cancer while on polychemotherapy</td>
<td>After 90-d follow-up CR: 39% PR: 39% SD: 52% PD: 9%</td>
<td>Median survival ECOG Performance Status 0: 6.8 mo ECOG Performance Status 1-3: 2.6 mo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coldwell (2007)[43]</td>
<td>44 patients with hepatic metastases at three hospitals who failed 1st-, 2nd-, or 3rd-line treatment for primary breast tumor and were not candidates for RFA, TACE, resection, IMRT, or SRT</td>
<td>After 12-wk follow-up PR: 47%</td>
<td>No radiation-related liver failures were observed Median survival: &gt;14 mo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI: confidence interval; CR: complete response; ECOG: Eastern Cooperative Oncology Group; IMRT: intensity-modulated radiotherapy; ORR: response rate; OS: overall survival; PD: progressive disease; PR: partial response; RE: radioembolization; RFA: radiofrequency ablation; SD: stable disease; SRT: stereotactic radiotherapy; TACE: transarterial chemoembolization; TTP: time to progression.
OTHER METASTATIC TUMORS IN THE LIVER

Data on the use of radioembolization in other tumors metastatic to the liver are limited and included numerous methodologic limitations such as patient heterogeneity, lack of a control group, and patient numbers too small to draw meaningful conclusions. For example, a retrospective data analysis was reported in 2014 by Michl on RE for liver metastases from pancreatic cancer. Nineteen patients were included, 16 of whom had received previous palliative chemotherapy.\[86\] Median local PFS in the liver was 3.4 months (range 0.9 to 45.0). Median OS was nine months (range 0.9 to 53.0) and one-year survival was 24%. Adverse effects were grade <3 (e.g., nausea, vomiting, fatigue, fever, abdominal pain) in the short term and long-term effects included liver abscess, gastroduodenal ulceration, cholestasis and cholangitis, ascites, and spleen infarction. The lack of a control group precludes conclusions about any survival benefits and complication rates of RE.

RADIOEMBOLIZATION AS A BRIDGE TO HEPATIC RESECTION

In 2013, Vouche reported on 83 patients treated with radioembolization as a technique to control or limit tumor progression in unresectable, unilobar hepatic disease and to hypertrophy a small future liver remnant.\[87\] Patients included in the study had right unilobar disease with HCC (n=67), cholangiocarcinoma (n=8), or metastatic CRC (n=8). One month after radioembolization, significant right lobe atrophy (p=0.003), left lobe hypertrophy (p<0.001), and future liver remnant hypertrophy (p<0.001) were observed and remained during follow-up. Successful right lobectomy was later performed in five patients, and six patients received liver transplants. However, further studies are needed to assess radioembolization as a bridge to hepatic resection.

PRACTICE GUIDELINE SUMMARY

NATIONAL COMPREHENSIVE CANCER NETWORK GUIDELINES

All the following statements are category 2A recommendations unless specified.

Primary Hepatocellular Carcinoma

National Comprehensive Cancer Network (NCCN) guidelines for hepatocellular carcinoma principles of local regional therapy state that all patients with HCC should be evaluated for potential curative therapies.\[9\] They recommend considering locoregional therapies for patients who are not candidates for surgical curative treatments, or as part of a strategy to bridge patients for other curative therapies. They also state that “all tumors irrespective of location may be amenable to arterially directed therapies [including bland TAE, TACE and TACE with drug-eluting beads, and RE with yttrium-90 microspheres] provided that the arterial blood supply to the tumor may be isolated without excessive non-target treatment.”

NCCN discussion indicates that there is limited evidence available on the utility of radioembolization as a bridge to liver transplant for patients on a liver transplant waiting list. However, most NCCN member centers use radioembolization as a bridge to transplant.

Primary Intrahepatic Cholangiocarcinoma

Recommendations for unresectable intrahepatic cholangiocarcinoma (ICC) include chemotherapy, clinical trial, radiotherapy, arterially directed therapies, and supportive care.\[9\] The guideline follows that of the HCC locoregional therapy recommendations, above.
Metastatic Colorectal Cancer

Use of intra-arterial embolization including RE is a recommendation for highly selected patients with chemotherapy-resistant/refractory disease without obvious systemic disease, with predominant hepatic metastases.\[6, 88\] Additionally, for hepatic metastases that are not optimally resectable, portal vein embolization and yttrium-90 radioembolization are among the options that can be considered.

Metastatic Neuroendocrine Tumors

For unresectable liver metastases (carcinoid or neuroendocrine tumors of the pancreas, e.g., islet cell), recommendations include hepatic regional therapy which includes radioembolization.\[1\]

Metastatic Breast Cancer

Current recommendations do not address the use of radioembolization in the treatment of metastatic breast cancer.\[89\]

Metastatic Melanoma

Current recommendations do not address the use of radioembolization in the treatment of metastatic cutaneous melanoma.\[90\] Regarding treating of liver metastases from uveal melanoma, the guidelines state that SIRT has been reported in retrospective studies and one prospective study and that more research is needed.\[91\]

AMERICAN COLLEGE OF RADIOLOGY APPROPRIATENESS CRITERIA®

The American College of Radiology (ACR) published Appropriateness Criteria for radiologic management of hepatic malignancy. \[92\]

Primary Hepatocellular Carcinoma

ACR Appropriateness Criteria consider radioembolization with beta-emitting Y90 beads to be an emerging treatment option for HCC, with outcomes similar to those with transarterial chemoembolization (TACE) and transarterial embolization (TAE), but with the possibility of less patient discomfort and toxicity. The guideline also reports that radioembolization has "shown the ability to effectively downstage patients for potential transplant or resection. Therefore, ACR recommendations are that radioembolization may be appropriate for solitary HCC tumor <3cm, and usually appropriate, particularly in the presence of portal vein thrombosis or extensive bilobar disease, for solitary HCC tumor of 5 cm and for multiple tumors, at least one of which is >5cm.

Metastatic Colorectal Cancer

The ACR reports that published evidence suggests that TACE and radioembolization provide similar survival benefit and may be appropriate for patients with metastatic liver-dominant colorectal tumors \(\geq 5\) cm, or for solitary colorectal liver metastasis.

Metastatic Neuroendocrine Tumors

The ACR reports increasing research into the use of radioembolization in this patient population, with early small studies suggesting therapeutic equivalency with more traditional
arterial embolization techniques. Radioembolization is recommended as usually appropriate for symptomatic neuroendocrine metastases in the liver when medication fails to control symptoms.

AMERICAN COLLEGE OF RADIOLOGY/AMERICAN SOCIETY FOR RADIATION ONCOLOGY/SOCIETY OF INTERVENTIONAL RADIOLOGY ET AL

A joint practice parameter from the American College of Radiology (ACR), American Brachytherapy Society (ABS), American College of Nuclear Medicine (ACNM), American Society for Radiation Oncology (ASTRO), Society of Interventional Radiology (SIR), and Society of Nuclear Medicine and Molecular Imaging (SNMMI) on selective internal radiation therapy list indications for radioembolization which include, but are not limited to:[93]

- Unresectable and/or inoperable primary or secondary liver malignancies that are liver dominant but not necessarily exclusive to the liver; and
- Performance status that will allow them to benefit from the therapy (e.g., ECOG performance status of 0 or 1 or KPS of 70 or more); and
- Life expectancy of at least three months

RADIOEMBOLIZATION BRACHYTHERAPY ONCOLOGY CONSORTIUM

Members met as an independent group of experts in interventional radiology, radiation oncology, nuclear medicine, medical oncology, and surgical oncology. Using level 2A evidence (panel consensus with low-level evidence), 14 recommendations were made. They concluded that there was sufficient evidence to support the safety and efficacy of yttrium-90 microsphere therapy and that its use requires multidisciplinary management, adequate patient selection, and meticulous angiographic technique. They also stated that the initiation of clinical trials was necessary to further define the role of yttrium-90 microsphere therapy in relation to other currently available therapies.[94]

SUMMARY

TRANSARTERIAL EMBOLIZATION WITH NON-RADIOACTIVE AGENTS

There is enough research to show that transarterial embolization (TAE) with non-radioactive agents improves health outcomes for people with cancer and various conditions. Therefore, transarterial embolization (TAE) with non-radioactive agents may be considered medically necessary for any indication.

TRANSARTERIAL CHEMOEMBOLIZATION

There is enough research to show that transarterial chemoembolization (TACE) improves health outcomes for people with cancer and various conditions. Therefore, transarterial chemoembolization (TACE) may be considered medically necessary for any indication.

RADIOEMBOLIZATION

Primary Hepatocellular Carcinoma (HCC)

Studies have demonstrated that radioembolization is comparable to transarterial chemoembolization (TACE), which is considered to be the therapy of choice for patients with
unresectable primary hepatocellular carcinoma (HCC) in terms of tumor response and overall survival. However, disadvantages of TACE include the necessity of multiple treatment sessions and hospitalization, its contraindication in patients with portal vein thrombosis, and its poorer tolerance by patients. Therefore, radioembolization may be considered medically necessary for the treatment of unresectable primary HCC or as a bridge to transplantation in primary HCC.

**Metastatic Colorectal Cancer in the Liver**

A major cause of morbidity and mortality in patients with colorectal disease metastatic to the liver is liver failure, as this disease tends to progress to diffuse, liver-dominant involvement. Therefore, the use of radioembolization to decrease tumor bulk and/or halt the time to tumor progression and liver failure may lead to prolonged progression free and overall survival in patients with no other treatment options (i.e., those with chemotherapy refractory liver-dominant disease). Other uses include palliation of symptoms from tumor bulk. Radioembolization for the treatment of unresectable hepatic metastases from colorectal cancer may be considered medically necessary in carefully selected patients, when criteria are met.

There is insufficient evidence on the use of radioembolization for the treatment of unresectable hepatic metastases from colorectal cancer when the patient does not meet criteria. Therefore, radioembolization for the treatment of unresectable hepatic metastases from colorectal cancer is considered investigational when criteria are not met.

**Metastatic Neuroendocrine Tumors in the Liver**

Studies of radioembolization for treatment of metastatic neuroendocrine tumors in the liver have included heterogeneous patient populations, making interpretation of survival data difficult. However, relief of symptoms from carcinoid syndrome has been reported in a proportion of patients. Surgical debulking of liver metastases has shown palliation of hormonal symptoms; similarly, debulking by radioembolization may lead to symptom relief in some patients. Therefore, radioembolization for the treatment of unresectable hepatic metastases from neuroendocrine tumors may be medically necessary in carefully selected patients when criteria are met.

There is insufficient evidence on the use of radioembolization for the treatment of hepatic metastases from neuroendocrine tumors when the patient does not meet criteria. Therefore, radioembolization for the treatment of hepatic metastases from neuroendocrine tumors is considered investigational when criteria are not met.

**Metastatic Melanoma in the Liver**

In patients with uveal melanoma, the liver is the most common site of metastatic disease. Studies of radioembolization for treatment of metastatic melanoma (uveal or cutaneous) in the liver consists of one comparative study and several relatively small observational studies. In general, these studies predict good tumor response to radioembolization and report significant increases in overall survival compared to those treated with best supportive care. Therefore, radioembolization may be considered medically necessary for the treatment of diffuse, symptomatic hepatic metastases from melanoma when criteria are met.

*These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.*
There is insufficient evidence on the use of radioembolization for the treatment of hepatic metastases from melanoma when the patient does not meet criteria. Therefore, radioembolization for the treatment of hepatic metastases from melanoma is considered investigational when criteria are not met.

Primary Intrahepatic Cholangiocarcinoma (ICC)

The current evidence on the use of radioembolization (RE) in patients with primary intrahepatic cholangiocarcinoma (ICC) is limited to data from small studies that do not compare the health outcomes of RE with other treatments. These study designs make interpretation of the data on tumor response and survival difficult to interpret. However, ICC is a rare tumor, so large comparative studies may never become available. The available studies have consistently reported beneficial effects in patients who are not candidates for surgical tumor resection. Because there are currently limited treatment options for these patients, radioembolization may be medically necessary for the treatment of unresectable primary ICC. Since surgical resection is currently the preferred treatment for these tumors, radioembolization is considered investigational for resectable primary ICC.

Miscellaneous Metastatic Tumors in the Liver

The current evidence on the use of radioembolization in intrahepatic cholangiocarcinoma and metastatic tumors in the liver other than those from colorectal carcinoma, melanoma or neuroendocrine tumors is too limited to draw meaningful conclusions due to methodologic limitations such as small numbers of heterogeneous patients. Therefore, radioembolization for these other tumors, including metastatic tumors from breast and pancreatic cancer, is considered investigational.

REFERENCES


20. S Belinson, Y Yang, R Chopra, V Shankaran, D Samson, N Aronson. Local Therapies for Unresectable Primary Hepatocellular Carcinoma [Internet]. AHRQ Comparative Effectiveness Reviews. 2013. PMID: 23844445


These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.
Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


42. GA van Hazel, V Heinemann, NK Sharma, et al. SIRFLOX: Randomized Phase III Trial Comparing First-Line mFOLFOX6 (Plus or Minus Bevacizumab) Versus mFOLFOX6 (Plus or Minus Bevacizumab) Plus Selective Internal Radiation Therapy in Patients With Metastatic Colorectal Cancer. *J Clin Oncol.* 2016;34:1723-31. PMID: 26903575


These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


62. A Frilling, AK Clift, A Braat, et al. Radioembolisation with 90Y microspheres for neuroendocrine liver metastases: an institutional case series, systematic review and


"Radioembolization for Primary and Metastatic Tumors of the Liver." Policy No. 8.01.43

**CODES**

**NOTE:** CPT code 37243 can be used for both radioactive and non-radioactive embolization procedures performed for numerous conditions/locations. Embolization codes requiring prior authorization are listed on the “Pre-authorization List” web page. There may be codes related to embolization, such as CPT 37242 which may be used for prostate artery embolization, that do not require prior approval. Embolization codes not listed on the pre-authorization website do not require prior approval.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>37242</td>
<td>Vascular embolization or occlusion, inclusive of all radiological supervision and interpretation, intraprocedural roadmapping, and imaging guidance necessary to complete the intervention; arterial, other than hemorrhage or tumor (eg, congenital or acquired arterial malformations, arteriovenous malformations, arteriovenous fistulas, aneurysms, pseudoaneurysms)</td>
</tr>
<tr>
<td></td>
<td>37243</td>
<td>Vascular embolization or occlusion, inclusive of all radiological supervision and interpretation, intraprocedural roadmapping, and imaging guidance necessary to complete the intervention; for tumors, organ ischemia, or infarction</td>
</tr>
<tr>
<td></td>
<td>75894</td>
<td>Transcatheter therapy, embolization, any method, radiological supervision and interpretation</td>
</tr>
<tr>
<td></td>
<td>77778</td>
<td>Unlisted procedure, medical radiation physics, dosimetry and treatment devices, and special services</td>
</tr>
<tr>
<td></td>
<td>79445</td>
<td>Radiopharmaceutical therapy, by intra-arterial particulate administration</td>
</tr>
<tr>
<td>HCPCS</td>
<td>C2616</td>
<td>Brachytherapy source, non-stranded, yttrium-90, per source</td>
</tr>
<tr>
<td></td>
<td>S2095</td>
<td>Transcatheter occlusion or embolization for tumor destruction, percutaneous, any method, using yttrium-90 microspheres</td>
</tr>
</tbody>
</table>

**Date of Origin:** December 2010
Orthopedic Applications of Stem Cell Therapy, Including Bone Substitutes Used with Autologous Bone Marrow

Effective: July 1, 2022

Next Review: October 2022
Last Review: June 2022

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Mesenchymal stem cells (MSCs) are multipotent cells (also called “stromal multipotent cells”) that possess the ability to differentiate into various tissues including organs, trabecular bone, tendon, articular cartilage, ligaments, muscle, and fat. Potential uses of MSCs for orthopedic applications include treatment of damaged bone, cartilage, ligaments, tendons, and intervertebral discs.

MEDICAL POLICY CRITERIA

Note: Use of platelet rich plasma is addressed in Medicine Policy No. 77 (see Cross References section). This policy does not apply to the use of unmanipulated bone marrow aspirate for spinal indications which may be considered medically necessary.

I. Mesenchymal stem cell therapy, including but not limited to manipulated or unmanipulated bone marrow, fat, and amnion cells, is considered investigational for all orthopedic applications, including but not limited to use in repair or regeneration of musculoskeletal tissue.
II. Allograft bone products containing viable stem cells are considered *investigational* for all orthopedic applications, including but not limited to demineralized bone matrix (DBM) with stem cells.

III. Synthetic bone graft substitutes that must be combined with autologous bone marrow are considered *investigational* for all orthopedic applications.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

### CROSS REFERENCES

1. [Autologous Blood-Derived Growth Factors as a Treatment for Wound Healing and Other Conditions](#), Medicine, Policy No. 77
2. [Progenitor Cell Therapy for the Treatment of Damaged Myocardium Due to Ischemia](#), Medicine, Policy No. 100
3. [Stem-cell Therapy for Peripheral Arterial Disease](#), Medicine, Policy No. 141
4. [Autologous Chondrocyte Implantation for Focal Articular Cartilage Lesions](#), Surgery, Policy No. 87

### BACKGROUND

MSCs are associated with the blood vessels within bone marrow, synovium, fat, and muscle where they can be mobilized for endogenous repair, as occurs with healing of bone fractures. Stimulation of endogenous MSCs is the basis of procedures such as bone marrow stimulation (e.g., microfracture) and harvesting/grafting of autologous bone for fusion. Bone-marrow aspirate is considered to be the most accessible source and, thus, the most common place to isolate MSCs for treatment of musculoskeletal disease. However, harvesting MSCs from bone marrow requires an additional procedure that may result in donor-site morbidity. In addition, the number of MSCs in bone marrow is low, and the number and differentiation capacity of bone marrow-derived MSCs decreases with age, limiting their efficiency when isolated from older patients.

Tissues such as muscle, cartilage, tendon, ligaments, and vertebral discs show limited capacity for endogenous repair. Tissue engineering techniques are being developed to improve the efficiency of repair or regeneration of damaged musculoskeletal tissues. Tissue engineering focuses on the integration of biomaterials with MSCs and/or bioactive molecules such as growth factors. In vivo, the fate of stem cells is regulated by signals in the local 3-dimensional microenvironment from the extracellular matrix and neighboring cells. It is believed that the success of tissue engineering with MSCs will also require an appropriate 3-dimensional scaffold or matrix, culture conditions for tissue-specific induction, and implantation techniques that provide appropriate biomechanical forces and mechanical stimulation. Given that each tissue type requires different culture conditions, induction factors (e.g., signaling proteins, cytokines, growth factors), and implantation techniques, each preparation must be individually examined. The ability to induce cell division and differentiation, without adverse effects such as the formation of neoplasms, remains a significant concern.

The U.S. Food and Drug Administration (FDA) stated:

*“Cell-based therapies are one of the most rapidly advancing approaches intended to repair, replace, restore, or regenerate cells, tissues and organs. They can be applied to damage caused by disease, injury, or aging. Many cell-based therapies use immature cells (stem cells) that are expanded outside of the body. The expanded cells are sometimes used in their...***
immature state, but they are often manufactured into more mature cells before they are given to a patient. The resulting cells are intended to repair cell or tissue damage (efficacy) without unintended serious consequences such as tumors, severe immune reactions, or unwanted tissue development (safety). Manufacturing of large numbers of cells outside the natural environment of the human body may lead to ineffective or dangerous cells, so it is important to understand and carefully control the production process and to define measures that reliably predict safety and efficacy of the cell-based products.”[1]

REGULATORY STATUS

Concentrated autologous MSCs do not require approval by the U.S. Food and Drug Administration (FDA).

Demineralized bone matrix (DBM), which is processed allograft bone, is considered minimally processed tissue and does not require FDA approval. At least four commercially available DBM products are reported to contain viable stem cells:

- AlloStem® (AlloSource) is partially demineralized allograft bone seeded with adipose-derived MSCs
- Map3™ (RTI surgical) contains cortical cancellous bone chips, DBM, and multipotent adult progenitor cells
- OsteoCell Plus® (NuVasive): an allograft cellular bone matrix containing native MSCs.
- Trinity Evolution Matrix™ (Orthofix): an allograft that is processed and cryopreserved to maintain viable adult MSCs and osteoprogenitor cells.

Whether these products can be considered minimally manipulated tissue is debated. A product would not meet the criteria for FDA regulation part 1271.10 if it is dependent upon the metabolic activity of living cells for its primary function. Otherwise, a product would be considered a biologic product and would need to demonstrate safety and efficacy for the product’s intended use with an investigational new drug and Biologics License Application (BLA).

Other products contain DBM and are designed to be mixed with bone marrow aspirate. Some of the products that are currently available are:

- Fusion Flex™ (Wright Medical): a dehydrated moldable DBM scaffold that will absorb autologous bone marrow aspirate.
- Ignite® (Wright Medical): an injectable graft with DBM that can be combined with autologous bone marrow aspirate.

Other commercially available products are intended to be mixed with bone marrow aspirate and have received 510(k) clearance, such as:

- CopiOs sponge or paste (Zimmer): synthetic bone graft material consisting of mineralized, lyophilized collagen.
- Collage™ Putty (Orthofix): Composed of type-1 bovine collagen and beta Tri-calcium phosphate.
- Vitoss® (Stryker, developed by Orthovita): composed of beta tricalcium phosphate.
- nanOss® Bioactive (RTI Surgical, developed by Pioneer Surgical): nanostructured hydroxyapatite and an open structured engineered collagen carrier.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
No products using engineered MSCs have been approved by the FDA for orthopedic applications.

In 2008, the FDA determined that the mesenchymal stem cells sold by Regenerative Sciences for use in the Regenexx™ procedure would be considered drugs or biological products and thus require submission of a New Drug Application (NDA) or Biologics Licensing Application (BLA) to the FDA. In 2014, a federal appellate court upheld FDA’s power to regulate adult stem cells as drugs and biologics and ruled that the Regenexx cell product fell within FDA’s authority to regulate human cells, tissues, and cellular and tissue-based products (HCT/Ps) (Section 351).[2] To date, no NDA or BLA has been approved by the FDA for this product. As of 2015, the expanded stem cell procedure is only offered in the Cayman Islands. Regenexx™ network facilities in the U.S. provide same-day stem cell and blood platelet procedures, which do not require FDA approval.

**EVIDENCE SUMMARY**

At this time, the literature consists mainly of articles describing the potential of stem cell therapy for orthopedic applications in humans, along with basic science experiments on sources of mesenchymal stem cells (MSCs), regulation of cell growth and differentiation, and development of scaffolds.[3] Although the evidence base has been steadily increasing, authors indicate that the technology is in an early stage of development. In order to assess the safety and efficacy of orthopedic applications of MSCs and allograft bone products, such as demineralized bone matrix, high-quality randomized trials (RCTs) are required that compare health outcomes with versus without the use of these products.

**CARTILAGE DEFECTS**

**Systematic Reviews**

A systematic review by Borakati (2018) included 13 studies comparing patients with osteoarthritis who were treated either with MSCs or with a control treatment that was identical other than the inclusion of MSCs (i.e., studies using chondrogenic cellular therapy as a control were not included).[4] Pain assessment results were noted for each of the controlled studies, resulting in a pooled standardized mean difference (SMD) of -1.27 (95% confidence interval [CI] -1.95 to -0.58) in favor of the group treated with MSCs. Reviewers reported a Z-statistic effect size of 3.62, again in favor of the groups treated with MSCs (p<0.001); although they noted the high heterogeneity across controlled studies (I²=92%). Additionally, 34 uncontrolled studies (n=737 patients) were summarized and evaluated qualitatively: reviewers noted consistent cartilage regrowth and reduction of pain following treatment with MSCs in these studies; however, as pain medication was often given concurrently, interpretation of the latter outcome is limited.

Emadedin (2018) reported a triple-blind placebo-controlled phase 1/2 trial of expanded MSCs in 47 patients with OA of the knee.[5] Compared to the placebo group, the MSC group showed statistically significant improvements in WOMAC pain and function subscales but not VAS. The WOMAC stiffness subscale improved to a similar extent in the two groups. Minimum Clinically Important Improvement and Patient Acceptable Symptom State were not significantly different between the two groups. Study limitations included the short duration of follow-up, statistical analysis, and lack of information regarding use of analgesic medications.

Iijima (2018) published a systematic review of MSC treatment for knee osteoarthritis, which
included 35 studies. Of these, only seven were RCTs. Meta-analysis results indicated that there was improvement in knee pain (SMD -1.45, 95% CI -1.94 to -0.96), cartilage quality (SMD -1.99, 95% CI -3.51 to -0.47), and self-reported function (SMD 1.50, 95% CI 1.09 to 1.92), however the authors stated that the evidence quality was “very low” to “low,” and emphasized the need for high-quality RCTs.

Another 2018 systematic review on stem cell therapy for articular cartilage repair noted similar concerns regarding the quality of the evidence. The review included 46 studies that evaluated MSCs from a variety of sources, most of which were case reports and case series. The authors noted that among these, “18 studies erroneously referred to adipose tissue-derived stromal vascular fractions as "adipose-derived MSCs," 2 studies referred to peripheral blood-derived progenitor cells as "peripheral blood-derived MSCs," and 1 study referred to bone marrow aspirate concentrate as "bone marrow-derived MSCs."

Cui (2016) published a systematic review on 18 studies looking at the effect of MSC in treating patients with osteoarthritis. MSC treatment in patients with KOA showed continual efficacy for 24 months compared with their pretreatment condition. Effectiveness of MSCs was improved at 12 and 24 months post-treatment, compared with at three and six months. There was no dose response association in the MSC numbers. This review only included four randomized trials while the remaining 14 studies were non-randomized and had methodological limitations.

Xu (2015) published a meta-analysis on the effect of MSCs for articular cartilage degeneration treatment, including 11 controlled trials (n=558). No critical appraisal of the quality of the included studies was reported. MSC treatment significantly improved the American Orthopedic Foot and Ankle Society Scale (SMD 0.91, 95% confidence interval [CI], 0.52 to 1.29) and the Osteo-Arthritis Outcome Score (SMD 2.81, 95% CI 2.02 to 3.60). Comprehensive evaluation indexes, such as the American Knee Society Knee Score System (SMD -0.12, 95% CI -1.02 to 0.78), the Hospital for Special Surgery Knee Rating Scale (SMD 0.24, 95% CI -0.56 to 1.05) and the International Knee Documentation Committee (SMD -0.21, 95% CI -0.77 to 0.34), were no different between MSC use and other treatments. The reviewers concluded that there was no obvious advantage regarding the application of stem cells to treat cartilage injury, compared with other treatments.

Filardo (2013) conducted a systematic review of mesenchymal stem cells for the treatment of cartilage lesions. They identified 72 preclinical papers and 18 clinical reports. Of the 18 clinical reports, none were randomized, five were comparative, six were case series, and seven were case reports. In two clinical studies the source of MSCs was adipose tissue, in five it was bone marrow concentrate, and in 11 studies the source of MSCs was bone marrow-derived. The authors reached the following conclusion:

“Despite the growing interest in this biological approach for cartilage regeneration, knowledge on this topic is still preliminary, as shown by the prevalence of preclinical studies and the presence of low-quality clinical studies. Many aspects have to be optimized, and randomized controlled trials are needed to support the potential of this biological treatment for cartilage repair and to evaluate advantages and disadvantages with respect to the available treatments.”

The source of MSCs may have an impact on outcomes, but this is not well understood, and the available literature uses multiple different sources of MSC. Because of the uncertainty over whether these products are equivalent, the summary of the key evidence to date is grouped by source of MSC.
Randomized Controlled Trials

Cartilage Defects: MSCs Expanded from Bone Marrow

Wong (2013) reported on the use of cultured MSCs in 56 patients with osteoarthritis who underwent medial opening-wedge high tibial osteotomy and microfracture of a cartilage lesion. Bone marrow was harvested at the time of microfracture and the MSCs were isolated and cultured. After three weeks, the cells were assessed for viability and delivered to the clinic, where patients received an intra-articular injection of MSCs suspended in hyaluronic acid (HA) or, for controls, intra-articular injection of HA alone. The primary outcome was the International Knee Documentation Committee (IKDC) score at six months, one year, and two years. Secondary outcomes were the Tegner and Lysholm scores through two years and the Magnetic Resonance Observation of Cartilage Repair Tissue (MOCART) scoring system by MRI at one year. All patients completed the two-year follow-up. After adjusting for age, baseline scores, and time of evaluation, the group treated with MSCs showed significantly better scores on the IKDC (mean difference 7.65 on 0 to 100 scale, p=0.001), Lysholm (mean difference, 7.61 on 0 to 100 scale, p=0.02), and Tegner (mean difference 0.64 on a 0 to 10 scale, p=0.02). Blinded analysis of MRI results found higher MOCART scores in the MSC group. The group treated with MSCs had a higher proportion of patients who had complete cartilage coverage of their lesions (32% vs 0%), greater than 50% cartilage cover (36% vs 14%) and complete integration of the regenerated cartilage (61% vs 14%).

A controlled, double-blind clinical trial was conducted with a group of 47 patients with radiographic and symptomatic knee osteoarthritis. Three groups were randomized for intra-articular injections: autologous bone marrow-derived culture-expanded MSCs (n=16); autologous bone marrow-derived culture-expanded MSCs with platelet-rich plasma (PRP) (n=14); and corticosteroid (n=17). The results of the study show Knee Injury and Osteoarthritis Outcome Score (KOOS) is significantly improved at one month (p=0.003) with MSCs and by one year both MSCs and MSCS + PRP show the highest percentage of improvement.

Cartilage Defects: MSCs Concentrated from Bone Marrow

A small RCT published by Vega (2015) that assessed the efficacy of bone marrow derived MSCs as a treatment for knee osteoarthritis, randomizing 30 patients with chronic knee pain unresponsive to conservative treatments and showing radiological evidence of osteoarthritis. Fifteen patients were treated with allogeneic bone marrow MSCs by intra-articular injection, while 15 controls received intra-articular hyaluronic acid (HA). Clinical outcomes were followed for one year and included evaluations of pain, disability, and quality of life. Articular cartilage quality was assessed by quantitative magnetic resonance imaging T2 mapping. The MSC-treated patients displayed significant improvement in algofunctional indices versus the active controls. Quantification of cartilage quality by T2 relaxation measurements showed a significant decrease in poor cartilage areas, with cartilage quality improvements in MSC-treated patients.

Cartilage Defects: Adipose-Derived MSCs

The literature on adipose-derived MSCs for articular cartilage repair is very limited, coming from two research groups in Korea. One of the groups appears to have been providing this treatment as an option for patients for a number of years and recently published a RCT that evaluated cartilage healing after high tibial osteotomy (HTO) in 52 patients with osteoarthritis of the medial compartment. Patients were randomly assigned to HTO with application of...
platelet-rich plasma (PRP) or HTO with application of PRP plus MSCs. MSCs from adipose
tissue were obtained through liposuction from the buttocks. The tissue was centrifuged and the
stromal vascular fraction mixed with PRP for injection. A total of 44 patients completed second
look arthroscopy and one- and two-year clinical follow-up. There were statistically significant
differences for PRP only versus PRP+MSC on the Knee Injury and Osteoarthritis Outcome
Score (KOOS) subscales for pain (74±5.7 vs. 81.2±6.9, p<0.001) and symptoms (75.4±8.5 vs.
82.8±7.2, p=0.006). There were also statistically significant differences on the final pain score
for the PRP only versus PRP+MSC groups (16.2±4.6 vs. 10.2±5.7, p<0.001), but the Lysholm
score, which is more scientifically proven, was not significantly different between the PRP only
and PRP+MSC groups (80.6±13.5 vs. 84.7±16.2, all respectively, p=0.36). Articular cartilage
healing was rated as improved with MSCs following video review of second-look arthroscopy;
blinding of this measure is unclear. There are a number of limitations of this study, including
the small sample size, short duration of follow-up, and significant improvements on only some
of the outcomes. All of the significant differences in outcomes were modest in magnitude, and
as a result, there is uncertainty regarding the clinical significance of the findings.

This group also published a trial comparing treatment with adipose-derived MSCs, fibrin glue,
and microfracture to microfracture alone. A total of 80 patients with a single International
Cartilage Repair Society grade III/IV symptomatic cartilage defect on the femoral condyle were
randomized to receive one of the treatments. The mean follow-up time was 27.4 months. At
follow-up, the MSC + fibrin glue + microfracture group had significantly greater improvements
in the Knee Injury and Osteoarthritis Outcome Score pain and symptom subscores than the
microfracture alone group (p=0.034 and 0.005, respectively). There were no significant
differences between groups for the activities of daily living, sports and recreation, or quality of
live subscores. Second-look arthroscopies were performed in 57 of the 80 patients, with no
significant differences between groups. The lack of blinding in this study limits the conclusions
that can be drawn from its results.

A multisite prospective double-blinded randomized placebo-controlled clinical trial was
conducted in adult patients with symptomatic knee osteoarthritis. The trial included 39
eligible patients injected with high-dose, low-dose, or placebo stromal vascular fraction
medium obtained from liposuction for intra-articular administration of progenitor cells and
mesenchymal stem cells derived from adipose tissue. After six months, change in WOMAC
score was 83.9%, 51.5%, and 25.0%, respectively, and at one year was 89.5%, 68.2%, and
0%, respectively. Significant changes when compared with placebo revealed a dose
dependent improvement in osteoarthritis symptoms and pain at six months (high dose, p=0.04;
low does, p=0.02) and at one year (high dose, p=0.006; low dose, p=0.009).

Cartilage Defects: MSCs from Peripheral Blood

A 2013 report described a small randomized controlled trial with autologous peripheral blood
MSCs for focal articular cartilage lesions. Fifty patients with grade 3 and 4 lesions of the
knee joint underwent arthroscopic subchondral drilling followed by five weekly injections of HA.
Half of the patients were randomly allocated to receive injections of peripheral blood stem cells
or no further treatment. There were baseline differences in age between the groups, with a
mean age of 38 for the treatment group compared to 42 for the control group. The peripheral
blood stem cells were harvested after stimulation with recombinant human granulocyte colony-
stimulating factor, divided in vials, and cryopreserved. At six months after surgery, HA and
MSC were re-administered over three weekly injections. At 18 months after surgery, second
look arthroscopy on 16 patients in each group showed significantly (p=.022) higher histological
scores (by about 10%) for the MSC group (1,066 vs. 957 by independent observers) while blinded evaluation of MRI showed a statistically significant (p=0.013) higher morphologic score (9.9 vs. 8.5). There was no difference in International Knee Documentation Committee (IKDC) scores between the two groups at 24 months after surgery. It is uncertain how differences in patient age at baseline may have affected the response to subchondral drilling.

**Cartilage Defects: MSCs from Synovial Tissue**

Akgun (2015) reported a small (n=14) investigator-blinded RCT that compared matrix-induced autologous MSCs from synovial tissue versus matrix-induced autologous chondrocyte implantation (MACI). Both chondrocytes from cartilage and MSCs from synovia were harvested in an arthroscopic procedure, expanded in culture, and then cultured on a collagen membrane for two days. Implantation was performed with the cells facing the subchondral bone. Follow up evaluations were made through 24 months post-procedure. Outcomes on the KOOS subscales and the VAS pain score were statistically better in the MSC group than the MACI group (p<0.05) at the six-month follow up, although it is not clear if the difference observed would be considered clinically significant. Studies with larger samples sizes and follow-up supported by histological analyses are necessary to determine long-term outcomes of this treatment.

**Section Summary**

The evidence base on MSCs for cartilage repair is increasing, although nearly all studies to date have been performed in Asia with a variety of methods of MSC preparation. Four randomized studies have reported improvements in histologic and morphologic outcomes. Three of these studies also reported improvements in functional outcomes. A meta-analysis of 13 studies found a consistent reduction of pain in groups treated with MSCs, although the studies included were highly heterogeneous and did not consistently distinguish between improvements due to MSCs and those due to pain medication. The method of preparation used in one positive study was to obtain MSCs from bone marrow at the time of microfracture, culture (expand) over a period of three weeks, and then inject into the knee in a carrier of HA. Another randomized trial, using MSCs from peripheral blood, found improvements in histologic and morphologic outcomes, but not functional outcomes, following stimulation with recombinant human granulocyte colony-stimulating factor. A third small RCT found that MSCs from synovial tissue and cultured in collagen resulted in outcomes at least as good as those following MACI.

**FUSION AND NON-UNION**

There is limited evidence on the use of allografts with stem cells for fusion of the extremities or spine or for the treatment of non-union. No RCTs for this indication were identified.

Eastlack (2014) reported outcomes from a series of 182 patients who were treated with anterior cervical discectomy and fusion using Osteocel Plus in a PEEK cage and anterior plating. At 24 months, 74% of patients (180/249 levels treated) were available for follow-up. These patients had significant improvements in clinical outcomes; 87% of levels achieved solid bridging and 92% of levels had range of motion less than 3°. With 26% loss to follow-up at 24 months and lack of a standard of care control group, interpretation of these results is limited.

One retrospective series from 2009 was identified on the use of Trinity MSC bone allograft for revision surgery of the foot and ankle. Twenty-three patients were included who had
undergone revision foot and/or ankle surgery for residual malunion, non-union, or significant segmental bone loss. Patients were followed to the point of radiographic and clinical union, which occurred at a median of 72.5 days for 21 of the 23 patients (91.3%). However, these outcomes do not permit conclusions because of a lack of a control group for comparison with patients who received stem-cell therapy.

**Section Summary**

Current evidence is insufficient to determine whether the use of stem cell results in superior outcomes such as higher fusion rates, or lower rates of reoperations and adverse events.

**MENISCETOMY**

Vangsness (2014) reported an industry-sponsored phase 1/2 randomized, double-blind, multicenter study of cultured allogeneic MSCs (Chondrogen™, Osiris Therapeutics) injected into the knee after partial meniscectomy.[21] The 55 patients were randomized to intra-articular injection of either 50´106 allogeneic MSCs, 150´106 allogeneic MSCs in HA, or HA vehicle control at 7 to 10 days after meniscectomy. The cultured MSCs were derived from bone-marrow aspirates from unrelated donors. At two-year follow-up, three patients in the low-dose MSC group had significantly increased meniscal volume measured by MRI (with an a priori determined threshold of at least 15%) compared to none in the control group and none in the high-dose MSC group. There was no significant difference between the groups in the Lysholm Knee Scale. On subgroup analysis, patients with osteoarthritis who received MSCs had a significantly greater reduction in pain at two years compared with patients who received HA alone. This appears to be a post hoc analysis and should be considered preliminary. No serious adverse events were thought to be related to the investigational treatment.

**Section Summary**

Current evidence for the use of stem cells as an adjunct to meniscectomy is limited to a single preliminary RCT. The outcomes of this study must be validated in large, long-term, randomized controlled trials.

**OSTEONECROSIS**

Several randomized comparative trials have been identified that evaluated the use of MSCs for osteonecrosis of the femoral head.

**Osteonecrosis: MSCs Expanded from Bone Marrow**

Zhao (2012) reported a randomized trial that included 100 patients (104 hips) with early stage femoral head osteonecrosis treated with core decompression and expanded bone marrow MSCs versus core decompression (CD) alone.[22] At 60 months after surgery, two of the 53 hips (3.7%) treated with MSCs continued to have progressive disease and underwent vascularized bone grafting, compared with 10 of 44 hips (23%) in the decompression group who had disease progression and underwent either vascularized bone grafting (n=5) or total hip replacement (n=5). In addition, treatment with MSC improved Harris Hip scores compared to CD and decreased the volume of the necrotic lesion of the hips preoperatively classified at stage IC, IIB, and IIC (p<0.05, respectively; stage IIA, P=0.06, respectively).

**Osteonecrosis: MSCs Concentrated from Bone Marrow**
A 2017 randomized, double-blind trial was conducted using autologous bone marrow concentrate in 38 patients with stage three osteonecrosis.[23] A control group of core decompression plus saline injection was compared to patients receiving core decompression plus MSC implantation. The primary outcome was needing total hip replacement and secondary outcomes were clinical symptoms such as pain and functional ability. There was no difference between groups on any outcomes including total hip replacement requirements, clinical tests, or radiologic evidence.

Another small trial randomized 40 patients (51 hips) with early stage femoral head osteonecrosis to core decompression plus concentrated bone marrow MSCs or core decompression alone.[24] Blinding of assessments in this small trial was not described. Harris Hip Score (HHS) was significantly improved in the MSC group (scores of 83.65 and 82.42; p<0.05) compared with core decompression (scores of 76.68 and 77.39). Kaplan-Meier analysis showed improved hip survival in the MSC group (mean of 51.9 weeks) compared with the core decompression group (mean of 46.7 weeks). There were no significant differences between the groups in the radiographic assessment or MRI results. The conflicting report of improvement via HHS compared to no observable improvement via MRI, may point to the need for study blinding to control for confounding bias toward treatment.

Section Summary

Two small studies reported improvement in the Harris Hip Score in patients with osteonecrosis of the femoral head treated with core decompression and MSCs, although it was not reported if the patients or investigators were blinded to the treatment group. Hip survival was significantly improved following treatment with either expanded or concentrated MSCs. The effect appears to be larger with expanded MSCs compared with concentrated MSCs. However, a double-blind RCT found no difference between MSC treatment or saline injection, when combined with core decompression. Additional studies with a larger number of patients are needed to permit greater certainty regarding the effect of this treatment on health outcomes.

PRACTICE GUIDELINE SUMMARY

American College of Rheumatology and Arthritis Foundation

In 2019, guidelines from the American College of Rheumatology and Arthritis Foundation on osteoarthritis (OA) of the hand, hip, and knee gave a strong recommendation against stem cell injections in patients with knee and/or hip OA, noting the heterogeneity in preparations and lack of standardization of techniques.[25] No recommendation was made for hand OA, since efficacy of stem cells has not been evaluated.

American Academy of Orthopaedic Surgeons

A 2020 guideline from American Association of Orthopaedic Surgeons on the management of glenohumeral joint OA, endorsed by several other societies, states that injectable biologics such as stem cells cannot be recommended in the treatment glenohumeral joint OA.[26] There was consensus from the panel that better standardization and high-quality evidence from clinical trials is needed to provide definitive evidence on the efficacy of biologics in glenohumeral OA. The strength of evidence was rated as no reliable scientific evidence to determine benefits and harms. The 2013 guideline on treatment of osteoarthritis of the knee does not address stem cell injections.
American Association of Neurological Surgeons

In 2014, the American Association of Neurological Surgeons guidelines on fusion procedures for degenerative disease of the lumbar spine relevant to this evidence review have indicated that “The use of demineralized bone matrix (DBM) as a bone graft extender is an option for 1- and 2-level instrumented posterolateral fusions. Demineralized Bone Matrix: Grade C (poor level of evidence).”[27]

Summary

There is not enough research to know if or how well mesenchymal stem cells (MSCs) or allograft bone products containing stem cells work to treat people with orthopedic conditions. No clinical guidelines based on research recommend MSC treatment or allograft bone products containing stem cells for people with orthopedic conditions. Therefore, use of stem cells for orthopedic applications is considered investigational.

References

6. H Iijima, T Isho, H Kuroki, M Takahashi, T Aoyama. Effectiveness of mesenchymal stem cells for treating patients with knee osteoarthritis: a meta-analysis toward the establishment of effective regenerative rehabilitation. NPJ Regenerative medicine. 2018;3:15. PMID: 30245848


## CODES

**NOTE:** There are no specific codes for orthopedic applications of stem cell therapy. The appropriate CPT code for reporting this procedure is 20999, or the code for an unlisted procedure of the body area on which the procedure is performed.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0566T</td>
<td>Autologous cellular implant derived from adipose tissue for the treatment of osteoarthritis of the knees; tissue harvesting and cellular implant creation</td>
</tr>
<tr>
<td></td>
<td>0717T</td>
<td>Injection of cellular implant into knee joint using ultrasound guidance, unilateral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Autologous adipose-derived regenerative cell (ADRC) therapy for partial thickness rotator cuff tear; adipose tissue harvesting, isolation and preparation of harvested cells, including incubation with cell dissociation enzymes, filtration, washing and concentration of ADRCs</td>
</tr>
<tr>
<td>Codes</td>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>0718T</td>
<td></td>
<td>Autologous adipose-derived regenerative cell (ADRC) therapy for partial thickness rotator cuff tear; injection into supraspinatus tendon including ultrasound guidance, unilateral</td>
</tr>
<tr>
<td>0737T</td>
<td></td>
<td>Xenograft implantation into the articular surface</td>
</tr>
<tr>
<td>20939</td>
<td></td>
<td>Bone marrow aspiration for bone grafting, spine surgery only, through separate skin or fascial incision (List separately in addition to code for primary procedure)</td>
</tr>
<tr>
<td>20999</td>
<td></td>
<td>Unlisted procedure, musculoskeletal system, general</td>
</tr>
<tr>
<td>21899</td>
<td></td>
<td>Unlisted procedure, neck or thorax</td>
</tr>
<tr>
<td>22899</td>
<td></td>
<td>Unlisted procedure, spine</td>
</tr>
<tr>
<td>23929</td>
<td></td>
<td>Unlisted procedure, shoulder</td>
</tr>
<tr>
<td>24999</td>
<td></td>
<td>Unlisted procedure, humerus or elbow</td>
</tr>
<tr>
<td>25999</td>
<td></td>
<td>Unlisted procedure, forearm or wrist</td>
</tr>
<tr>
<td>26989</td>
<td></td>
<td>Unlisted procedure, hands or fingers</td>
</tr>
<tr>
<td>27299</td>
<td></td>
<td>Unlisted procedure, pelvis or hip joint</td>
</tr>
<tr>
<td>27599</td>
<td></td>
<td>Unlisted procedure, femur or knee</td>
</tr>
<tr>
<td>27899</td>
<td></td>
<td>Unlisted procedure, leg or ankle</td>
</tr>
<tr>
<td>28899</td>
<td></td>
<td>Unlisted procedure, foot or toes</td>
</tr>
<tr>
<td>29999</td>
<td></td>
<td>Unlisted procedure, arthroscopy</td>
</tr>
<tr>
<td>38206</td>
<td></td>
<td>Blood-derived hematopoietic progenitor cell harvesting for transplantation, per collection; autologous</td>
</tr>
<tr>
<td>38230</td>
<td></td>
<td>Bone marrow harvesting for transplantation; allogeneic</td>
</tr>
<tr>
<td>38232</td>
<td></td>
<td>Bone marrow harvesting for transplantation; autologous</td>
</tr>
<tr>
<td>38241</td>
<td></td>
<td>Bone Marrow or blood-derived peripheral stem cell transplantation; autologous</td>
</tr>
</tbody>
</table>

**Date of Origin:** September 2011
**Coverage of Treatments Provided in a Clinical Trial**

**Effective:** January 1, 2022

**Next Review:** November 2022  
**Last Review:** November 2021

**IMPORTANT REMINDER**

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

**DESCRIPTION**

Effective January 1, 2014, the Affordable Care Act (ACA) requires group health plans or a health insurance issuer offering group or individual health insurance coverage to provide coverage for routine patient costs associated with participating in an approved clinical trial.[1] This policy is written to assist in applying Sec. 2709 of the ACA, Coverage for Individuals Participating in Approved Clinical Trials.

**MEDICAL POLICY CRITERIA**

Routine patient costs associated with approved clinical trials may be considered medically necessary for qualified individuals with respect to treatment of cancer or other life threatening disease or condition, when the Affordable Care Act definitions for clinical trial participation are met.

- See Background for definitions.
- See Policy Guidelines for clinical trial registry resource.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

**POLICY GUIDELINES**

[ClinicalTrials.gov](https://clinicaltrials.gov) includes a registry of publicly and privately supported clinical studies.
LIST OF INFORMATION NEEDED FOR REVIEW

It is critical that the list of information below is submitted for review to determine if the policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

1. Pertinent History and Physical, including specific diagnosis and treatment history
2. Clinical trial name and the NCT number
3. Phase of the trial
4. Currently planned, requested interventions
5. Anticipated possible interventions

CROSS REFERENCES

1. COVID-19 Testing, Laboratory, Policy No. 74

BACKGROUND

DEFINITIONS

• Routine patient costs
  o Routine patient costs *include* all items and services consistent with the coverage provided in the plan (or coverage) that is typically covered for a qualified individual who is not enrolled in a clinical trial.
  o Routine patient costs *do not include* the investigational item, device, or service, itself; items and services that are provided solely to satisfy data collection and analysis needs and that are not used in the direct clinical management of the patient; or a service that is clearly inconsistent with widely accepted and established standards of care for a particular diagnosis.

• Approved clinical trial

An approved clinical trial is defined as a phase I, phase II, phase III, or phase IV clinical trial that is conducted in relation to the prevention, detection, or treatment of cancer or other life-threatening disease or condition, and that is described by any of the following:

  o The study or investigation is approved or funded by one or more of the following:
    ▪ The National Institutes of Health
    ▪ The Centers for Disease Control and Prevention
    ▪ The Agency for Health Care Research and Quality
    ▪ The Centers for Medicare & Medicaid Services
    ▪ A cooperative group or center of any of the above four entities or the Department of Defense or the Department of Veterans Affairs
    ▪ A qualified non-governmental research entity identified in the guidelines issued by the National Institutes of Health for center support grants
    ▪ The Department of Veterans Affairs, the Department of Defense or the Department of Energy if the study or investigation has been reviewed and approved through a system of peer review that the Secretary determines to be comparable to the system of peer review of studies and

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
investigations used by the National Institutes of Health, and assures unbiased review of the highest scientific standards by qualified individuals who have no interest in the outcome of the review; OR

- The study or investigation is conducted under an investigational new drug application reviewed by the Food and Drug Administration; OR
- The study or investigation is a drug trial that is exempt from having such an investigational new drug application.

- Life-threatening condition

A life-threatening condition is defined as any disease or condition from which the likelihood of death is probable unless the course of the disease or condition is interrupted.

- Qualified individual

A participant who is a beneficiary in a health plan who is eligible to participate in an approved clinical trial according to the trial protocol with respect to treatment of cancer or another life threatening disease or condition and either:

- The referring health care professional is a participating health care provider and has concluded that the individual’s participation in such trial would be appropriate based upon the individual meeting the clinical trial eligibility requirements; or
- The participant or beneficiary provides medical and scientific information establishing that the individual’s participation in such trial would be appropriate based upon the individual meeting the clinical trial eligibility requirements.

REFERENCES


CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S9988</td>
<td></td>
<td>Services provided as part of a phase I clinical trial</td>
</tr>
<tr>
<td>S9990</td>
<td></td>
<td>Services provided as part of a phase II clinical trial</td>
</tr>
<tr>
<td>S9991</td>
<td></td>
<td>Services provided as part of a phase III clinical trial</td>
</tr>
</tbody>
</table>

Date of Origin: November 2013
Confocal Laser Endomicroscopy

Effective: November 1, 2021

Next Review: July 2022
Last Review: September 2021

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Confocal laser endomicroscopy (CLE), also known as confocal fluorescent endomicroscopy and optical endomicroscopy, allows in vivo microscopic imaging of cells during endoscopy. CLE is proposed for a variety of purposes, especially as a real-time alternative to histology during colonoscopy and for targeting areas to undergo biopsy in patients with inflammatory bowel disease and Barrett esophagus.

MEDICAL POLICY CRITERIA

Use of confocal laser endomicroscopy is considered investigative for all indications.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Analysis of Human DNA in Stool Samples as a Technique for Colorectal Cancer Screening, Genetic Testing, Policy No. 12
2. In Vivo Analysis of Colorectal Polyps, Medicine, Policy No. 104
BACKGROUND

CLE involves using light from a low-power laser to illuminate tissue and, subsequently, the same lens detects light reflected from the tissue through a pinhole. The term confocal refers to having both illumination and collection systems in the same focal plane. Light reflected and scattered at other geometric angles that is not reflected through the pinhole is excluded from detection, which dramatically increases the special resolution of CLE images.

Endoscope-based and probe-based systems have been cleared by the U.S. Food and Drug Administration (FDA). Endoscope-based systems incorporate a confocal probe onto the tip of a conventional endoscope. Image collection scan rates vary by device. Probe-based systems place a probe through the biopsy channel of a conventional endoscope. Depth of imaging and field of view varies by device. As pointed out in review articles, the limited viewing area emphasizes the need for careful conventional endoscopy to target the areas for evaluation. Both CLE systems are optimized using a contrast agent. The most widely used agent is intravenous fluorescein, which is FDA-approved for ophthalmologic imaging of blood vessels when used with a laser scanning ophthalmoscope.

Unlike techniques such as chromoendoscopy, which are primarily intended to improve the sensitivity of colonoscopy, CLE is unique in that it is designed to immediately characterize the cellular structure of lesions. CLE can thus potentially be used to make a diagnosis of polyp histology, particularly in association with screening or surveillance colonoscopy, which could allow for small hyperplastic lesions to be left in place rather than removed and sent for histologic evaluation. This would reduce risks associated with biopsy and reduce the number of biopsies and histologic evaluations. Another key potential application of CLE technology is targeting areas for biopsy in patients with Barrett esophagus undergoing surveillance endoscopy. This is an alternative to conducting random biopsies during surveillance and has the potential to reduce the number of biopsies and/or improve the detection of dysplasia. Other potential uses of CLE under investigation include better diagnosis and differentiation of conditions such as gastric metaplasia, lung cancer, and bladder cancer.

As noted previously, limitations of CLE systems include a limited viewing area and depth of view. An additional limitation is the lack of standardized systems for classifying lesions viewed with CLE devices. Although there is not currently an internationally accepted classification system for colorectal lesions, two systems have been developed that have been used in a number of studies conducted in different countries. These are the Mainz criteria for endoscopy-based CLE devices and the Miami classification system for probe-based CLE devices.[1] Lesion classification systems are less developed for non-gastrointestinal lesions viewed by CLE devices, e.g., those in the lung or bladder. Another potential limitation of CLE is the learning curve for obtaining high-quality images and classifying lesions. Although several recent studies have found that the ability to acquire high-quality images and interpret them accurately can be learned relatively quickly, these studies were limited to colorectal applications of CLE.[2, 3]

Regulatory Status

Several CLE devices have been cleared for marketing by the FDA. These include:

**Cellvizio® (Mauna Kea Technologies):** This device consists of a confocal laser system, proprietary software, a flat-panel display and miniaturized fiber optic probes. Since 2006, Mauna Kea has received ten FDA approvals for Cellvizio® systems, most recently in May 2016.
EC-3870CLIK Confocal Video Colonoscope (Pentax Medical Company): This is an endoscopy-based CLE system which consists of the EC-3870CLIK, Confocal Video Colonoscope (K042741) and the ISC-1000 Pentax Confocal Laser System (K042740). The device must be used with a Pentax Video Processor. According to FDA materials, the intended use of the device is to provide optical and microscopic visualization of and therapeutic access to the lower gastrointestinal tract.

EVIDENCE SUMMARY

COLORECTAL LESIONS

Ideally, the evaluation of the safety and efficacy of confocal laser endomicroscopy (CLE) as a diagnostic tool would be based on randomized controlled trials (RCTs) comparing CLE to conventional diagnostic methods, such as biopsy with histology for analysis of colorectal lesions. The evidence for the use of CLE is best evaluated in the framework of a diagnostic test, as the test provides diagnostic information that assists in treatment decisions. Validation of the clinical use of any diagnostic test focuses on three main principles:

1. The analytic validity of the test, which refers to the technical accuracy of the test in detecting abnormal histology that is present or in excluding an abnormality that is absent;
2. The clinical validity of the test, which refers to the diagnostic performance of the test (sensitivity, specificity, positive and negative predictive values) in detecting clinical disease; and
3. The clinical utility of the test, i.e., how the results of the diagnostic test will be used to change management of the patient and whether these changes in management lead to clinically important improvements in health outcomes.

Multiple studies have evaluated the diagnostic accuracy of CLE for patients undergoing screening or surveillance colonoscopy. Several systematic reviews of studies evaluating the diagnostic accuracy of CLE compared to a reference standard have been published. Descriptions of several systematic reviews and representative diagnostic accuracy studies are included below.

Systematic Reviews

A 2018 systematic review by Lord analyzed the diagnostic accuracy of several optical imaging techniques for in vivo lesion characterization in colonic inflammatory bowel disease (IBD). A total of 22 studies were identified assessing performance of virtual chromoendoscopy, dye-based chromoendoscopy, magnification endoscopy, and confocal laser endomicroscopy. A bivariate meta-analysis was performed. Pooled sensitivities of real-time CLE, magnification endoscopy, virtual chromoendoscopy, and dye-based chromoendoscopy were 91% (95% CI 66% to 98%), 90% (95% CI 77% to 96%), 86% (95% CI 62% to 95%), and 67% (95% CI 44% to 84%), respectively. Pooled specificities were 97% (95% CI 94% to 98%), 87% (95% CI 81% to 91%), 87% (95% CI 72% to 95%), 86% (95% CI 72% to 94%), for the same methods, respectively. The authors concluded that real-time CLE is highly accurate for differentiating neoplastic from non-neoplastic lesions in patients with colonic IBD, but also note that most
CLE studies were performed by single expert users within tertiary centers, which may confound results.

In 2013, Su reviewed studies on the efficacy of CLE for discriminating colorectal neoplasms from non-neoplasms.[5] Studies needed to use histologic biopsy as the reference standard and in which the pathologist and endoscopist were blinded to each other’s findings. Included studies also used a standardized CLE classification system. Patient populations included individuals at increased risk of colorectal cancer due to personal or family history, patients with previously identified polyps, and/or patients with IBD. Two reviewers independently assessed the quality of individual studies using the modified Quality Assessment Of Diagnostic Accuracy Studies (QUADAS) tool, and studies considered to be at high risk of bias were excluded from further consideration. A total of 15 studies with 719 adult patients were found to be eligible for the systematic review. All were single-center trials and two were available only as abstracts. In all the studies, suspicious lesions were first identified by conventional white-light endoscopy with or without chromoendoscopy and then further examined by CLE. A pooled analysis of the 15 studies found an overall sensitivity of CLE of 94% (95% CI 0.88 to 0.97) and specificity of 95% (95% CI 0.89 to 0.97), compared to histology. Six of the studies included patients at increased risk of colorectal cancer (CRC) who were undergoing surveillance endoscopy, five studies included patients with colorectal polyps and four studies included patients with IBD. In a predefined subgroup analysis by indication for screening, the pooled sensitivity and specificity for surveillance studies was 94% (95% CI, 90% to 97%) and 98% (95% CI 97% to 99%), respectively. For patients presenting with colorectal polyps, the pooled sensitivity of CLE was 91% (95% CI 87% to 94%) and specificity was 85% (95% CI 78% to 90%). For patients with IBD, the pooled sensitivity was 83% (95% CI 70% to 92%) and specificity was 90% (95% CI 87% to 93%). In other predefined subgroup analyses, the summary sensitivity and specificity was significantly higher (p<0.001) in studies of endoscopy-based CLE (97% and 99%, respectively) than studies of probe-based CLE (87% and 82%, respectively). In addition, the summary sensitivity and specificity was significantly higher (p<0.01) with real-time CLE in which the macroscopic endoscopy findings were known (96% and 97%, respectively) than with blinded CLE in which recorded confocal images were subsequently analyzed without knowledge of macroscopic endoscopy findings (85% and 82%, respectively).

Another systematic review was published in 2013 by Dong.[6] The investigators included studies that assessed the diagnostic accuracy of CLE compared with conventional endoscopy. They did not explicitly state that the reference standard was histologic biopsy, but this was the implied reference standard. A total of six studies were included in a meta-analysis. All of the studies were prospective, and at least five included blinded interpretation of CLE findings (in one study, it was unknown whether interpretation was blinded). In a pooled analysis of data from all six studies, the sensitivity was 81% (95% CI 77% to 85%) and the specificity was 88% (95% CI 85% to 90%). The authors also conducted a subgroup analysis by type of CLE used. When findings from the two studies on endoscopy-based CLE were pooled, the sensitivity was 82% (95% CI 69% to 91%) and the specificity was 94% (95% CI 91% to 96%). Two studies may not have been a sufficient number to obtain a reliable estimate of diagnostic accuracy. When findings from the 4 studies on probe-based endoscopy were pooled, the sensitivity was 81% (95% CI 76% to 85%) and the specificity was 75% (95% CI 69% to 81%).

A 2013 systematic review by Wanders searched for studies that reported diagnostic accuracy of studies on any of several new technologies used to differentiate between colorectal neoplasms and non-neoplasms.[7] To be included in the review, studies needed to use the technology to differentiate between non-neoplastic and neoplastic lesions and to use
histopathology as the reference standard. Blinding was not an inclusion criterion. Eleven eligible studies were identified that included an analysis of CLE. A pooled analysis of study findings yielded an estimated sensitivity of 93.3% (95% CI 88.4 to 96.2) and a specificity of 89.9% (95% CI 81.8% to 94.6%). A meta-analysis limited to the five studies that used endoscopy-based CLE found a sensitivity of 94.8% (95% CI 90.6% to 98.92%) and a specificity of 94.4% (95% CI 90.7% to 99.2%). When findings of the six studies on probe-based CLE were pooled, the sensitivity was 91.5% (86.0% to 97.0%) and the specificity was 80.9 (95% CI 69.4% to 92.4%).

Nonrandomized Studies

Ohmiya (2017) evaluated the ability of CLE to differentiate among ulcerative colitis (UC)-associated neoplasia (differentiated type or undifferentiated type), sporadic adenoma, and circumscribed regenerative lesions. The authors examined 12 patients with suspected UC-associated neoplasia with probe-based CLE and compared findings with pathological diagnoses determined by magnifying chromoendoscopy with crystal violet and narrow band imaging. Sensitivity, specificity, and accuracy of CLE were 100%, 83%, and 92%, respectively. The authors stated that CLE was helpful in evaluating suspected UC-associated neoplasia, but it is limited by the small sample size.

In 2017, Kim evaluated probe-based CLE for feasibility and safety in evaluating colorectal submucosa following removal of colorectal neoplasms. Colorectal submucosa were classified as negative or indicative of carcinoma infiltration. The results were compared to pathological findings. The sensitivity, specificity, and accuracy of the classifications were 91.7%, 86.8%, and 88.0%, respectively. The authors concluded that CLE is useful but that large-scale prospective studies are needed.

In a 2012 study by Shadid two methods of analyzing CLE images, real-time diagnosis and blinded review of video images after endoscopy (known as "offline" diagnosis), were compared. The study included 74 patients with a total of 154 colorectal lesions. Eligibility criteria were similar to the Buchner study (see above); the included patients undergoing surveillance or screening colonoscopy. Patients underwent white-light colonoscopy and identified polyps were also evaluated with virtual chromoendoscopy and probe-based CLE. Intravenous fluorescein sodium was administered after the first polyp was identified. At the time of examination, an endoscopist made a real-time diagnosis based on CLE images. Based on that diagnosis, the patient underwent polypectomy, biopsy or endoscopic mucosal resection, and histopathologic analysis was done on the specimens. The CLE images were then de-identified and then reviewed offline by the same endoscopist at least one month later. At the second review, the endoscopist was blinded to the endoscopic and histopathologic diagnosis. Of the 154 polyps, 74 were found by histopathologic analysis to be non-neoplastic and 80 were neoplastic (63 tubular adenomas, 12 tubulovillous adenomas, three mixed hyperplastic-adenoma polyps and two adenocarcinomas). Overall, there was not a statistically significant difference in the diagnostic accuracy of real-time CLE diagnosis and blinded offline CLE diagnosis (i.e., confidence intervals overlapped). The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for real-time CLE diagnosis was 81%, 76%, 87%, and 79%, respectively. For offline diagnosis, these numbers were 88%, 77%, 81% and 85%, respectively. However, in the subgroup of 107 smaller polyps, less than 10 mm in size, the accuracy of real-time CLE was significantly lower than offline CLE. For the smaller polyps, sensitivity, specificity, PPV and NPV of real-time CLE was 71%, 83%, 78%, and 78% and for offline CLE was 86%, 78%, 76%, and 87%, all respectively. For larger polyps, in
contrast, there was a nonsignificant trend in favor of better diagnostic accuracy with real-time compared to offline CLE.

A 2011 study by Hlavaty included patients with ulcerative colitis or Crohn disease. Thirty patients were examined with standard white-light colonoscopy, chromoendoscopy and an endoscopy-based CLE system. An additional 15 patients were examined only with standard colonoscopy. All lesions identified by white-light colonoscopy or chromoendoscopy were examined using CLE to identify neoplasia using the Mainz classification system. Suspicious lesions underwent biopsy and, additionally, random biopsies were taken from four quadrants every 10 cm per the standard surveillance colonoscopy protocol. All specimens underwent histologic analysis by a gastrointestinal pathologist who was blinded to the CLE diagnosis. Diagnostic accuracy of CLE was calculated for examinable lesions only. Compared to histologic diagnosis, the sensitivity of CLE for diagnosing low-grade and high-grade intraepithelial neoplasia was 100%, the specificity was 98.4%, the PPV was 66.7%, and the NPV was 100%. However, whereas CLE was able to examine 28 of 30 (93%) flat lesions, it could examine only 40 of 70 (57%) protruding polyps. Moreover, 6 of 10 (60%) dysplastic lesions, including three of five low-grade and high-grade intraepithelial neoplasms were not evaluable by CLE. It is also worth noting that the diagnostic accuracy of chromoendoscopy was similar to that of CLE. The sensitivity, specificity, PPV and NPV of chromoendoscopy was 100%, 97.9%, 75%, and 100%, respectively.

A 2011 study by Xie included 116 consecutive patients who had polyps found during CLE; one patient was excluded from the analysis. All patients had an indication for colonoscopy (19 were undergoing surveillance postpolypectomy, two had a family history of colorectal cancer, three had IBD and 91 were seeking a diagnosis). All patients first underwent white-light colonoscopy. Endoscopy-based CLE was used on the first polyp identified during withdrawal of the endoscope (i.e., one polyp per patient was analyzed). Intravenous fluorescein sodium was used. Real-time diagnosis of the polyp was performed based on criteria used at the study center (which is adapted from the Mainz classification system). The polyps were biopsied or were removed and histopathologic diagnosis was determined. Real-time CLE diagnosis correctly identified 109 of 115 (95%) adenomas or hyperplastic polyps. Four adenomas were misdiagnosed by CLE as hyperplastic polyps (two were tubulovillous adenomas and two were tubulovillous adenomas) and two hyperplastic polyps were misdiagnosed as adenomas. The overall sensitivity, specificity, PPV, and NPV of CLE diagnosis was 93.9% (95% CI 85.4% to 97.6%), 95.9% (95% CI 86.2% to 98.9%), 96.9% (95% CI 89% to 99%), and 94.8% (95% CI 89.1% to 97.6%), respectively. For polyps less than 10 mm, the CLE diagnosis had a sensitivity of 90.3% and specificity of 95.7%, and for polyps 10 mm and larger, sensitivity was 97.1% and specificity was 100%.

In 2010, Buchner published findings on 75 patients who had a total of 119 polyps. Patients were eligible for study participation if they were undergoing surveillance or screening colonoscopy or undergoing evaluation of known or suspected polyps identified by other imaging modalities or endoscopic resection of larger flat colorectal neoplasia. White-light colonoscopy was used as the primary screening method. When a suspicious lesion was identified, it was evaluated by virtual chromoendoscopy system and a probe-based CLE system. Intravenous fluorescein sodium was administered after the first polyp was identified. Following the imaging techniques, the appropriate intervention, i.e., polypectomy, biopsy, or endoscopic mucosal resection, of lesions were performed and all resected specimens underwent histopathologic analysis by a pathologist blinded to CLE information. Confocal images of the 199 polyps were evaluated after all procedures were completed; the evaluator
was blinded to histology diagnosis and endoscopic appearance of the lesion. Diagnosis of confocal images used modified Mainz criteria; polyps were classified as benign or neoplastic. According to histopathologic analysis, there were 38 hyperplastic polyps and 81 neoplastic lesions (58 tubular adenomas, 15 tubulovillous adenomas and 4 adenocarcinomas). CLE correctly identified 74 of 81 neoplastic polyps (sensitivity, 91%; 95% CI 83% to 96%). In addition, CLE correctly identified 29 of 38 hyperplastic polyps (specificity, 76%; 95% CI, 60% to 89%). In contrast, virtual chromoendoscopy correctly identified 62 neoplastic polyps (sensitivity, 77%; 95% CI 66% to 85%) and 27 hyperplastic polyps (specificity, 71%; 95% CI 54% to 85%).

Section Summary

Multiple studies have evaluated the accuracy of confocal laser endoscopy compared with histopathology for diagnosing colorectal lesions. In three published systematic reviews, pooled estimates of overall sensitivity of CLE ranged from 81% to 94% and pooled estimates of specificity ranged from 88% to 95%. Although the reported diagnostic accuracy tended to be relatively high, it is not clear whether the accuracy is high enough to replace biopsy/polypectomy and histologic analysis.

BARRETT ESOPHAGUS

The ideal study would determine whether CLE with targeted biopsy can distinguish Barrett’s Esophagus (BE) without dysplasia from BE with low- and high-grade dysplasia. In addition, study results would need to determine if CLE with target biopsy led to fewer biopsies of benign tissue compared to surveillance with random biopsies. The ideal study to address the above questions would include an unselected clinical population of patients with BE presenting for surveillance and would randomly assign patients to CLE with targeted biopsy or a standard biopsy protocol without CLE. Relevant outcomes include diagnostic accuracy for detecting dysplasia, the detection rate for dysplasia, and the number of biopsies. Several studies with most or all of these elements of study design were identified, including randomized controlled trials (RCTs).

Systematic Reviews

In 2017, Xiong published a systematic review and meta-analysis to assess the accuracy of within-patient comparisons of narrow band imaging and CLE for the diagnosis of high-grade dysplasia and esophageal adenocarcinoma in BE patients.[14] The quality of studies was assessed using the QUADAS-2 tool. A total of five studies with 251 patients were included in the meta-analysis. The pooled sensitivities were not significantly different, with values of 62.8% (95% CI 0.56 to 0.69, I²=94.6%) for narrow band imaging and 72.3% (95% CI 0.66 to 0.78, I²=89.3%) for CLE. Pooled specificities were also not significantly different (narrow band imaging 85.3% [95% CI 0.84 to 0.87, I²=92.1%] vs CLE 83.8% [95% CI 0.82 to 0.85, I²=96.8%]). The pooled additional detection rate of CLE compared to narrow band imaging for per-lesion detection of neoplasia was 19.3% (95% CI 0.05 to 0.33, I²=74.6%).

In 2016, Xiong published a meta-analysis of prospective studies evaluating the diagnostic accuracy of CLE in patients with BE and using histopathologic analysis as the criterion standard.[15] Studies were not required to compare CLE to standard four-quadrant biopsy. Fourteen studies were included. Three were reported to have a high risk of bias and the rest a low risk of bias. There was no statistically significant publication bias. In a pooled analysis of seven studies (n=473 patients) reporting a per-patient analysis, the sensitivity of CLE for
detecting neoplasia was 89% (95% CI, 82% to 94%) and the specificity was 83% (95% CI 78% to 86%). The pooled positive and negative likelihood ratios were 6.53 (95% CI, 3.12 to 13.4) and 0.17 (95% CI 0.11 to 0.29, respectively). Reviewers did not report PPV or NPV. Sensitivity and specificity were similar to those reported below in the 2014 meta-analysis by Gupta. Limitations to this analysis include heterogeneity of the results and a lack of relationship between the diagnostic odds ratio and the characteristics of the studies.

Gupta (2014) conducted a systematic review and meta-analysis to evaluate the diagnostic accuracy of the CLE-based targeted biopsies in detecting high grade dysplasia (HGD)/adenocarcinoma compared with four-quadrant random biopsies.[16] All the studies that compared the diagnostic yield from CLE-based targeted biopsies to detect HGD/adenocarcinoma with a gold standard of histopathology were included and a meta-analysis was carried out to estimate the pooled sensitivity, specificity, and positive and negative likelihood. Seven studies with 345 patients and 3080 lesions were included in the meta-analysis. All the studies had reported per-lesion analyses; however, only four of the seven studies had data reported on per-patient analyses. 'Per-lesion' analysis for the diagnosis of HGD/adenocarcinoma yielded a pooled sensitivity and specificity of 68% (95% CI of 64-73%) and 88% (95% CI 87 to 89%), respectively. The pooled positive and negative likelihood ratios were 6.56 (95% CI 3.61 to 11.90) and 0.24 (95% CI 0.09 to 0.63), respectively. Similar numbers were calculated on the basis of 'per-patient' basis, which showed a pooled sensitivity and specificity of 86% (95% CI 74 to 96%) and 83% (95% CI 77 to 88%), respectively. The pooled positive and negative likelihood ratios were 5.61 (95% CI 2.00 to 15.69) and 0.21 (95% CI 0.08 to 0.59), respectively. Authors noted that CLE, by providing targeted biopsies, has a good diagnostic accuracy in identifying HGD/EAC; however, the overall prevalence of HGD/EAC in the studies included was much higher than what would be seen in clinical practice and these results should be interpreted with caution. Due to its relatively low sensitivity and negative predictive value, CLE may currently not replace standard biopsy techniques for the diagnosis of HGD/EAC in Barrett's esophagus.

In 2013, a meta-analysis by Wu of observational studies and RCTs focused on the diagnostic accuracy of CLE for detecting neoplasia in BE patients.[17] In a pooled analysis of data from four studies that reported per-patient accuracy of CLE, the pooled sensitivity for detection of neoplasia was 89% (95% CI 0.80% to 0.95%), and the pooled specificity was 75% (95% CI 69% to 81%). Seven studies reported per-location accuracy of CLE. The pooled sensitivity for CLE was 70% (95% CI 65% to 74%) and the pooled specificity was 91% (95% CI 90% to 92%). This study did not address other outcomes such as number of biopsies and did not compare CLE for detection of neoplasia in patients with BE with white-light endoscopy.

Randomized Controlled Trials

In 2013, Canto published findings from a single-blind multicenter RCT conducted at academic centers with experienced endoscopists.[18] The trial included consecutive patients undergoing endoscopy for routine surveillance of BE or for suspected or known neoplasia. Patients were randomized to high-definition white-light endoscopy with random biopsy (n=98) or white-light endoscopy with endoscopy-based CLE and targeted biopsy (n=94). In the white-light endoscopy-only group, four-quadrant random biopsies were taken every one to two cm of the entire length of the BE for patients undergoing surveillance and every one cm in patients with suspected neoplasia. In the CLE group, biopsy specimens were obtained only when there was CLE evidence of neoplasia. The final pathology diagnosis was the reference standard. A per-patient analysis of diagnostic accuracy for diagnosing BE-related neoplasia found a sensitivity...
of 40% with white-light endoscopy alone and 95% with white-light endoscopy plus CLE. Specificity was 98% with white-light endoscopy alone and 92% with white-light endoscopy plus CLE. When the analysis was done on a per-biopsy specimen basis, when CLE was added, the sensitivity was substantially higher and the specificity was slightly lower. The median number of biopsies per patient was significantly higher in the white-light endoscopy group compared with the group that also received CLE (4 vs 2, p<0.001). The investigators conducted an analysis of the number of cases in which CLE resulted in a different diagnosis. Thirty-two of 94 (34%) patients in the white-light plus CLE group had a correct change in dysplasia grade after CLE compared to the initial endoscopic findings. Six of the 32 (19%) patients had lesions and the remaining 26 did not. In 21 of the 26 patients without lesions, CLE changed the plan from biopsy to no biopsy. The remaining 62 of 94 (65%) patients in the white-light endoscopy plus CLE group had concordant diagnoses with the two techniques. The study was conducted at academic centers and used endoscopy-based CLE. Findings may not be generalizable to other clinical settings or to probe-based CLE.

In 2011, Sharma published an international, multicenter RCT that included 122 consecutive patients presenting for surveillance of BE or endoscopic treatment of high-grade dysplasia or early carcinoma.[19] This study was described in the systematic review and meta-analysis described by Gupta in the previous section. Patients were randomly assigned to receive, in random order, both standard white-light endoscopy and narrow-band imaging. Following these two examinations, which were done in a blinded fashion, the location of lesions was unblinded and, subsequently, all patients underwent probe-based CLE. All examinations involved presumptive diagnosis of suspicious lesions. Also, in both groups, after all evaluations were performed, there were biopsies of all suspicious lesions, as well as biopsies of random locations (four quadrants every two cm). Histopathologic analysis was the reference standard. Twenty-one patients were excluded from the analysis. Of the remaining 101 patients, 66 (65%) were found on histopathologic analysis to have no dysplasia, four (4%) had low-grade dysplasia, six (6%) had high-grade dysplasia and 25 (25%) had early carcinoma. The sensitivity of CLE with white-light endoscopy for detecting high-grade dysplasia or early carcinoma was 68.3% (95% CI, 60.0% to 76.7%), which was significantly higher than white-light endoscopy alone; 34.2% (95% CI 25.7% to 42.7%, p=0.002). However, the specificity of CLE and white-light endoscopy was significantly lower than white-light endoscopy alone: 92.7% (95% CI 90.8% to 94.6%) versus 87.8% (95% CI 85.5% to 90.1%; p<0.001). For white-light endoscopy alone, the PPV was 42.7% (32.8% to 52.6%) and the NPV was 89.8% (95% CI 87.7% to 92.0%). For white-light endoscopy with probe-based CLE, the PPV was 47.1% (95% CI 39.7% to 54.5%) and the NPV was 94.6% (95% CI 92.9% to 96.2%). White-light endoscopy alone missed 79 of 120 (66%) areas with high-grade dysplasia or early carcinoma and white-light endoscopy with CLE missed 38 (32%) areas. On a per-patient basis, 31 patients were diagnosed with high-grade dysplasia or early carcinoma. White-light endoscopy alone failed to identify four of these patients (sensitivity, 87%), whereas white-light endoscopy and CLE failed to identify two patients (sensitivity, 93.5%).

Another RCT was published in 2012 by Bertani in Italy; this was a single-center study.[20] The study compared the dysplasia detection rate of biopsies obtained by standard white-light endoscopy only to the detection rate with standard endoscopy followed by probe-based CLE in patients with BE who were enrolled in a surveillance program. One hundred consecutive patients were included, and 50 were randomly assigned to each group. In both groups, targeted biopsies of suspicious lesions and random four-quadrant biopsies (one biopsy every one cm) were taken. The authors described the criteria they used for classifying CLE images as dysplastic or neoplastic. According to histopathologic analysis, the reference standard,
high-grade dysplasia, was diagnosed in three patients and low-grade dysplasia was diagnosed in 16 patients, for an overall detection rate of 19 in 100 (19%) cases. Five cases were in the standard endoscopy group (one case of high-grade dysplasia and four cases of low-grade dysplasia) and 14 were in the CLE group (two cases of high-grade dysplasia and 12 cases of low-grade dysplasia). No suspicious lesions were identified in the standard endoscopy group and thus, only random biopsies were performed. In the CLE group, no suspicious lesions were identified when patients were initially evaluated with standard endoscopy but CLE detected areas suspicious for neoplasia in 21 of 50 (42%) of patients. All the cases of dysplasia were in patients with areas suspicious for neoplasia at CLE but not standard endoscopy. The sensitivity, specificity, PPV and NPV of probe-based CLE for detecting dysplasia were 100%, 83%, 67%, and 100%, respectively. Overall, the mean number of biopsies did not differ between groups (mean of 6.6 per patient in the standard endoscopy group and 6.1 in the CLE group, p=0.77), so the increased detection rate in the CLE group cannot be explained by a larger number of biopsies.

A single-center crossover RCT was published in 2009 by Dunbar. This study was able to evaluate whether CLE can reduce the biopsy rate. This study was described in the systematic review and meta-analysis described by Gupta (2014) in the previous section. Forty-six patients with BE were enrolled, and 39 (95%) completed the study protocol. Of these, 23 were undergoing BE surveillance and 16 had BE with suspected neoplasia. All patients received endoscopy-based CLE and standard endoscopy, in random order. One endoscopist performed all CLE procedures and another endoscopist performed all standard endoscopy procedures; endoscopists were blinded to the finding of the other procedure. During the standard endoscopy procedure, biopsies were taken of any discrete lesions followed by four-quadrant random biopsy (every one cm for suspected neoplasia and every two cm for BE surveillance). During the CLE procedure, only lesions suspicious of neoplasia were biopsied. Endoscopists interpreted CLE images using the Confocal Barrett’s Classification system, developed in a previous research study. Histopathologic analysis was the reference standard. Among the 16 study completers with suspected high-risk dysplasia, there were significantly fewer biopsies per patient with CLE compared to standard endoscopy (mean of 9.8 biopsies vs 23.9 biopsies per patient, p=0.002). Although there were fewer biopsies, the mean number of biopsy specimens showing high-grade dysplasia or cancer was similar in the two groups: 3.1 during CLE and 3.7 during standard endoscopy, respectively. The diagnostic yield for neoplasia was 33.7% with CLE and 17.2% with standard endoscopy. None of the 23 patients undergoing BE for surveillance were found to have high-grade dysplasia or cancer. The mean number of mucosal specimens obtained for patients in this group was 12.6 with white-light endoscopy and 1.7 with CLE (p<0.001).

Nonrandomized Studies

Richardson (2019) conducted a prospective study at eight centers to compare probe-based CLE to conventional histology using the Seattle Protocol (random 4-quadrant biopsy) to identify intestinal metaplasia among 172 patients undergoing screening or surveillance endoscopy for BE. Endoscopists recruited for the study were early users of CLE with less than two years of experience and no formal pathology training. All patients underwent a standardized endoscopy with white light and narrow band imaging evaluation, identification of landmarks, and recording of columnar lined esophagus visualized according to the Prague classification. Patients then received fluorescein followed by optical biopsy; images were interpreted both in real time and immediately following the procedure. After CLE images were acquired, esophageal biopsies were taken via the Seattle Protocol. Endoscopists were able to
identify intestinal metaplasia among 99 patients (57.6%) using CLE compared to 46 patients (27%) using the Seattle Protocol (p<0.0001). Dysplasia was identified in 6 patients using CLE compared to 2 patients using the Seattle Protocol (both of which were also identified via CLE). Confocal laser endomicroscopy also identified significantly more patients with intestinal metaplasia compared to the Seattle Protocol among those with visible columnar lined esophagus (75 vs. 31 patients, respectively; p<0.0001), but not among those without columnar lined esophagus (24 vs. 15 patients; p=0.067). Identification of intestinal metaplasia was not found to be significantly different when comparing CLE to expert review.

Section Summary

Several RCTs and a meta-analysis of RCTs and non-randomized, observational studies suggest that CLE has high accuracy for identifying dysplasia in patients with BE. A 2014 meta-analysis found that the pooled sensitivity, specificity, and negative predictive value in available studies is not sufficiently high to replace the standard Seattle protocol, according to criteria adopted by the American Society for Gastrointestinal Endoscopy (ASGE).

The sensitivity of CLE in the individual studies was higher than for white-light endoscopy alone, but the specificity was not consistently higher. There are limited data comparing standard protocols using random biopsies to protocols using CLE and targeted biopsies, so data are inconclusive regarding the potential for CLE to reduce the number of biopsies in patients with BE undergoing surveillance without compromising diagnostic accuracy. Moreover, studies do not appear to use a consistent approach to classifying lesions viewed using CLE as dysplastic.

PANCREATIC DISEASES

Systematic Reviews

Konjetia (2020) conducted a systematic review to evaluate the diagnostic performance and safety of needle-based confocal laser endomicroscopy (nCLE) in the diagnosis of pancreatic cystic lesions. Seven studies were included in the review with a total sample size of 324 patients. The pooled sensitivity, specificity, positive likelihood ratio, and negative likelihood ration of nCLE was 85% (95% CI, 71-93), 99% (95% CI, 90-100), 78.66 (95% CI, 7.99-774.68), and 0.15 (95% CI, 0.07-0.31) respectively. The diagnostic accuracy as measured by summary receiver operating characteristic curve was 99%. The results showed that nCLE may be effective in diagnostic evaluation of PCLs, however a large amount of heterogeneity was present in the analysis which is consistent with prior reviews.

In 2020, Facciorusso published a meta-analysis of needle-based confocal laser endomicroscopy (nCLE) in pancreatic cystic lesions. Ten studies with a total of 536 patients met inclusion criteria. Three studies were rated as low-quality and the rest as high quality using the Newcastle/Ottawa scale. There was no evidence of publication bias. Diagnostic outcomes from the included studies were pooled using a random-effects mode. Overall pooled diagnostic accuracy was 88.6% (83.7 to 93.4%; I²=41.73%). Pooled sensitivity and specificity of nCLE were calculated from nine studies to be 82.4% and 96.6%, respectively. A direct comparison between the diagnostic sensitivity of nCLE and endoscopic ultrasound-guided fine needle aspiration (FNA) was conducted. No statistically significant difference was reported (OR=1.51, 0.34 to 6.68), although the authors cautioned that there was high heterogeneity.

Also in 2020, Chin published a systematic review on the role of needle-based confocal laser endomicroscopy in the evaluation of pancreatic cystic lesions. Twelve studies were
included, six retrospective and six prospective. No meta-analysis was completed. The accuracy of nCLE was between 46% and 95%, although only one study reported accuracy below 71%. The reported incidence of acute pancreatitis, the most common complication related to nCLE, was 1.3% to 12%.

Nonrandomized Studies

Hao (2020) published a study to evaluate the diagnostic efficacy of EUS-guided nCLE in solid pancreatic lesions (SPLs) and pancreatic cystic lesions (PCLs).[26] A total of 172 patients were enrolled and underwent EUS-nCLE. The reported mean sensitivity, specificity, negative predictive value, positive predictive value and accuracy of the nCLE in diagnosis of pancreatic ductal adenocarcinoma were 90.3%, 89.5%, 93.3%, 85.0% and 90.0%, respectively.

Nakaoka (2020) reported a study of 30 patients who underwent endoscopic retrograde cholangiopancreatography with pCLE for the evaluation of indeterminate pancreatic diseases.[27] Compared to cytology, the diagnostic accuracy (96.7% vs. 76.7%; p=0.0227) and the sensitivity (91.7% vs. 41.7%; p=0.0094) of pCLE for pancreatic ductal adenocarcinoma was significantly higher. The diagnostic accuracy (93.3% vs. 63.3%; p=0.0048) and the specificity (90.9% vs. 50%; p=0.0029) for pancreatitis were significantly higher for pCLE than for cytology. However, the diagnostic accuracies of the two methods did not significantly differ for main duct intrapapillary mucinous neoplasms.

Haghighi (2019) reported results of a study to determine the diagnostic utility of nCLE compared to endoscopic ultrasound-guided FNA (EUS-FNA) for PCLs.[28] A total of 32 patients diagnosed with PCL who had undergone nCLE and FNA over a 10-year period within a major urban teaching hospital were included. The diagnoses in the included patients were serous cystadenoma (n=13), intraductal papillary mucinous neoplasms (n=7), mucinous cystic neoplasms (n=2), well-differentiated neuroendocrine tumors (n=2), cysts (n=2), benign pancreatic lesions (n=2), adenocarcinoma (n=1), gastrointestinal stromal tumor (GIST; n=1), and lymphangioma (n=1). The diagnostic accuracy varied by diagnosis. The highest diagnostic accuracy was for intraductal papillary mucinous neoplasms (n=7, vs. 100% for nCLE compared to 42.8% for EUS-FNA, n=3), while the diagnostic accuracy rate for serous cystadenoma was 69.2% (n=9; vs. 76.9% for EUS-FNA, n=10). Overall, the sensitivity, specificity, PPV, and NPV were 91.7%, 87.5%, 84.6%, and 93.3%, respectively, for nCLE and 80.0%, 92.3%, 88.9%, and 85.7%, respectively, for EUS-FNA.

ASSESSING THE ADEQUACY OF ENDOSCOPIC TREATMENT OF GASTROINTESTINAL LESIONS

Evidence is not clear regarding whether use of CLE improves the determination of residual disease compared with conventional techniques (i.e. white-light endoscopy). In 2014, Ypsilantis published a systematic review of the literature.[29] They included retrospective and prospective studies that reported diagnostic accuracy of CLE for the detection of residual disease after endoscopic mucosal resection (EMR) of gastrointestinal lesions. After examining full-text articles, a total of three studies (one RCT and two prospective, non-randomized comparative studies) met the eligibility criteria. Studies included patients with BE, gastric neoplasia, and colorectal neoplasia. There was significant heterogeneity among studies. In a per-lesion meta-analysis, pooled sensitivity of CLE for detecting neoplasia was 91% (95% CI 83% to 96%), and pooled specificity was 69% (95% CI 61 to 76%). Based on the small number of studies and heterogeneity among studies, the authors concluded that evidence on the
usefulness of CLE in assessing the adequacy of EMR is weak. The single RCT was published in 2012 by Wallace[30] This multicenter trial included patients with BE who were undergoing ablation. After an initial attempt at ablation, patients were randomized to follow-up with either with high-definition white light (HDWL) endoscopy or HDWL endoscopy plus CLE. The primary outcome was the proportion of optimally treated patients, defined as those with no evidence of disease at follow-up, and those with residual disease who were identified and treated. Enrollment in the study was halted after an interim analysis showed no difference between groups. Among the 119 patients who had enrolled by the time of the interim analysis, 15 (26%) of 57 in the HDWL group and 17 (27%) of 62 in the HDWL plus CLE group were optimally treated; the difference was not statistically significant. Moreover, other outcomes were similar in the two groups.

Section summary

There is insufficient evidence that CLE improves upon standard practice for assessing the adequacy of endoscopic treatment of gastrointestinal lesions. The single RCT on this topic was stopped early because an interim analysis reported that CLE did not improve upon high-definition white light endoscopy.

OTHER POTENTIAL APPLICATIONS OF CLE


PRACTICE GUIDELINE SUMMARY

AMERICAN GASTROENTEROLOGICAL ASSOCIATION

In 2011 the American Gastroenterological Association (AGA) published a position statement on the management of Barrett esophagus.[75] The statement includes the following recommendations regarding endoscopic surveillance of Barrett esophagus:

The AGA suggest that endoscopic surveillance be performed in patients with Barrett esophagus (weak recommendation, moderate-quality evidence).

The AGA suggest the following surveillance intervals (weak recommendation, low-quality evidence):

- No dysplasia: three to five years
- Low-grade dysplasia: 6 to 12 months
- High-grade dysplasia in the absence of eradication therapy: three months

For patients with Barrett esophagus who are undergoing surveillance, the AGA recommended:
- Endoscopic evaluation be performed using white light endoscopy (strong recommendation, moderate-quality evidence).
- Four-quadrant biopsy specimens be taken every 2 cm (strong recommendation, moderate-quality evidence).
- Specific biopsy specimens of any mucosal irregularities be submitted separately to the pathologist (strong recommendation, moderate-quality evidence).
- Four-quadrant biopsy specimens be obtained every 1 cm in patients with known or suspected dysplasia (strong recommendation, moderate-quality evidence).

The AGA recommend against requiring chromoendoscopy or advanced imaging techniques for the routine surveillance of patients with Barrett esophagus at this time (weak recommendation, low-quality evidence).

**AMERICAN SOCIETY FOR GASTROINTESTINAL ENDOSCOPY**

In 2019, the American Society for Gastrointestinal Endoscopy (ASGE) published a guideline on screening and surveillance of Barrett’s esophagus.[76] The guideline includes the following recommendation regarding surveillance of dysplasia in patients with Barrett’s esophagus: “In patients with BE undergoing surveillance, we suggest against routine use of CLE compared with WLE with Seattle protocol biopsy sampling (conditional recommendation, low quality of evidence).”

The ASGE published a guideline (2006; reaffirmed in 2011) on the role of endoscopy in the surveillance of premalignant conditions of the upper gastrointestinal (GI) tract.[77] Regarding the use of confocal endoscopy as an adjunct to white-light endoscopy, the guidelines stated that this technique is “still in development.” The guideline also included the following statements on surveillance of patients with BE:

The cost effectiveness of surveillance in patients without dysplasia is controversial. Surveillance endoscopy is appropriate for patients fit to undergo therapy, should endoscopic/histologic findings dictate. For patients with established Barrett’s esophagus of any length and with no dysplasia, after 2 consecutive examinations within 1 year, an acceptable interval for additional surveillance is every 3 years.

Patients with high-grade dysplasia are at significant risk for prevalent or incident cancer. Patients who are surgical candidates may elect to have definitive therapy. Patients who elect surveillance endoscopy should undergo follow-up every 3 months for at least 1 year, with multiple large capacity biopsy specimens obtained at 1 cm intervals. After 1 year of no cancer detection, the interval of surveillance may be lengthened if there are no dysplastic changes on 2 subsequent endoscopies performed at 3-month intervals. High-grade dysplasia should be confirmed by an expert GI pathologist.

Surveillance in patients with low-grade dysplasia is recommended. The significance of low-grade dysplasia as a risk factor for cancer remains poorly defined; therefore, the optimal interval and biopsy protocol has not been established. A follow-up EGD (screening esophagastroduodenoscopy) (i.e., at 6 months) should be performed with concentrated biopsies in the area of dysplasia. If low-grade dysplasia is confirmed, then one possible management scheme would be surveillance at 12 months and yearly thereafter as long as dysplasia persists.
In 2012, the ASGE stated the following in their guideline on the role of endoscopy in BE and other premalignant conditions of the esophagus: "Adjuncts to white-light endoscopy used to improve the sensitivity for the detection of BE and dysplastic BE include chromoendoscopy, electrical enhanced imaging, magnification, and confocal endoscopy."

The ASGE Technology Committee published a Technology Status Evaluation Report on CLE in 2014. The report concluded that CLE is an emerging technology with the potential to improve patient care. However, before the technology can be widely accepted, further studies are needed in the following areas:

- Use of CLE outside of the academic setting, particularly the applicability of the technology in community settings.
- The learning curve of CLE image interpretation and any additional time needed to perform the procedure.
- The clinical efficacy of the technology compared to other available advanced imaging technologies.
- Approaches to CLE imaging and image interpretation.

In 2016, based on a systematic review of 102 studies conducted between 2004 and 2015, the ASGE concluded additional clinical trials on CLE are still necessary.

**SUMMARY**

There is not enough research to know if or how well confocal laser endomicroscopy (CLE) works to improve health outcomes for people with any condition. This does not mean that it does not work, but more research is needed to know. Therefore, use of CLE with endoscopy is considered investigational for all indications.

**REFERENCES**


60. S Han, P Tatman, S Mehrotra, et al. Combination of ERCP-Based Modalities Increases Diagnostic Yield for Biliary Strictures. * Digestive diseases and sciences*. 2020. PMID: 32430658


---

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.
Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


## CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0397T</td>
<td>Endoscopic retrograde cholangiopancreatography (ERCP), with optical endomicroscopy (List separately in addition to code for primary procedure)</td>
</tr>
<tr>
<td></td>
<td>43206</td>
<td>Esophagoscopy, flexible, transoral; with optical endomicroscopy</td>
</tr>
<tr>
<td></td>
<td>43252</td>
<td>Esophagogastroduodenoscopy, flexible, transoral; with optical endomicroscopy</td>
</tr>
<tr>
<td></td>
<td>88375</td>
<td>Optical endomicroscopic image(s), interpretation and report, real-time or referred, each endoscopic session.</td>
</tr>
</tbody>
</table>

**Date of Origin:** July 2014
Gender Affirming Interventions for Gender Dysphoria

Effective: September 1, 2022

Next Review: September 2022
Last Review: April 2022

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

This policy addresses interventions for gender dysphoria, a marked incongruence between one's experienced/expressed gender and assigned gender.

MEDICAL POLICY CRITERIA

Notes:

- Member contracts for covered services vary. Member contract language takes precedence over medical policy.


- This policy does not address the following interventions:
  - Psychotherapy, which may be considered medically necessary for gender dysphoria; and
  - Medications such as hormonal therapy (see Cross References).
I. Gender affirming interventions for gender dysphoria may be considered medically necessary when either of the following criteria is met:

A. For member contracts subject to Washington’s Gender Affirming Treatment Act (SSB 5313), all of the following criteria are met (1. – 5.):

1. At least 2 licensed mental health professionals have diagnosed gender dysphoria, and recommend the intervention (Note: only 1 mental health professional referral is required for breast/chest surgery); and

2. Six continuous months of hormone therapy as appropriate to the patient’s gender goals unless hormones are not clinically indicated for the individual (Notes: hormonal therapy is not required prior to breast/chest surgery); and

3. At least 6 months of living in a role that is congruent with the patient’s identity; and

4. The request is for treatment(s) as prescribed by the treating provider because of, related to, or consistent with a person's gender expression or identity and is prescribed in accordance with accepted standards of care; and

5. Either of the following is met:
   a. Request is for genital surgery and patient has reached the age of majority (defined as age 18 in Washington state); or
   b. Request is not for genital surgery.

B. For all other member contracts, both of the following criteria are met (1. – 2.):

1. All of following general criteria are met (a. – d.):
   a. Age at least 18 years (Note: age requirement will not be applied to breast/chest surgery with documented provider determination of medical necessity of earlier intervention); and
   b. At least 2 licensed mental health professionals have diagnosed gender dysphoria, and recommend the intervention (Note: only 1 mental health professional referral is required for breast/chest surgery); and
   c. Six continuous months of hormone therapy as appropriate to the patient’s gender goals unless hormones are not clinically indicated for the individual (Note: hormonal therapy is not required prior to breast/chest surgery); and
   d. At least 6 months of living in a role that is congruent with the patient’s identity.

2. One or more of the following criteria are met:
   a. The request is for any of the following procedures:
      i. Clitoroplasty
      ii. Hysterectomy (Note: Hysterectomy is considered medically necessary without routine review and is not required to meet Criterion I.B.1.)
      iii. Labiaplasty

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
iv. Breast/chest surgery (i.e., breast augmentation, breast reduction, mastectomy, mastopexy, nipple/areola reconstruction/repositioning)

v. Metoidioplasty

vi. Orchiectomy

vii. Penectomy

viii. Penile prostheses implantation

ix. Phallic reconstruction/Phalloplasty

x. Salpingo-oophorectomy

xi. Scrotoplasty

xii. Testicular prostheses implantation

xiii. Urethroplasty

xiv. Vaginectomy

xv. Vaginoplasty

b. Clinical documentation is submitted expressly documenting that the intervention would improve otherwise documented significant gender dysphoria and the request is for one or more of the following procedures:

i. Hair removal

ii. Hair transplantation

iii. Endometrial ablation when all of the following criteria are met:

a.) Hysteroscopy, sonohysterography (SIS), or pelvic ultrasound has been performed and report is provided; and

b.) Endometrial sampling or dilation and curettage (D&C) has been performed or is planned according to any of the following:

i.) Endometrial sampling or D&C has been performed and report is provided. The histopathology report is provided showing absence of endometrial hyperplasia or uterine cancer; or

ii.) Endometrial sampling or D&C has been performed and report is provided. The histopathology report is provided, but inadequate tissue was obtained for diagnosis; or

iii.) Cervical stenosis precludes endometrial sampling, and D&C is planned concomitantly with ablation procedure.

II. Gender affirming surgical interventions for gender dysphoria are considered not medically necessary for gender dysphoria when either of the following is met:

A. For member contracts subject to Washington’s Gender Affirming Treatment Act (SSB 5313), when Criterion I.A. is not met; or

B. For member contracts not subject to Washington’s Gender Affirming Treatment Act (SSB 5313), when any of the following is met:
1. Interventions listed in Criterion I.B.2 that do not meet the medical necessity criteria listed in Criterion I.B.1.; or
2. Interventions not listed in Criterion I.B.2. including, but not limited to abdominoplasty, blepharoplasty, brow lift, calf implants, cheek/malar implants, chin/nose implants, collagen injections, face-lift, facial bone reduction, forehead lift, lip reduction, liposuction, neck tightening, panniculectomy, pectoral implants, reduction thyroid chondroplasty, rhinoplasty, suction-assisted lipoplasty of the waist, voice modification surgery, and revision to a previous gender affirming surgery because of dissatisfaction with the appearance.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

LIST OF INFORMATION NEEDED FOR REVIEW

REQUIRED DOCUMENTATION:
The information below must be submitted for review to determine whether policy criteria are met. If any of these items are not submitted, it could impact the review and decision outcome:

- History and Physical/Chart Notes
  - Documentation of therapy requested if applicable
  - Documentation of patient capacity to make decisions/consent to treatment

- For medical treatment or mastectomy:
  - Documentation that a licensed mental health professional has diagnosed gender dysphoria
  - Documentation of length of time living as desired gender
  - Documentation of length of time therapy occurred including licensure of therapist
  - For patients under the age of 18, documented provider determination of medical necessity of earlier intervention

- For all surgical treatments:
  - Documentation that at least 2 licensed mental health professionals have diagnosed gender dysphoria and recommend surgical treatment

- For all surgical treatments, excluding breast/chest surgery:
  - Documentation of hormonal therapy including length of time administered
  - Documented treatment plan including if planned procedures are reversals

- For procedures in Criteria I.B.2.b.:
  - Documentation that the intervention would improve otherwise documented significant gender dysphoria

- In addition to the above, for endometrial ablation:

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Endometrial histopathological report or documentation cervical stenosis precludes endometrial sampling and D&C is planned to be completed concomitantly with ablation procedure.

Hysteroscopy, sonohysterography (SIS), or pelvic ultrasound report

CROSS REFERENCES

1. Endometrial Ablation, Surgery, Policy No. 01
2. Cosmetic and Reconstructive Surgery, Surgery, Policy No. 12
3. Reconstructive Breast Surgery/Mastopexy, and Management of Breast Implants, Surgery, Policy No. 40
4. Reduction Mammaplasty, Surgery, Policy No. 60
5. Autologous Fat Grafting to the Breast and Adipose-derived Stem Cells, Surgery, Policy No. 182
6. Medication Policy Manual, Do a find (Ctrl+F) and enter drug name in the find bar to locate the appropriate policy.

BACKGROUND

This policy supports applicable professional association statements,[1-5] and is also intended to support the Affordable Care Act (ACA) Section 1557 final implementing regulations published on May 18, 2016, and applicable state requirements[6].

MEDICAL AND SURGICAL INTERVENTIONS FOR GENDER DYSPHORIA

A clinical diagnosis of gender dysphoria is required prior to intervention for the disorder. Gender affirming interventions typically include psychotherapy, hormone therapy and in some cases surgical procedures. Psychotherapy followed by hormone therapy is often the first medical treatment sought, although not all transgender individuals on hormone therapy choose to undergo gender affirming surgery.[2]

Gender Dysphoria

Gender dysphoria is defined by the Diagnostic and Statistical Manual of Mental Disorders DSM-5 Diagnostic Criteria as follows:[7]

Gender Dysphoria in Children 302.6 (F64.2)

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months’ duration, as manifested by at least six of the following (one of which must be Criterion 1):

1. A strong desire to be of the other gender or an insistence that one is the other gender (or some alternative gender, different from one's assigned gender).
2. In boys (assigned gender), a strong preference for cross-dressing or simulating female attire; or in girls (assigned gender), a strong preference for wearing only typical masculine clothing and a strong resistance to wearing of typical feminine clothing.
3. A strong preference for cross-gender roles in make-believe play or fantasy play.
4. A strong preference for toys, games, or activities stereotypically used or engaged in by the other gender.
5. A strong preference for playmates of the other gender.
6. In boys (assigned gender), a strong rejection of typically masculine toys, games and activities and a strong avoidance of rough-and-tumble play; or in
girls (assigned gender), a strong rejection of typically feminine toys, games and activities.
7. A strong dislike of one's sexual anatomy.
8. A strong desire for the primary and/or secondary sex characteristics that match one's experienced gender.

B. The condition is associated with clinically significant distress or impairment in social, school, or other important areas of functioning.

*Specify if:*

**With a disorder of sex development** (e.g., a congenital adrenogenital disorder such as 255.2 [E25.0] congenital adrenal hyperplasia or 259.0 [E34.50] androgen insensitivity syndrome).

**Coding note:** Code the disorder of sex development as well as gender dysphoria.

**Gender Dysphoria in Adolescents and Adults 302.85 (F64.1)**

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months’ duration, as manifested by at least two of the following:

1. A marked incongruence between one’s experienced/expressed gender and primary and/or secondary sex characteristics (or in young adolescents, the anticipated secondary sex characteristics).
2. A strong desire to be rid of one’s primary and/or secondary sex characteristics because of a marked incongruence with one’s experienced/expressed gender (or in young adolescents, a desire to prevent the development of the anticipated secondary sex characteristics).
3. A strong desire for the primary and/or secondary sex characteristics of the other gender.
4. A strong desire to be of the other gender (or some alternative gender different from one’s assigned gender).
5. A strong desire to be treated as the other gender (or some alternative gender different from one’s assigned gender).
6. A strong conviction that one has the typical feelings and reactions of the other gender (or some alternative gender different from one’s assigned gender).

B. The condition is associated with clinically significant distress or impairment in social, occupational, or other important areas of functioning.

*Specify if:*

**With a disorder of sex development** (e.g., a congenital adrenogenital disorder such as 255.2 [E25.0] congenital adrenal hyperplasia or 259.0 [E34.50] androgen insensitivity syndrome).

**Coding note:** Code the disorder of sex development as well as gender dysphoria.

*Specify if:*

**Post transition:** The individual has transitioned to full-time living in the desired gender (with or without legalization of gender change) and has undergone (or is preparing to have) at least one cross-sex medical procedure or treatment regimen- namely regular cross-sex hormone treatment or gender reassignment.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
surgery confirming the desired gender (e.g., penectomy, vaginoplasty in the natal male; mastectomy or phalloplasty in the natal female).

Psychotherapy

Psychotherapy provided by a mental health professional typically includes an initial assessment of gender identity and dysphoria, the historical development of gender dysphoric feelings, and severity of resulting stress caused by the condition. The goal of therapy is to assess, diagnose, and discuss treatment options, if needed, and is typically required prior to hormone therapy and/or surgical treatment.

Hormone Therapy

Hormone therapy is undertaken in order to feminize or masculinize individuals' bodies to conform to their desired gender identities. For transgender individuals, hormone replacement therapy (HRT) causes the development of many of the secondary sexual characteristics of their gender identity. Prescribed hormones differ depending upon the natal gender of the individual. For individuals seeking to feminize, hormone treatment may include estradiol, finasteride, and spironolactone. For individuals seeking to masculinize, hormone treatment may include androgenic hormones such as testosterone.

Surgical Interventions

Surgical intervention for gender dysphoria differs depending upon the gender assigned at birth. For individuals who are feminizing, surgery may involve removal of the testicles and penis and the creation of a pseudo vagina, clitoris, and labia. Complications of feminizing genital surgery may include necrosis of the vagina and labia, neovaginal prolapse, fistulas from the bladder or bowel into the vagina, stenosis of the urethra, and small or short vaginas.

For individuals who are masculinizing, surgery may involve removal of the uterus, ovaries, and vagina, and creation of a neophallus and scrotum with scrotal and/or penile prostheses. The creation of a neophallus for these patients is a multistage reconstructive procedure. Currently, techniques for penile reconstruction procedures vary and complications may include frequent urinary tract stenoses and fistulas, donor site scarring and necrosis of the neophallus. In addition, breast size does not significantly decrease with hormonal therapy and as a result, masculinizing patients may choose to undergo mastectomy to remove breast tissue. For many patients this may be the only surgery undertaken. Mastectomy may involve a complete resection of all breast tissue; however, the nipple/areola sparing technique is typically performed to preserve the nipple/areola. For those who are taking androgen hormones, menstruation usually ceases with the medication intervention alone. In those who experience continued uterine bleeding other hormonal regimes may be attempted, or endometrial ablation.

There are various additional surgical procedures which may be sought in order to complete the physical gender transformation and align an individual to their gender identity. However, conflicting opinions exist regarding whether these procedures are essential in treating gender dysphoria.

The WPATH recommends that patients, "engage in 12 continuous months of living in a gender role that is congruent with their gender identity…" prior to gender reassignment surgery so that patients may socially adjust to their desired gender role. WPATH notes that changing a
gender role may have personal and social consequences which should be adequately explored prior to undergoing an irreversible surgery.

### EVIDENCE SUMMARY

Evidence regarding interventions for gender dysphoria in transgender individuals primarily consists of systematic reviews consisting of small cohort studies. Randomized clinical trials (RCTs) comparing gender dysphoria interventions with no intervention are ideal. However, there are challenges in conducting RCTs for these interventions due to several factors, such as small patient populations, and ethical concerns regarding the high morbidity and mortality rates associated with no intervention. Therefore, large RCTs are not anticipated. This policy relies on the following systematic reviews and non-randomized studies, as well as professional association recommendations to support applicable federal and state requirements.

### SYSTEMATIC REVIEWS

Wernick (2019) published a systematic review of the psychological benefits of gender affirming surgery. A total of 33 studies met inclusion criteria. The key concepts searched were quality of life, gender-confirmation surgical procedures, and transgender persons. Sixteen of the identified studies addressed compared pre- and post-surgical data, while 17 studies compared between-group differences. No meta-analysis was completed. Most studies demonstrated a trend of better mental health in transgender individuals who underwent surgeries, but not all reported improvements were statistically significant. The systematic review concluded that gender affirming surgery may lead to psychological benefits for individuals with gender dysphoria and that more research is needed to understand the factors that contribute to the outcomes following these surgeries.

Berli (2017) published a review of the available literature regarding facial gender confirmation surgery (FGCS). The literature search went through December, 2016. The evidence was evaluated using the Oxford Centre for Evidence-Based Medicine suggestions for levels of evidence. Based on their findings, Berli and colleagues recommended that the next World Professional Association for Transgender Health (WPATH) Standards of Care version should include specific FGCS procedures. The authors also recommended replacing the historical term, facial feminization surgery (FFS) with more inclusive terminology – facial gender confirmation surgery. The body of evidence regarding FGCS is limited to case reports and case series. The authors found most data did not include quality-of-life outcome measures, and when reported, standardized instruments were not utilized. FGCS procedures were categorized by the authors as structural (e.g., forehead reconstruction, rhinoplasty), and secondary nonstructural procedures (e.g., blepharoplasty, upper lip shortening techniques). The review was limited by the paucity of data on FGCS as a treatment for gender dysphoria. In addition, methodological limitations of the review included but were not limited to, lack of transparent study selection and a transparent, comprehensive assessment of study quality and risk of bias. These limitations prohibit conclusions about overall health outcomes.

In 2009, Murad assessed quality of life and other psychosocial outcomes of transgendered individuals with gender identity disorder (GID), receiving hormonal therapy as part of gender affirmation surgery. Twenty-eight cohort studies were included in the review which included pooled data from 1,833 patients with GID (1,093 male-to-female [MTF] and 801 female-to-male [FTM]). Significant improvements were reported after gender affirmation compared to pre-treatment status: 80% of patients reported improvement in gender dysphoria (95% CI = 68-89%; 8 studies) 78% reported significant improvement in psychological symptoms (95% CI =
56-94%; 7 studies) 80% reported significant improvement in quality of life (95% CI = 72-88%; 16 studies); and 72% reported significant improvement in sexual function (95% CI = 60-81%; 15 studies). Significant study heterogeneity was reported for all outcomes. Although the authors acknowledge the low quality of evidence used in the analysis, gender affirmation that included hormonal interventions in patients with GID was thought to likely improve symptoms of gender dysphoria and overall quality of life.

In 2009, Elamin evaluated the use of sex steroids on cardiovascular risk in transgender individuals.[13] A total of 16 studies were included in the review with a total of 1,471 male-to-female (MTF) patients and 651 female-to-male (FTM) patients. Steroid use was associated with increased serum triglycerides in both MTF and FTM patients and a nonsignificant effect on HDL-cholesterol and systolic blood pressure in FTM patients. Authors noted that the quality of evidence was low due to methodological limitations of included studies, including but not limited to, heterogeneity of patient population and variable follow-up periods and uncontrolled study design.

Nonrandomized Studies

Primary evidence is limited to cohort studies with a variety of methodological limitations, including but not limited to small sample size, short-term follow-up, lack of comparison group, and varied treatment methods. Despite these limitations, significant improvements in quality of life, psychological comorbidities, and sexual functioning were consistently reported in patients who received gender-confirming medical treatments.[14] Below are summaries of representative publications.

Morrison (2020) performed a prospective, international, multicenter cohort study to assess outcomes of facial feminization surgery.[15] Outcomes reported were facial feminization outcome scores, satisfaction, and cephalometric analysis of femininity. A total of 66 consecutive patients at two clinics were enrolled. The increase in median facial feminization score from pre- to six months post-surgery was statistically significant, from 47.2 to 80.6 (p<0.0001). Cephalometric measures, including glabellar angle, nasolabial angle, and forehead inclination, were significantly more feminine after surgery. Mean satisfaction, measured on a five-point Likert scale, was 3.0 at <1-month post-surgery and six months post-surgery. No general measures of quality of life or mental health were reported.

Imbimbo (2009) evaluated the clinical and psychosocial profile of male-to-female transgenders who had undergone reconstructive surgery.[16] The average age of patients was 31 years old, 72% had high educational levels, half of patients’ contemplated suicide at some point prior to surgery and 4% had attempted suicide. Improved sex life satisfaction was reported in 75% of patients, with almost all patients’ reporting satisfaction with their new sexual status. Additional studies sought to evaluate the sociodemographic profile of transgender individuals with GID in an effort to better characterize and provide treatment for this population.[17]

Heylens (2014) assessed comorbidities and psychosocial factors at various phases of the gender affirmation process in 57 patients with GID.[18] The Symptom Checklist-90 (SCL-90) was administered at three time points: baseline, after the start of hormone therapy, and after sex reassignment surgery (SRS) (also known as [aka] gender affirmation surgery). Psychopathological parameters include overall psychoneurotic distress, anxiety, agoraphobia, depression, somatization, paranoid ideation/psychoticism, interpersonal sensitivity, hostility, and sleeping problems and the psychosocial parameters consist of relationship, living
situation, employment, sexual contacts, social contacts, substance abuse, and suicide attempt. The greatest improvement in psychoneurotic distress was observed after the initiation of hormone therapy (p<0.001). In addition, significant decreases in anxiety, depression, interpersonal sensitivity and hostility were reported after hormone therapy. No significant differences were observed in pre- and postoperative assessments.

Fisher (2013) described clinical and sociodemographic features of 140 transmen (n=48) and transwomen (n=92) with GID and without affirmation surgery.[19] The following assessment tests were administered: the Body Uneasiness Test (a self-rating scale exploring different areas of body-related psychopathology), Symptom Checklist-90 Revised (a self-rating scale to measure psychological state), and the Bem Sex Role Inventory (a self-rating scale to evaluate gender role). Authors reported that transmen displayed significantly better social functioning than transwoman.

Gorin-Lazard (2013) reported a case series which assessed a variety of gender dysphoria symptoms with hormonal treatment preceding gender affirmation surgery. Pre- and post-hormone treatment self-esteem (Social Self-Esteem Inventory), mood (Beck Depression Inventory), QoL (Subjective Quality of Life Analysis), and global functioning (Global Assessment of Functioning) scores were compared in 49 patients.[20] Hormone therapy was reported to be an independent factor in greater self-esteem, a reduction in depression, and improved QoL scores.

Gomez-Gil (2012) evaluated symptoms of social distress, anxiety and depression in 187 transgendered individuals.[21] Of those included in the study, 120 had undergone hormonal sex-reassignment (SR) (aka gender affirmation) treatment and 67 had not. Social anxiety was assessed with the Social Anxiety and Distress Scale (SADS) and depression and anxiety were assessed with the Hospital Anxiety and Depression Scale (HADS). The non-hormone group was reported to be significantly younger than the treatment group (mean age 25.9 vs. 33.6 years, p=0.001) and was less likely to have undergone surgical interventions (p<0.001). After adjusting for confounding factors, the authors reported that patients who were receiving hormone treatment had significantly lower prevalence of depression, anxiety, and social anxiety than those not receiving hormones.

Johansson (2010) reported long-term (five-year) outcomes of transgendered individuals (n=42) with GID who had completely transitioned (n=32), were in progress (n=5) or who were on hormone therapy (n=5).[22] Authors reported that no patient regretted affirmation and clinicians rated the global outcome as favorable in 62% of the cases, compared to 95% according to the patients themselves, with no differences between the subgroups. At follow-up, more than 90% of patients reported stable or improved work situations, partner relations and sex-life. However, 5-15% of patients reported dissatisfaction with hormonal treatment, results of surgery, total gender affirmation procedure, or their present general health.

Asscheman (2011) evaluated the long-term (one-year) effects of cross-sex hormones in 966 male-to-female (MTF) and 365 female-to-male (FTM) transgendered individuals.[23] MTF patients received different doses of estrogen and cyproterone acetate and FTM patients received parenteral/oral testosterone esters or testosterone gel. Hormone treatment levels varied at pre-and post-surgical affirmation time points. High mortality rates were reported in the MTF group when compared to the general population (51%); however, this increased rate was due to non-hormone-related causes such as suicide, acquired immunodeficiency syndrome.
(AIDS), cardiovascular disease, drug abuse and other unknown causes. No significant increase in mortality was observed in FTM patients compared to the general population.

Summary

The evidence is limited by a lack of well-designed studies comparing the safety and effectiveness gender affirming surgery to no treatment or to hormone therapy alone. There are challenges in conducting these large studies, and therefore such studies are not expected in the near future. Although additional research is needed, the research addressing genital and chest surgeries has consistently suggested significant improvement in symptoms and overall quality of life. With regard to other surgeries, such as facial feminization and body contouring, evidence is insufficient to show an improvement in health outcomes.

PRACTICE GUIDELINE SUMMARY

WORLD PROFESSIONAL ASSOCIATION FOR TRANSGENDER HEALTH

The World Professional Association for Transgender Health (WPATH) is a multidisciplinary professional society representing the specialties of medicine, psychology, social sciences and law that has published clinical guidelines regarding health services for patients with gender disorders. In 2011, WPATH approved the update of their evidence and consensus-based guideline regarding, the Standards of Care (SOC) for the Health of Transsexual, Transgender, and Gender Nonconforming Peoples, 7th Version. The 8th version is anticipated in 2021. WPATH guidelines describe surgical procedures as “irreversible changes to the body.” Therefore, WPATH guidelines recommend the appropriate care should be taken to ensure patients have sufficient time (at least 24 hours) to consider all the information and can provide informed consent. WPATH notes, “(t)hese surgeries may be performed once there is written documentation that this assessment has occurred and that the person has met the criteria for a specific surgical treatment. By following this procedure, mental health professionals, surgeons, and patients share responsibility for the decision to make irreversible changes to the body.”

Physical Interventions for Adolescents

WPATH guidelines state that physical interventions for adolescents fall into three categories or stages:

1. Fully reversible interventions. These involve the use of GnRH analogues to suppress estrogen or testosterone production and consequently delay the physical changes of puberty. Alternative treatment options include progestins (most commonly medroxyprogesterone) or other medications (such as spironolactone) that decrease the effects of androgens secreted by the testicles of adolescents who are not receiving GnRH analogues. Continuous oral contraceptives (or depot medroxyprogesterone) may be used to suppress menses.

2. Partially reversible interventions. These include hormone therapy to masculinize or feminize the body. Some hormone-induced changes may need reconstructive surgery to reverse the effect (e.g., gynaecomastia caused by estrogens), while other changes are not reversible (e.g., deepening of the voice caused by testosterone).

3. Irreversible interventions. Reversible and irreversible interventions are outlined in the standards of care, specifying intervention sequencing in adolescents. It is also stated that “[t]wo goals justify intervention with puberty suppressing hormones: (i) their use gives adolescents more time to explore their gender nonconformity and other developmental issues;
and (ii) their use may facilitate transition by preventing the development of sex characteristics that are difficult or impossible to reverse if adolescents continue on to pursue sex reassignment."

Referral for Surgery

WPATH guidelines indicate that surgical interventions can be initiated by a referral from a qualified mental health professional. One or two referrals may be required depending upon the type of surgery requested. “The mental health professional provides documentation—in the chart and/or referral letter—of the patient’s personal and treatment history, progress, and eligibility.” WPATH guidelines specifically recommend the following:

- One referral from a qualified mental health professional is needed for breast/chest surgery (e.g., mastectomy, chest reconstruction, or augmentation mammoplasty).
- Two referrals—from qualified mental health professionals who have independently assessed the patient—are needed for genital surgery (i.e., hysterectomy/salpingo-oophorectomy, orchiectomy, genital reconstructive surgeries).

Criteria for Breast/Chest Surgery (One Referral)

WPATH lists the following criteria for mastectomy and creation of a male chest in FTM patients:

1. Persistent, well-documented gender dysphoria;
2. Capacity to make a fully informed decision and to consent for treatment;
3. Age of majority in a given country;
4. If significant medical or mental health concerns are present, they must be reasonably well controlled.

Hormone therapy is not a prerequisite.

Criteria for Genital Surgery (Two Referrals)

WPATH lists the following criteria for genital surgery:

1. Persistent, well-documented gender dysphoria;
2. Capacity to make a fully informed decision and to consent for treatment;
3. Age of majority in a given country;
4. If significant medical or mental health concerns are present, they must be reasonably well controlled.
5. 12 continuous months of hormone therapy as appropriate to the patient’s gender goals (unless hormones are not clinically indicated for the individual).

In addition, WPATH made specific recommendations regarding breast augmentation procedures:

Breast Augmentation

The WPATH guidelines recommend patients undergo hormone therapy for a minimum of 12 months prior to augmentation surgery and lists specific criteria for breast augmentation (implants/lipofilling).

Other Procedures
The WPATH guidelines state that “while most professionals agree that genital surgery and mastectomy cannot be considered purely cosmetic, opinions diverge as to what degree other surgical procedures (e.g. breast augmentation, facial feminization surgery) can be considered purely reconstructive.” The guidelines go on to say that “[a]lthough it may be much easier to see a phalloplasty or a vaginoplasty as an intervention to end lifelong suffering, for certain patients an intervention like a reduction rhinoplasty can have a radical and permanent effect on their quality of life, and therefore is much more medically necessary than for somebody without gender dysphoria.”

**Physical Effects of Hormone Therapy**

The WPATH guidelines outline the time course of the physical changes that are induced by feminizing/masculinizing hormone therapy. Some of the effects of masculinizing hormones that are relevant to other interventions discussed here are:

- **facial/body hair growth**: expected onset three to six months, expected maximum effect three to five years
- **cessation of menses**: expected onset two to six months
- **vaginal atrophy**: expected onset three to six months, expected maximum effect one to two years
- **deepened voice**: expected onset 3 – 12 months, expected maximum effect one to two years

Some of the effects of feminizing hormones that are relevant to other interventions discussed here are:

- **breast growth**: expected onset three to six months, expected maximum effect two to three years
- **decreased testicular volume**: expected onset three to six months, expected maximum effect two to three years
- **thinning and slowed growth of body and facial hair**: expected onset 6 – 12 months, expected maximum effect over three years

**THE ENDOCRINE SOCIETY**

In 2017, the Endocrine Society in conjunction with American Association of Clinical Endocrinologists, American Society of Andrology, European Society for Pediatric Endocrinology, European Society of Endocrinology, Pediatric Endocrine Society, and World Professional Association for Transgender Health published updated guidelines for the treatment of gender-dysphoric/gender-incongruent persons.[10] The guideline employed transparent methods for evidence review and for rating the quality of evidence. Guidelines were referenced as **recommendations** or **suggestions**, by the numbers 1 and 2, respectively. Evidence was ranked as **very low-quality | ⬤ ⬤ ⬤ ⬤; low quality | ⬤ ⬤ ⬤; moderate quality | ⬤ ⬤ ⬤; and high quality | ⬤ ⬤ ⬤ ⬤.** The consortium made the following statements:

1.0 Evaluation of Youth and Adults

1.1 We advise that only trained mental health professionals (MHPs) who meet the following criteria should diagnose gender dysphoria (GD)/ gender incongruence in adults: (1) competence in using the Diagnostic and Statistical Manual of Mental Disorders (DSM) and/or the International Statistical Classification of Diseases and Related Health
Problems (ICD) for diagnostic purposes, (2) the ability to diagnose GD/ gender incongruence and make a distinction between GD/gender incongruence and conditions that have similar features (e.g., body dysmorphic disorder), (3) training in diagnosing psychiatric conditions, (4) the ability to undertake or refer for appropriate treatment, (5) the ability to psychosocially assess the person’s understanding, mental health, and social conditions that can impact gender-affirming hormone therapy, and (6) a practice of regularly attending relevant professional meetings. (Ungraded Good Practice Statement)

1.2. We advise that only MHPs who meet the following criteria should diagnose GD/gender incongruence in children and adolescents: (1) training in child and adolescent developmental psychology and psychopathology, (2) competence in using the DSM and/or the ICD for diagnostic purposes, (3) the ability to make a distinction between GD/gender incongruence and conditions that have similar features (e.g., body dysmorphic disorder), (4) training in diagnosing psychiatric conditions, (5) the ability to undertake or refer for appropriate treatment, (6) the ability to psychosocially assess the person’s understanding and social conditions that can impact gender-affirming hormone therapy, (7) a practice of regularly attending relevant professional meetings, and (8) knowledge of the criteria for puberty blocking and gender-affirming hormone treatment in adolescents. (Ungraded Good Practice Statement)

1.3. We advise that decisions regarding the social transition of prepubertal youths with GD/gender incongruence are made with the assistance of an MHP or another experienced professional. (Ungraded Good Practice Statement).

1.4. We recommend against puberty blocking and gender-affirming hormone treatment in prepubertal children with GD/gender incongruence. (1 | ✦✦✦✦)

1.5. We recommend that clinicians inform and counsel all individuals seeking gender-affirming medical treatment regarding options for fertility preservation prior to initiating puberty suppression in adolescents and prior to treating with hormonal therapy of the affirmed gender in both adolescents and adults. (1 | ✦✦✦✦)

2.0 Treatment of Adolescents

2.1. We suggest that adolescents who meet diagnostic criteria for GD/gender incongruence, fulfill criteria for treatment, and are requesting treatment should initially undergo treatment to suppress pubertal development. (2 | ✦✦✦✦)

2.2. We suggest that clinicians begin pubertal hormone suppression after girls and boys first exhibit physical changes of puberty. (2 | ✦✦✦✦)

2.3. We recommend that, where indicated, GnRH analogues are used to suppress pubertal hormones. (1 | ✦✦✦✦)

2.4. In adolescents who request sex hormone treatment (given this is a partly irreversible treatment), we recommend initiating treatment using a gradually increasing dose schedule after a multidisciplinary team of medical and MHPs has confirmed the persistence of GD/gender incongruence and sufficient mental capacity to give informed consent, which most adolescents have by age 16 years. (1 | ✦✦✦✦).

2.5. We recognize that there may be compelling reasons to initiate sex hormone treatment prior to the age of 16 years in some adolescents with GD/gender incongruence, even though there are minimal published studies of gender-affirming hormone treatments administered before age 13.5 to 14 years. As with the care of adolescents 16 years of age, we recommend that an expert multidisciplinary team of medical and MHPs manage this treatment. (1 | ✦✦✦✦).
2.6. We suggest monitoring clinical pubertal development every 3 to 6 months and laboratory parameters every 6 to 12 months during sex hormone treatment. (2 | ⬤ ⬤ ○ ○)

3.0 Hormonal Therapy for Transgender Adults

3.1. We recommend that clinicians confirm the diagnostic criteria of GD/gender incongruence and the criteria for the endocrine phase of gender transition before beginning treatment. (1 | ⬤ ⬤ ⬤ ○)

3.2. We recommend that clinicians evaluate and address medical conditions that can be exacerbated by hormone depletion and treatment with sex hormones of the affirmed gender before beginning treatment. (1 | ⬤ ⬤ ⬤ ○)

3.3. We suggest that clinicians measure hormone levels during treatment to ensure that endogenous sex steroids are suppressed and administered sex steroids are maintained in the normal physiologic range for the affirmed gender. (2 | ⬤ ○ ○ ○)

3.4. We suggest that endocrinologists provide education to transgender individuals undergoing treatment about the onset and time course of physical changes induced by sex hormone treatment. (2 | ⬤ ○ ○ ○)

4.0 Adverse Outcome Prevention and Long-term Care

4.1. We suggest regular clinical evaluation for physical changes and potential adverse changes in response to sex steroid hormones and laboratory monitoring of sex steroid hormone levels every 3 months during the first year of hormone therapy for transgender males and females and then once or twice yearly. (2 | ⬤ ⬤ ○ ○)

4.2. We suggest periodically monitoring prolactin levels in transgender females treated with estrogens. (2 | ⬤ ⬤ ○ ○)

4.3. We suggest that clinicians evaluate transgender persons treated with hormones for cardiovascular risk factors using fasting lipid profiles, diabetes screening, and/or other diagnostic tools. (2 | ⬤ ⬤ ○ ○)

4.4. We recommend that clinicians obtain bone mineral density (BMD) measurements when risk factors for osteoporosis exist, specifically in those who stop sex hormone therapy after gonadectomy. (1 | ⬤ ⬤ ○ ○)

4.5. We suggest that transgender females with no known increased risk of breast cancer follow breast-screening guidelines recommended for non-transgender females. (2 | ⬤ ⬤ ○ ○)

4.6. We suggest that transgender females treated with estrogens follow individualized screening according to personal risk for prostatic disease and prostate cancer. (2 | ⬤ ○ ○ ○)

4.7. We advise that clinicians determine the medical necessity of including a total hysterectomy and oophorectomy as part of gender-affirming surgery. (Ungraded Good Practice Statement)

5.0 Surgery for Sex Reassignment and Gender Confirmation

5.1. We recommend that a patient pursue genital gender-affirming surgery only after the MHP and the clinician responsible for endocrine transition therapy both agree that surgery is medically necessary and would benefit the patient’s overall health and/or well-being. (1 | ⬤ ⬤ ○ ○)

5.2. We advise that clinicians approve genital gender-affirming surgery only after completion of at least 1 year of consistent and compliant hormone treatment, unless hormone therapy is not desired or medically contraindicated. (Ungraded Good Practice Statement)
5.3. We advise that the clinician responsible for endocrine treatment and the primary care provider ensure appropriate medical clearance of transgender individuals for genital gender-affirming surgery and collaborate with the surgeon regarding hormone use during and after surgery. (Ungraded Good Practice Statement)

5.4. We recommend that clinicians refer hormone treated transgender individuals for genital surgery when: (1) the individual has had a satisfactory social role change, (2) the individual is satisfied about the hormonal effects, and (3) the individual desires definitive surgical changes. (1 | ⬤ ⬤ ⬤ ⬤)

5.5. We suggest that clinicians delay gender-affirming genital surgery involving gonadectomy and/or hysterectomy until the patient is at least 18 years old or legal age of majority in his or her country. (2 | ⬤ ⬤ ⬤).

5.6. We suggest that clinicians determine the timing of breast surgery for transgender males based upon the physical and mental health status of the individual. There is insufficient evidence to recommend a specific age requirement. (2 | ⬤ ⬤ ⬤ ⬤)

AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGY

In 2017 and 2011, the American College of Obstetricians and Gynecology (ACOG) published committee opinions regarding care for transgender adolescents, and health care services for transgendered individuals, respectively. [24, 25] Although these guidelines are not based on evidence, ACOG does make the following statements:

“Obstetrician–gynecologists should be prepared to assist or refer transgender individuals for routine treatment and screening as well as hormonal and surgical therapies. Hormonal and surgical therapies for transgender patients may be requested, but should be managed in consultation with health care providers with expertise in specialized care and treatment of transgender patients.”

Regarding adolescents, ACOG highlights age-specific concerns with a focus on medical management.

“Consensus guidelines support initiating medical therapy after an adolescent has an established diagnosis of transgender identity and has reached Tanner stage II development.”

In addition, ACOG guidelines made specific recommendations regarding hormone therapy, surgery and screening for both female-to-male and male-to-female patients:

**Female-to-Male Transgender Individuals**

**Hormones**

Methyltestosterone injections every 2 weeks are usually sufficient to suppress menses and induce masculine secondary sex characteristics. Before receiving androgen therapy, patients should be screened for medical contraindications and have periodic laboratory testing, including hemoglobin and hematocrit to evaluate for polycythemia, liver function tests, and serum testosterone level assessments (goal is a mid-normal male range of 500 microgram/dL), while receiving the treatment.

**Surgery**

Hysterectomy, with or without salpingo-oophorectomy, is commonly part of the surgical process. An obstetrician–gynecologist who has no specialized expertise in transgender
care may be asked to perform this surgery, and also may be consulted for routine reasons such as dysfunctional bleeding or pelvic pain. Reconstructive surgery should be performed by a urologist, gynecologist, plastic surgeon, or general surgeon who has specialized competence and training in this field.

**Screening**

Age-appropriate screening for breast cancer and cervical cancer should be continued unless mastectomy or removal of the cervix has occurred. For patients using androgen therapy who have not had a complete hysterectomy, there may be an increased risk of endometrial cancer and ovarian cancer.

**Male-to-Female Transgender Individuals**

**Hormones**

Estrogen therapy results in gynecomastia, reduced hair growth, redistribution of fat, and reduced testicular volume. All patients considering therapy should be screened for medical contraindications. After surgery, doses of estradiol, 2–4 mg/d, or conjugated equine estrogen, 2.5 mg/d, are often sufficient to keep total testosterone levels to normal female levels of less than 25 ng/dL. Nonoral therapy also can be offered. It is recommended that male-to-female transgender patients receiving estrogen therapy have an annual prolactin level assessment and visual field examination to screen for prolactinoma.

**Surgery**

Surgery usually involves penile and testicular excision and the creation of a neovagina. Reported complications of surgery include vaginal and urethral stenosis, fistula formation, problems with remnants of erectile tissue, and pain. Vaginal dilation of the neovagina is required to maintain patency. Other surgical procedures that may be performed include breast implants and nongenital surgery, such as facial feminization surgery.

**Screening**

Age-appropriate screening for breast and prostate cancer is appropriate for male-to-female transgender patients. Opinion varies regarding the need for Pap testing in this population. In patients who have a neocervix created from the glans penis, routine cytologic examination of the neocervix may be indicated. The glands are more prone to cancerous changes than the skin of the penile shaft, and intraepithelial neoplasia of the glans is more likely to progress to invasive carcinoma than is intraepithelial neoplasia of other penile skin.

**SUMMARY**

For member contracts subject to Washington’s Gender Affirming Treatment Act (SSB 5313)

For member contracts subject to the Washington Gender Affirming Treatment Act (SSB 5313), criteria for gender affirming interventions are based on the research, guidelines developed using the available evidence and expert clinical consensus, and on the Act.
Therefore, for member contracts subject to the Washington Gender Affirming Treatment Act (SSB 5313), gender affirming interventions for gender dysphoria may be considered medically necessary when specified policy criteria are met.

For member contracts subject to the Washington Gender Affirming Treatment Act (SSB 5313), criteria for gender affirming interventions are based on the Act and on guidelines developed using the available evidence and expert clinical consensus. Therefore, for these members, when these criteria are not met, gender affirming interventions for gender dysphoria are considered not medically necessary.

For member contracts not subject to Washington’s Gender Affirming Treatment Act (SSB 5313)

The research lacks well-designed studies comparing the safety and effectiveness of no intervention for gender dysphoria with interventions such as gender affirming surgery. However, there are challenges in conducting large studies to evaluate existing treatments, and such studies are not expected in the near future. Although additional research is needed, the research has consistently suggested significant improvement in symptoms and overall quality of life in those who have received certain interventions for gender dysphoria. The World Professional Association for Transgender Health (WPATH) Standards of Care (SOC) for the Health of Transsexual, Transgender, and Gender Nonconforming Peoples recommend that specific criteria are met prior to surgical interventions for gender dysphoria. These guidelines are based on evidence and expert clinical consensus and the included criteria were developed to promote optimal patient care. Therefore, gender affirming interventions for gender dysphoria may be considered medically necessary when specified policy criteria are met.

The World Professional Association for Transgender Health (WPATH) Standards of Care (SOC) for the Health of Transsexual, Transgender, and Gender Nonconforming Peoples recommend that specific criteria are met prior to surgical interventions for gender dysphoria. These guidelines are based on evidence and expert clinical consensus and the included criteria were developed to promote optimal patient care. Therefore, when these criteria are not met, gender affirming interventions for gender dysphoria are considered not medically necessary.

There are no evidence-based clinical practice guidelines that recommend gender affirming surgical interventions not listed in Criterion I.B.2. or revision to a previous gender affirming surgery because of dissatisfaction with the appearance improve health outcomes. Therefore, gender affirming surgical interventions not listed in Criterion I.B.2. and revision to a previous gender affirming surgery because of dissatisfaction with the appearance are considered not medically necessary.

The World Professional Association for Transgender Health (WPATH) Standards of Care (SOC) for the Health of Transsexual, Transgender, and Gender Nonconforming Peoples describe reversible and irreversible interventions, and the ideal order and timing of these approaches. Surgery as an intervention is considered irreversible by WPATH. Therefore, reversal of gender affirming surgery for gender dysphoria is considered not medically necessary.

REFERENCES

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


**CODES**

**NOTES:**

- Follicular unit extraction (FEU) of individual hairs is correctly coded with code 15775 or 15776 and is determined by the number of "punch grafts" performed. Be advised that standard CMS Medically Unlikely Edits (MUEs or Maximum Units of Service) will apply.
- Code 17999 should be reported for laser hair removal. This code may also be used for abdominoplasty or calf/pectoral implants.
- Codes 31552, 31554, 31580, 31584, 31587, and 31591 are not appropriate to use to represent voice modification. Unlisted code 31599 should be reported instead.
- Code 31899 should be reported for reduction thyroid chondroplasty (reduction of the thyroid cartilage or Adam's Apple).
- Code 40799 should be reported for lip reduction.
- Code 55899 should be reported for phallic reconstruction/phalloplasty.
- Codes 55970 and 55980 are non-specific. The specific procedure code(s) must be requested in place of these non-specific codes.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>11950</td>
<td>Subcutaneous injection of filling material (eg, collagen); 1 cc or less</td>
</tr>
<tr>
<td></td>
<td>11951</td>
<td>Subcutaneous injection of filling material (eg, collagen); 1.1 to 5.0 cc</td>
</tr>
<tr>
<td></td>
<td>11952</td>
<td>Subcutaneous injection of filling material (eg, collagen); 5.1 to 10.0 cc</td>
</tr>
<tr>
<td></td>
<td>11954</td>
<td>Subcutaneous injection of filling material (eg, collagen); over 10 cc</td>
</tr>
<tr>
<td></td>
<td>11970</td>
<td>Replacement of tissue expander with permanent implant</td>
</tr>
<tr>
<td></td>
<td>11971</td>
<td>Removal of tissue expander(s) without insertion of implant</td>
</tr>
<tr>
<td></td>
<td>15775</td>
<td>Punch graft for hair transplant; 1 to 15 punch grafts</td>
</tr>
<tr>
<td></td>
<td>15776</td>
<td>Punch graft for hair transplant; more than 15 punch grafts</td>
</tr>
<tr>
<td></td>
<td>15820</td>
<td>Blepharoplasty, lower eyelid</td>
</tr>
<tr>
<td></td>
<td>15821</td>
<td>;with extensive herniated fat pad</td>
</tr>
<tr>
<td></td>
<td>15822</td>
<td>Blepharoplasty, upper eyelid</td>
</tr>
<tr>
<td></td>
<td>15823</td>
<td>;with excessive skin weighting down lid</td>
</tr>
<tr>
<td></td>
<td>15824</td>
<td>Rhytidectomy; forehead</td>
</tr>
<tr>
<td></td>
<td>15825</td>
<td>Rhytidectomy; neck with platysmal tightening (platysmal flap, P-flap)</td>
</tr>
<tr>
<td></td>
<td>15826</td>
<td>Rhytidectomy; glabellar frown lines</td>
</tr>
<tr>
<td></td>
<td>15828</td>
<td>Rhytidectomy; cheek, chin, and neck</td>
</tr>
<tr>
<td></td>
<td>15829</td>
<td>Rhytidectomy; superficial musculoaponeurotic system (SMAS) flap</td>
</tr>
<tr>
<td></td>
<td>15830</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); abdomen, infraumbilical panniculectomy</td>
</tr>
<tr>
<td></td>
<td>15832</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); thigh</td>
</tr>
<tr>
<td></td>
<td>15833</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); leg</td>
</tr>
<tr>
<td></td>
<td>15834</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); hip</td>
</tr>
<tr>
<td></td>
<td>15835</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); buttock</td>
</tr>
<tr>
<td></td>
<td>15836</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); arm</td>
</tr>
<tr>
<td></td>
<td>15837</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); forearm or hand</td>
</tr>
<tr>
<td></td>
<td>15838</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); submental fat pad</td>
</tr>
<tr>
<td></td>
<td>15839</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy); other area</td>
</tr>
<tr>
<td></td>
<td>15847</td>
<td>Excision, excessive skin and subcutaneous tissue (includes lipectomy), abdomen (eg, abdominoplasty) (includes umbilical transposition and fascial plication) (List separately in addition to code for primary procedure)</td>
</tr>
<tr>
<td></td>
<td>15876</td>
<td>Suction assisted lipectomy; head and neck</td>
</tr>
<tr>
<td></td>
<td>15877</td>
<td>Suction assisted lipectomy; trunk</td>
</tr>
<tr>
<td></td>
<td>15878</td>
<td>Suction assisted lipectomy; upper extremity</td>
</tr>
<tr>
<td></td>
<td>15879</td>
<td>Suction assisted lipectomy; lower extremity</td>
</tr>
<tr>
<td></td>
<td>17380</td>
<td>Electrolysis epilation, each 30 minutes</td>
</tr>
<tr>
<td></td>
<td>17999</td>
<td>Unlisted procedure, skin, mucous membrane and subcutaneous tissue</td>
</tr>
<tr>
<td></td>
<td>19303</td>
<td>Mastectomy, simple, complete</td>
</tr>
<tr>
<td></td>
<td>19316</td>
<td>Mastopexy</td>
</tr>
<tr>
<td></td>
<td>19318</td>
<td>Breast reduction</td>
</tr>
<tr>
<td></td>
<td>19325</td>
<td>Breast augmentation with implant</td>
</tr>
<tr>
<td></td>
<td>19350</td>
<td>Nipple/areola reconstruction</td>
</tr>
<tr>
<td></td>
<td>19499</td>
<td>Unlisted procedure, breast</td>
</tr>
<tr>
<td></td>
<td>21120</td>
<td>Genioplasty; augmentation (autograft, allograft, prosthetic material)</td>
</tr>
<tr>
<td></td>
<td>21121</td>
<td>Genioplasty; sliding osteotomy, single piece</td>
</tr>
<tr>
<td></td>
<td>21122</td>
<td>Genioplasty; sliding osteotomies, 2 or more osteotomies (eg, wedge excision or bone wedge reversal for asymmetrical chin)</td>
</tr>
</tbody>
</table>

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
<table>
<thead>
<tr>
<th>Procedure Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>21123</td>
<td>Genioplasty; sliding, augmentation with interpositional bone grafts (includes obtaining autografts)</td>
</tr>
<tr>
<td>21209</td>
<td>Osteoplasty, facial bones; reduction</td>
</tr>
<tr>
<td>21270</td>
<td>Malar augmentation, prosthetic material</td>
</tr>
<tr>
<td>30400</td>
<td>Rhinoplasty, primary; lateral and alar cartilages and/or elevation of nasal tip</td>
</tr>
<tr>
<td>30410</td>
<td>; complete, external parts including bony pyramid, lateral and alar cartilages, and/or elevation of nasal tip</td>
</tr>
<tr>
<td>30420</td>
<td>; including major sepal repair</td>
</tr>
<tr>
<td>30430</td>
<td>Rhinoplasty, secondary; minor revision (small amount of nasal tip work)</td>
</tr>
<tr>
<td>30435</td>
<td>; intermediate revision (bony work with osteotomies)</td>
</tr>
<tr>
<td>30450</td>
<td>; major revision (nasal tip work and osteotomies)</td>
</tr>
<tr>
<td>31599</td>
<td>Unlisted procedure, larynx</td>
</tr>
<tr>
<td>31899</td>
<td>Unlisted procedure, trachea, bronchi</td>
</tr>
<tr>
<td>40799</td>
<td>Unlisted procedure, lips</td>
</tr>
<tr>
<td>53400</td>
<td>Urethroplasty; first stage, for fistula, diverticulum, or stricture (eg, Johannsen type)</td>
</tr>
<tr>
<td>53405</td>
<td>Urethroplasty; second stage (formation of urethra), including urinary diversion</td>
</tr>
<tr>
<td>53410</td>
<td>Urethroplasty, 1-stage reconstruction of male anterior urethra</td>
</tr>
<tr>
<td>53415</td>
<td>Urethroplasty, transpubic or perineal, 1-stage, for reconstruction or repair of prostatic or membranous urethra</td>
</tr>
<tr>
<td>53420</td>
<td>Urethroplasty, 2-stage reconstruction or repair of prostatic or membranous urethra; first stage</td>
</tr>
<tr>
<td>53425</td>
<td>Urethroplasty, 2-stage reconstruction or repair of prostatic or membranous urethra; second stage</td>
</tr>
<tr>
<td>53430</td>
<td>Urethroplasty, reconstruction of female urethra</td>
</tr>
<tr>
<td>54125</td>
<td>Amputation of penis; complete (Penectomy)</td>
</tr>
<tr>
<td>54400</td>
<td>Insertion of penile prosthesis; non-inflatable (semi-rigid)</td>
</tr>
<tr>
<td>54401</td>
<td>Insertion of penile prosthesis; inflatable (self-contained)</td>
</tr>
<tr>
<td>54405</td>
<td>Insertion of multi-component, inflatable penile prosthesis, including placement of pump, cylinders, and reservoir</td>
</tr>
<tr>
<td>54520</td>
<td>Orchietomy, simple (including subcapsular), with or without testicular prosthesis, scrotal or inguinal approach</td>
</tr>
<tr>
<td>54660</td>
<td>Insertion of testicular prosthesis</td>
</tr>
<tr>
<td>54690</td>
<td>Laparoscopy, surgical; orchietomy</td>
</tr>
<tr>
<td>55175</td>
<td>Scrotoplasty; simple</td>
</tr>
<tr>
<td>55180</td>
<td>Scrotoplasty; complicated</td>
</tr>
<tr>
<td>55899</td>
<td>Phallic reconstruction/Phalloplasty (Unlisted procedure, male genital system)</td>
</tr>
<tr>
<td>55970</td>
<td>intersex surgery; male to female</td>
</tr>
<tr>
<td>55980</td>
<td>intersex surgery; female to male</td>
</tr>
<tr>
<td>56625</td>
<td>Vulvectomy simple; complete</td>
</tr>
<tr>
<td>56800</td>
<td>Plastic repair of introitus</td>
</tr>
<tr>
<td>56805</td>
<td>Clitoroplasty for intersex state</td>
</tr>
<tr>
<td>57106</td>
<td>Vaginectomy, partial removal of vaginal wall</td>
</tr>
<tr>
<td>57110</td>
<td>Vaginectomy, complete removal of vaginal wall;</td>
</tr>
<tr>
<td>57291</td>
<td>Construction of artificial vagina; without graft</td>
</tr>
<tr>
<td>57292</td>
<td>Construction of artificial vagina; with graft</td>
</tr>
<tr>
<td>57295</td>
<td>Revision (including removal) of prosthetic vaginal graft; vaginal approach</td>
</tr>
<tr>
<td>57296</td>
<td>Revision (including removal) of prosthetic vaginal graft; open abdominal approach</td>
</tr>
<tr>
<td>57335</td>
<td>Vaginoplasty for intersex state</td>
</tr>
<tr>
<td>57426</td>
<td>Revision (including removal) of prosthetic vaginal graft, laparoscopic approach</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>58150</td>
<td>Total abdominal hysterectomy (corpus and cervix), with or without removal of</td>
</tr>
<tr>
<td></td>
<td>tube(s), with or without removal of ovary(s)</td>
</tr>
<tr>
<td>58180</td>
<td>Supracervical abdominal hysterectomy (subtotal hysterectomy), with or</td>
</tr>
<tr>
<td></td>
<td>without removal of tube(s), with or without removal of ovary(s)</td>
</tr>
<tr>
<td>58260</td>
<td>Vaginal hysterectomy, for uterus 250 g or less</td>
</tr>
<tr>
<td>58262</td>
<td>Vaginal hysterectomy, for uterus 250 g or less; with removal of tube(s),</td>
</tr>
<tr>
<td></td>
<td>and/or ovary(s)</td>
</tr>
<tr>
<td>58270</td>
<td>Vaginal hysterectomy, for uterus 250 g or less; with repair of enterocele</td>
</tr>
<tr>
<td>58275</td>
<td>Vaginal hysterectomy, with total or partial vaginectomy;</td>
</tr>
<tr>
<td>58290</td>
<td>Vaginal hysterectomy, for uterus greater than 250 g</td>
</tr>
<tr>
<td>58291</td>
<td>Vaginal hysterectomy, for uterus greater than 250 g; with removal of tube(s)</td>
</tr>
<tr>
<td></td>
<td>and/or ovary(s)</td>
</tr>
<tr>
<td>58353</td>
<td>Endometrial ablation, thermal, without hysteroscopic guidance</td>
</tr>
<tr>
<td>58356</td>
<td>Endometrial cryoablation with ultrasonic guidance, including endometrial</td>
</tr>
<tr>
<td></td>
<td>curettage, when performed</td>
</tr>
<tr>
<td>58541</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus 250 g or</td>
</tr>
<tr>
<td></td>
<td>less</td>
</tr>
<tr>
<td>58542</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus 250 g or</td>
</tr>
<tr>
<td></td>
<td>less; with removal of tube(s) and/or ovary(s)</td>
</tr>
<tr>
<td>58543</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus greater than</td>
</tr>
<tr>
<td></td>
<td>250 g</td>
</tr>
<tr>
<td>58544</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus greater than</td>
</tr>
<tr>
<td></td>
<td>250 g; with removal of tube(s) and/or ovary(s)</td>
</tr>
<tr>
<td>58550</td>
<td>Laparoscopy, surgical, with vaginal hysterectomy, for uterus 250 g or</td>
</tr>
<tr>
<td></td>
<td>less</td>
</tr>
<tr>
<td>58552</td>
<td>Laparoscopy, surgical, with vaginal hysterectomy, for uterus 250 g or</td>
</tr>
<tr>
<td></td>
<td>less; with removal of tube(s) and/or ovary(s)</td>
</tr>
<tr>
<td>58553</td>
<td>Laparoscopy, surgical, with vaginal hysterectomy, for uterus greater than</td>
</tr>
<tr>
<td></td>
<td>250 g</td>
</tr>
<tr>
<td>58554</td>
<td>Laparoscopy, surgical, with vaginal hysterectomy, for uterus greater than</td>
</tr>
<tr>
<td></td>
<td>250 g; with removal of tube(s) and/or ovary(s)</td>
</tr>
<tr>
<td>58563</td>
<td>Hysteroscopy, surgical; with endometrial ablation (eg. Endometrial ressection,</td>
</tr>
<tr>
<td></td>
<td>electrosurgical ablation, thermoablation)</td>
</tr>
<tr>
<td>58570</td>
<td>Laparoscopy, surgical, with total hysterectomy, for uterus 250 g or less</td>
</tr>
<tr>
<td>58571</td>
<td>Laparoscopy, surgical, with total hysterectomy, for uterus 250 g or less;</td>
</tr>
<tr>
<td></td>
<td>with removal of tube(s) and/or ovary(s)</td>
</tr>
<tr>
<td>58572</td>
<td>Laparoscopy, surgical, with total hysterectomy, for uterus greater than 250</td>
</tr>
<tr>
<td></td>
<td>g</td>
</tr>
<tr>
<td>58573</td>
<td>Laparoscopy, surgical, with total hysterectomy, for uterus greater than 250</td>
</tr>
<tr>
<td></td>
<td>g; with removal of tube(s) and/or ovary(s)</td>
</tr>
<tr>
<td>58720</td>
<td>Salpingo-oophorectomy, complete or partial, unilateral or bilateral (separate</td>
</tr>
<tr>
<td></td>
<td>procedure)</td>
</tr>
<tr>
<td>67900</td>
<td>Repair of brow ptosis (supraciliary, mid-forehead or coronal approach)</td>
</tr>
<tr>
<td>67901</td>
<td>Repair of blepharoptosis; frontalis muscle technique with suture or other</td>
</tr>
<tr>
<td></td>
<td>material (eg, banked fascia)</td>
</tr>
<tr>
<td>67902</td>
<td>Repair of blepharoptosis; frontalis muscle technique with autologous fascial</td>
</tr>
<tr>
<td></td>
<td>sling (includes obtaining fascia)</td>
</tr>
<tr>
<td>67903</td>
<td>Repair of blepharoptosis; (tarso) levator resection or advancement, internal</td>
</tr>
<tr>
<td></td>
<td>approach</td>
</tr>
<tr>
<td>67904</td>
<td>Repair of blepharoptosis; (tarso) levator resection or advancement, external</td>
</tr>
<tr>
<td></td>
<td>approach</td>
</tr>
<tr>
<td>67906</td>
<td>Repair of blepharoptosis; superior rectus technique with fascial sling (includes</td>
</tr>
<tr>
<td></td>
<td>obtaining fascia)</td>
</tr>
<tr>
<td>67908</td>
<td>Repair of blepharoptosis; conjunctivo-tarso-Muller's muscle-levator resection</td>
</tr>
<tr>
<td></td>
<td>(eg, Fasanella-Servat type)</td>
</tr>
<tr>
<td>67909</td>
<td>Reduction of overcorrection of ptosis</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>67950</td>
<td>Prosthesis, penile, inflatable</td>
</tr>
<tr>
<td>C1813</td>
<td>Prosthesis, penile, noninflatable</td>
</tr>
<tr>
<td>L8039</td>
<td>Breast prosthesis, not otherwise specified</td>
</tr>
<tr>
<td>L8600</td>
<td>Implantable breast prosthesis, silicone or equal</td>
</tr>
</tbody>
</table>

*Date of Origin: September 2014*
IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Bioengineered skin and soft tissue substitutes may be derived from human tissue (autologous or allogeneic), nonhuman tissue, synthetic materials, or a composite of these materials. Amniotic products may be derived from amnion, chorion, amniotic fluid, and umbilical cord. There are many potential applications for these products, including breast reconstruction, chronic full-thickness diabetic lower-extremity ulcers, venous ulcers, severe burns, knee osteoarthritis, plantar fasciitis, and ophthalmic conditions.

MEDICAL POLICY CRITERIA

Notes:

- Product-specific HCPCS codes are listed below in brackets, where applicable. Skin substitutes without a specific code may use Q4100/A4100.
- This policy does not apply to dural substitutes used during surgical procedures involving the central nervous system (brain and spinal cord) or to unprocessed cadaver skin allografts used as wound dressing.
I. **Breast reconstructive surgery** using any of the following allogeneic acellular dermal matrix products may be considered *medically necessary*:

A. AlloDerm® [Q4116]
B. AlloMend®
C. Cortiva® (AlloMax™)
D. DermACELL® [Q4122]
E. DermaMatrix™
F. FlexHD® [Q4128]
G. FlexHD® Pliable™
H. GraftJacket® [Q4107]

II. **Treatment of non-healing diabetic lower-extremity ulcers** that have not adequately responded following a 1-month period of conventional ulcer therapy, using any of the following tissue-engineered or amniotic skin substitutes, may be considered *medically necessary*:

A. Affinity® [Q4159]
B. AlloPatch® [Q4128]
C. AmnioBand® Membrane [Q4151]
D. Apligraf® [Q4101]
E. Biovance® [Q4154]
F. Dermagraft® [Q4106]
G. EpiCord® [Q4187]
H. EpiFix® [Q4186]
I. Grafix® [Q4132, Q4133]
J. Integra® Omnigraft™ Dermal Regeneration Matrix (also known as Omnigraft™) [Q4105]
K. Integra® Flowable Wound Matrix [Q4114]

III. **Treatment of chronic, noninfected, lower-extremity skin ulcers due to venous insufficiency** that have not adequately responded following a 1-month period of conventional ulcer therapy, using any of the following tissue-engineered skin substitutes, may be considered *medically necessary*:

A. Apligraf® [Q4101]
B. Oasis®™ Wound Matrix [Q4102]

IV. **Treatment of dystrophic epidermolysis bullosa** using the following tissue-engineered skin substitutes may be considered *medically necessary*:

A. OrCel® (for the treatment of mitten-hand deformity when standard wound therapy has failed and when provided in accordance with the humanitarian device exemption [HDE] specifications of the U.S. Food and Drug Administration [FDA]).

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
V. Treatment of second- and third-degree burns using any of the following tissue-engineered skin substitutes may be considered medically necessary:

A. Epicel® (for the treatment of deep dermal or full-thickness burns comprising a total body surface area ≥30% when provided in accordance with the HDE specifications of the FDA)

B. Integra® Dermal Regeneration Template [Q4105]

VI. Human amniotic membrane grafts not listed as investigational (see Policy Guidelines) may be considered medically necessary as a component of ophthalmologic surgery or repair, including but not limited to Prokera®, AmbioDisk™, or AmnioGraft®.

VII. Treatment of lower-extremity ulcers due to diabetes or venous insufficiency is considered not medically necessary when there has not been at least 1 month of conventional ulcer therapy.

VIII. The use of bioengineered skin and soft tissue substitutes for hernia repair or parastomal reinforcement is considered not medically necessary.

IX. All other uses of the amniotic membrane grafts and bioengineered skin and soft tissue substitutes listed above are considered investigational, including but not limited to tendon repair.

X. All other amniotic products and bioengineered skin or soft tissue substitutes not listed above are considered investigational (see Policy Guidelines).

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

POLICY GUIDELINES

Amniotic fluid is considered an amniotic product.

INVESTIGATIONAL PRODUCTS

The following amniotic products, placental products, and skin and soft tissue substitutes are considered investigational. There are many products available, and this list is not all-inclusive.

- ACell® UBM Hydrated/Lyophilized Wound Dressing
- AlloGen® [Q4212]
- AlloSkin™ [Q4115]
- AlloSkin™ AC [Q4141]
- AlloSkin™ RT [Q4123]
- AlloWrap® [Q4150]
- Altiply™ [Q4235]
- AmnioAmp-MP™ [Q4250]
- Amnioarmor™ [Q4188]
- AmnioBand®, particulate [Q4168]
- AmnioBind™ [Q4225]
- AmnioCore™ [Q4227]
- AmnioCyte™ Plus [Q4242]
- AmnioExcel® [Q4137]
- AmnioMatrix® [Q4139]
- Amnio-Maxx™ [Q4239]
- Amnion Bio/AxoBioMembrane™ [Q4211]
- Amniorepair® [Q4235]
- Amniotext™ [Q4245]
- Amniotext™ patch [Q4247]
- Amnio Wound [Q4181]
- AmnioWrap2™ [Q4221]
- Amniply™ [Q4249]
- Aongen™ Collagen Matrix
- APIS® [A2010]
- Architect® ECM, PX, FX [Q4147]
• Artacent® Cord [Q4216]
• Artacent® Wound [Q4169]
• Artacent® ac [Q4189, Q4190]
• ArthroFlex™ (Flex Graft) [Q4125]
• Ascent™ [Q4213]
• AxoGuard® Nerve Protector (AxoGen)
• Axolotl Ambient™, Cryo™ [Q4215]
• Axolotl Graft™, DualGraft™ [Q4210]
• BellaCell HD [Q4220]
• Biobrane®/Biobrane-L
• Bio-ConneKt® Wound Matrix [Q4161]
• BioDFence® [Q4140]
• BioDFence® Dryflex [Q4138]
• Biowound™, Plus, Xplus [Q4217]
• Cellesta™/Cellesta™ Duo [Q4184]
• Cellesta™ Cord [Q4214]
• Cellesta™ flowable amnion [Q4185]
• CLARIX 100 [Q4156]
• CLARIX Flo [Q4155]
• Cogenex® amniotic membrane [Q4229]
• Cogenex® flowable amnion [Q4230]
• CollaCare®
• CollaCare® Dental
• Collagen Wound Dressing (Oasis Research)
• CollaGUARD®
• CollaMend™
• CollaWound™
• Coll-e-Derm™ [Q4193]
• Collexa®
• Colliea®
• Conexa™
• CoreCyte™ [Q4240]
• Coreleader Colla-Pad
• CorMatrix®
• Corplex™ [Q4231, Q4232]
• CoreText™ or ProText™ [Q4246]
• Cryo-Cord™ [Q4237]
• Cygnus™ [Q4170]
• Cygnus™ Matrix [Q4199]
• Cymetra™ [Q4112]
• Cytal® (previously MatriStem®) [Q4118, Q4166]
• Dermadapt™ Wound Dressing
• Dermacyte® [Q4248]
• Derma-Gide® [Q4203]
• DermaPure™ [Q4152]
• DermaSpan™ [Q4126]
• Dermavest® [Q4153]
• Derm-Maxx [Q4238]
• DressSkin
• Endoform Dermal Template™
• ENDURAGen™
• Enverse™ [Q4258]
• EpiFix® Injectable [Q4145]
• Excellagen [Q4149]
• ExpressGraft™
• E-Z Derm™ [Q4136]
• FlowerAmnioFlo™ [Q4177]
• Flower AmnioPatch™ [Q4178]
• FlowerDerm™ [Q4179]
• Fluid Flow™, Fluid GF™ [Q4206]
• GammaGraft [Q4111]
• Graftjacket® Xpress, injectable [Q4113]
• Helicoll™ [Q4164]
• Human Health Factor 10 Patch™ (HHF10P™) [Q4224]
• Hyalomatrix® [Q4117]
• Hyalomatrix® PA
• hMatrix® [Q4134]
• InnovaMatrix™ [A2001]
• InnovaMatrix™ FS [A2013]
• Integra™ Matrix Wound Dressing [Q4108]
• Interfyl® [Q4171]
• Keramatrix® [Q4165]
• Kerecis™ [Q4158]
• Keroxx® [Q4202]
• MariGen™/Kerecis™ Omega3™
• MatriDerm®
• Matrion™ [Q4201]
• Matrix HD™
• Mediskin® [Q4135]
• Membrane Graft™/Membrane Wrap™ [Q4205]
• MemoDerm™ [Q4126]
• Microlyte® Matrix [A2005]
• Mirolyte® Matrix [A2005]
• Miroderm® biologic wound matrix [Q4175]
• Mirragen® [A2002]
LIST OF INFORMATION NEEDED FOR REVIEW

REQUIRED DOCUMENTATION:

The information below must be submitted for review to determine whether policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.
• History and physical/chart notes
• Indication for the requested service
• Documentation of symptoms, associated diagnoses and treatments
• Conservative treatment provided, if any
• Name of product to be used and indication

CROSS REFERENCES
None

BACKGROUND

BIOENGINEERED SKIN AND SOFT TISSUE SUBSTITUTES

Bioengineered skin and soft tissue substitutes may be either acellular or cellular. Acellular products (e.g., dermis with cellular material removed, synthetic products) contain a matrix or scaffold composed of materials such as collagen, hyaluronic acid, and fibronectin. Acellular dermal matrix (ADM) products can differ in a number of ways, including as species source (human, bovine, porcine), tissue source (e.g., dermis, pericardium, intestinal mucosa), additives (e.g., antibiotics, surfactants), hydration (wet, freeze-dried), and required preparation (multiple rinses, rehydration).

Cellular products contain living cells such as fibroblasts and keratinocytes within a matrix. The cells contained within the matrix may be autologous, allogeneic, or derived from other species (e.g., bovine, porcine). Skin substitutes may also be composed of dermal cells, epidermal cells, or a combination of dermal and epidermal cells, and may provide growth factors to stimulate healing. Bioengineered skin substitutes can be used as either temporary or permanent wound coverings.

There are many potential applications for artificial skin and soft tissue products. One large category is nonhealing wounds, which potentially encompasses diabetic neuropathic ulcers, vascular insufficiency ulcers, and pressure ulcers. A substantial minority of such wounds do not heal adequately with standard wound care, leading to prolonged morbidity and increased risk of mortality. For example, nonhealing lower-extremity wounds represent an ongoing risk for infection, sepsis, limb amputation, and death. Bioengineered skin and soft tissue substitutes have the potential to improve rates of healing and reduce secondary complications.

Other situations in which bioengineered skin products might substitute for living skin grafts include certain postsurgical states (e.g., breast reconstruction) in which skin coverage is inadequate for the procedure performed, or for surgical wounds in patients with compromised ability to heal. Second- and third-degree burns are another indication in which artificial skin products may substitute for auto- or allografts. Certain primary dermatologic conditions that involve large areas of skin breakdown (e.g., bullous diseases) may also be conditions in which artificial skin products can be considered as substitutes for skin grafts. ADM products are also being evaluated in the repair of other soft tissues including rotator cuff repair, following oral and facial surgery, hernias, and other conditions.

AMNIOTIC PRODUCTS

Human Amniotic Membrane
Human amniotic membrane (HAM) consists of two conjoined layers, the amnion, and chorion, and forms the innermost lining of the amniotic sac or placenta. When prepared for use as an allograft, the membrane is harvested immediately after birth, cleaned, sterilized, and either cryopreserved or dehydrated. Many products available using amnion, chorion, amniotic fluid, and umbilical cord are being studied for the treatment of a variety of conditions, including chronic full-thickness diabetic lower-extremity ulcers, venous ulcers, knee osteoarthritis, plantar fasciitis, and ophthalmic conditions. The products are formulated either as patches, which can be applied as wound covers, or as suspensions or particulates, or connective tissue extractions, which can be injected or applied topically.

Fresh amniotic membrane contains collagen, fibronectin, and hyaluronic acid, along with a combination of growth factors, cytokines, and anti-inflammatory proteins such as interleukin-1 receptor antagonist. There is evidence that the tissue has anti-inflammatory, antifibroblastic, and antimicrobial properties. HAM is considered nonimmunogenic and has not been observed to cause a substantial immune response. It is believed that these properties are retained in cryopreserved HAM and dehydrated HAM products, resulting in a readily available tissue with regenerative potential. In support, one dehydrated HAM product has been shown to elute growth factors into saline and stimulate the migration of mesenchymal stem cells, both in vitro and in vivo.

Use of a HAM graft, which is fixated by sutures, is an established treatment for disorders of the corneal surface, including neurotrophic keratitis, corneal ulcers and melts, following pterygium repair, Stevens-Johnson syndrome, and persistent epithelial defects. Amniotic membrane products that are inserted like a contact lens have more recently been investigated for the treatment of corneal and ocular surface disorders. Amniotic membrane patches are also being evaluated for the treatment of various other conditions, including skin wounds, burns, leg ulcers, and prevention of tissue adhesion in surgical procedures. Additional indications studied in preclinical models include tendonitis, tendon repair, and nerve repair. The availability of HAM opens the possibility of regenerative medicine for an array of conditions.

Amniotic Fluid

Amniotic fluid surrounds the fetus during pregnancy and provides protection and nourishment. In the second half of gestation, most of the fluid is a result of micturition and secretion from the respiratory tract and gastrointestinal tract of the fetus, along with urea. The fluid contains proteins, carbohydrates, peptides, fats, amino acids, enzymes, hormones, pigments, and fetal cells. Amniotic fluid has been compared with synovial fluid, containing hyaluronan, lubricant, cholesterol, and cytokines. Injection of amniotic fluid or amniotic fluid-derived cells is currently being evaluated for the treatment of osteoarthritis and plantar fasciitis.

REGULATORY STATUS

There are many artificial skin and soft-tissue products that are commercially available or in development. Information on specific products is available in a 2020 Technical Brief on skin substitutes for treating chronic wounds that was commissioned by the Agency for Healthcare Research and Quality.

The U.S. Food and Drug Administration (FDA) regulates human cells and tissues intended for implantation, transplantation, or infusion through the Center for Biologics Evaluation and Research. ADM and amniotic products are classified as banked human tissue and therefore,
not requiring FDA approval for homologous use. In 2017, the FDA published clarification of what is considered minimal manipulation and homologous use for human cells, tissues, and cellular and tissue-based products (HCT/Ps).[4]

HCT/Ps are defined as human cells or tissues that are intended for implantation, transplantation, infusion, or transfer into a human recipient. If an HCT/P does not meet the criteria below and does not qualify for any of the stated exceptions, the HCT/P will be regulated as a drug, device, and/or biological product and applicable regulations and premarket review will be required.

An HCT/P is regulated solely under section 361 of the PHS Act and 21 CFR Part 1271 if it meets all of the following criteria:

1. "The HCT/P is minimally manipulated;"
2. The HCT/P is intended for homologous use only, as reflected by the labeling, advertising, or other indications of the manufacturer’s objective intent;
3. The manufacture of the HCT/P does not involve the combination of the cells or tissues with another article, except for water, crystalloids, or a sterilizing, preserving, or storage agent, provided that the addition of water, crystalloids, or the sterilizing, preserving, or storage agent does not raise new clinical safety concerns with respect to the HCT/P; and
4. Either:
   i. The HCT/P does not have a systemic effect and is not dependent upon the metabolic activity of living cells for its primary function; or
   ii. The HCT/P has a systemic effect or is dependent upon the metabolic activity of living cells for its primary function, and:
      a. Is for autologous use;
      b. Is for allogeneic use in a first-degree or second-degree blood relative; or
      c. Is for reproductive use."

The guidance provides the following specific examples of homologous and non-homologous use for amniotic membrane:

a. "Amniotic membrane is used for bone tissue replacement to support bone regeneration following surgery to repair or replace bone defects. This is not a homologous use because bone regeneration is not a basic function of amniotic membrane.

b. An amniotic membrane product is used for wound healing and/or to reduce scarring and inflammation. This is not homologous use because wound healing and reduction of scarring and inflammation are not basic functions of amniotic membrane.

c. An amniotic membrane product is applied to the surface of the eye to cover or offer protection from the surrounding environment in ocular repair and reconstruction procedures. This is homologous use because serving as a covering and offering protection from the surrounding environment are basic functions of amniotic membrane."

The FDA noted the intention to exercise enforcement discretion for the next 36 months after publication of the guidance.
In 2003, Prokera® was cleared for marketing by the FDA through the 510(k) process for the ophthalmic conformer that incorporates amniotic membrane (K032104). The FDA determined that this device was substantially equivalent to the Symblepharon Ring. The Prokera® device is intended “for use in eyes in which the ocular surface cells have been damaged, or underlying stroma is inflamed and scarred.”[5] The development of Prokera®, a commercially available product, was supported in part by the National Institute of Health and the National Eye Institute.

AmnioClip (FORTECH GmbH) is a ring designed to hold the amniotic membrane in the eye without sutures or glue fixation. A mounting device is used to secure the amniotic membrane within the AmnioClip. The AmnioClip currently has CE approval in Europe.

**EVIDENCE SUMMARY**

Evidence reviews assess the clinical evidence to determine whether the use of technology improves health outcomes for patients. Broadly defined, health outcomes are the length of life, quality of life, and ability to function – including benefits and harms. The quality and credibility of the evidence depend on study design and conduct, minimizing bias and confounding that can generate incorrect findings. The randomized controlled trial (RCT) is preferred to assess efficacy; however, in some circumstances, nonrandomized studies may be adequate. RCTs are rarely large enough or long enough to capture less common adverse events and long-term effects. Other types of studies can be used for these purposes and to assess generalizability to broader clinical populations and settings of clinical practice.

The following is a summary of key literature to date.

**BREAST RECONSTRUCTION**

A meta-analysis by Lee and Mun (2016) included 23 studies (total n=6,199 cases) on implant-based breast reconstruction that were published between February 2011 and December 2014.[6] The analysis included an RCT and three prospective comparative cohort studies; the remainder was retrospective comparative cohort studies. Use of ADM did not affect the total complication rate (see Table 1). ADM significantly increased the risk of major infection, seroma, and flap necrosis, but reduced risks of capsular contracture and implant malposition. Use of ADM allowed for significantly greater intraoperative expansion (mean difference 79.63, 95% confidence interval [CI], 41.99 to 117.26, p<0.001) and percentage of intraoperative filling (mean difference 13.30, 95% CI 9.95 to 16.65, p<0.001), and reduced the frequency of injections to complete expansion (mean difference -1.56, 95% CI -2.77 to -0.35, p=0.01).

**Table 1. Meta-Analysis of Breast Reconstruction Outcomes with and without ADM**

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Relative Risk</th>
<th>95% Confidence Interval</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection</td>
<td>1.42</td>
<td>1.02 to 1.99</td>
<td>0.04</td>
</tr>
<tr>
<td>Seroma</td>
<td>1.41</td>
<td>1.12 to 1.78</td>
<td>0.004</td>
</tr>
<tr>
<td>Mastectomy flap necrosis</td>
<td>1.44</td>
<td>1.11 to 1.87</td>
<td>0.006</td>
</tr>
<tr>
<td>Unplanned return to the operating room</td>
<td>1.09</td>
<td>0.63 to 1.90</td>
<td>NS</td>
</tr>
<tr>
<td>Implant loss</td>
<td>1.00</td>
<td>0.68 to 1.48</td>
<td>NS</td>
</tr>
<tr>
<td>Total complications</td>
<td>1.08</td>
<td>0.87 to 1.34</td>
<td>NS</td>
</tr>
<tr>
<td>Capsular contracture</td>
<td>0.26</td>
<td>0.15 to 0.47</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Implant malposition</td>
<td>0.21</td>
<td>0.07 to 0.59</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Adapted from Lee and Mun (2016).[6]

ADM: acellular dermal matrix; NS: not significant.
A study by Davila (2013) used data from the American College of Surgeon’s National Surgical Quality Improvement Program to compare ADM-assisted tissue expander breast reconstruction (n=1,717) to submuscular tissue expander breast reconstruction (n=7,442) after mastectomy.[7] Complication rates did not differ significantly between the ADM-assisted (5.5%) and the submuscular tissue expander groups (5.3%, p=0.68). Rates of reconstruction-related complications, major complications, and 30-day reoperation did not differ significantly between cohorts.

**ALLODERM®**

**Randomized Controlled Trials**

McCarthy (2012) reported on a multicenter, blinded RCT of AlloDerm® in two-stage expander/implant reconstruction.[8] Seventy patients were randomized to AlloDerm® ADM-assisted tissue expander/implant reconstruction or to submuscular tissue expander/implant placement. The trial was adequately powered to detect clinically significant differences in immediate postoperative pain but underpowered to detect the secondary endpoint of pain during tissue expansion. There were no significant differences between the groups in the primary outcomes of immediate postoperative pain (54.6 AlloDerm® vs. 42.8 controls on a 100-point visual analog scale) or pain during the expansion phase (17.0 AlloDerm® vs. 4.6 controls) or in the secondary outcome of rate of tissue expansion (91 days AlloDerm® vs. 108 days controls) and patient-reported physical well-being. There was no significant difference in adverse events, although the total number of adverse events was small.

**Comparisons Between Products**

**AlloDerm® Versus AlloMax™**

Hinchcliff (2017) conducted an RCT that compared AlloDerm® with AlloMax™ (n=15 each) for implant-based breast reconstruction.[9] Complications were assessed 7, 14, and 30 days postoperatively and biopsies of the ADMs were taken during implant exchange. Vessel density in the AlloMax™ biopsies was higher than in the AlloDerm® biopsies. Complications were reported in 26.1% of AlloMax™ cases and 8.0% of AlloDerm® cases; these complication rates did not differ statistically with the 30 patients in this trial.

**AlloDerm® Versus DermaMatrix™**

Mendenhall (2017) conducted an RCT that compared AlloDerm® with DermaMatrix™ in 111 patients (173 breasts).[10] There were no significant differences in overall rates of complications (AlloDerm® 15.4%, DermaMatrix™ 18.3%, p=0.8) or implant loss (AlloDerm® 2.2%, DermaMatrix™ 3.7%, p=0.5) between the two ADMs.

**AlloDerm® Versus FlexHD®**

A retrospective review by Liu (2014) compared complication rates following breast reconstruction with AlloDerm® or FlexHD® in 382 consecutive women (547 breasts).[11] Eighty-one percent of the sample was immediate reconstruction: 165 used AlloDerm® and 97 used FlexHD®. Mean follow-up was 6.4 months. Compared with breast reconstruction without the use of AlloDerm® or FlexHD®, ADM had a higher rate of delayed healing (20.2% vs. 10.3%), although this finding might have been related to differences in fill volumes. In univariate analysis, there were no significant differences in complications (return to operating room, surgical site infection, seroma, hematoma, delayed healing, implant loss) between AlloDerm®
and FlexHD®. In multivariate analysis, there were no significant differences between AlloDerm® and FlexHD® for the return to the operating room, surgical site infection, seroma, or delayed healing. Independent risk factors for implant loss included the use of FlexHD®, single-stage reconstruction, and smoking.

**AlloDerm® Versus SimpliDerm®**

Tierney (2022) published a retrospective study compared AlloDerm® (n=69) with SimpliDerm® (n=38) for two-stage breast reconstruction following mastectomy. Adverse events were reported in 25.2% of patients and did not differ by group.

**AlloDerm® Versus FlexHD® Pliable and DermACELL®**

Chang and Liu (2017) reported on a prospective comparison of FlexHD® Pliable (32 breasts), AlloDerm® (22 breasts), and DermACELL® (20 breasts) in breast reconstruction. The choice of ADM was based on different years when each ADM was available for use at the investigators’ institution; patient demographics were comparable between groups. The pieces of ADM used were all the same size (8 × 16 cm) to eliminate an effect of size on outcomes. The time to drain removal was longer with AlloDerm® (26 days) than with FlexHD® (20 days) or DermACELL® (15 days, p=0.001). Complications were low (four in the Flex Pliable group, two in the AlloDerm® group, one in the DermACELL® group), with no significant differences between groups. At the time of exchange for a permanent implant or free flap reconstruction, all grafts had completely incorporated into the mastectomy skin flaps. No patients developed complications requiring removal of the ADM.

Pittman (2017) reported a retrospective pilot study of the use of AlloDerm® (50 breasts) and DermACELL® (50 breasts). The choice of ADM was based on products available during different years and patient demographics were similar between the two groups. Patients in the DermACELL® group had a significantly lower incidence of “red breast syndrome” (0% vs. 26%, p=0.001) and fewer days until drain removal (15.8 days vs. 20.6 days, p=0.017). There were no significant differences in the rates of other complications.

**Strattice™**

Dikmans (2017) reported on early safety outcomes from an open-label multicenter RCT that compared porcine ADM-assisted one-stage expansion with two-stage implant-based breast reconstruction (see Table 2). One-stage breast reconstruction with porcine ADM was associated with a higher risk of surgical complications, reoperation, and with removal of implant, ADM, or both (see Table 3). The trial was stopped early due to safety concerns, but it cannot be determined from this study design whether the increase in complications was due to the use of the xenogenic ADM or to the comparison between one-stage and two-stage reconstruction.

<table>
<thead>
<tr>
<th>Author</th>
<th>Countries</th>
<th>Sites</th>
<th>Dates</th>
<th>Participants</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dikmans (2017)</td>
<td>EU</td>
<td>8</td>
<td>2013-2015</td>
<td>Women intending to undergo skin-sparing mastectomy and immediate IBBR</td>
<td>Active: 59 patients (91 breasts)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>undergoing 1-stage IBBR with ADM</td>
<td>Comparator: 62 women (92 breasts)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>undergoing 2-stage IBBR</td>
</tr>
</tbody>
</table>
ADM: acellular dermal matrix; IBBR: implant-based breast reconstruction; RCT: randomized controlled trial.

Table 3. Summary of Key RCT Outcomes

<table>
<thead>
<tr>
<th>Study</th>
<th>Surgical Complications</th>
<th>Severe Adverse Events</th>
<th>Reoperation</th>
<th>Removal of Implant ADM, or Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dikmans (2017)</td>
<td>27 (46)</td>
<td>26 (29)</td>
<td>22 (37)</td>
<td>24 (26)</td>
</tr>
<tr>
<td>1-stage with ADM, n (%)</td>
<td>11 (18)</td>
<td>5 (5)</td>
<td>9 (15)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>3.81 (2.67 to 5.43), p&lt;0.001</td>
<td>3.38 (2.10 to 5.45), p&lt;0.001</td>
<td>8.80 (8.24 to 9.40), p&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

ADM: acellular dermal matrix; CI: confidence interval; OR: odds ratio; RCT: randomized controlled trial.

TENDON REPAIR

GraftJacket®

Barber (2012) reported an industry-sponsored multicenter RCT of augmentation with GraftJacket® human ADM for arthroscopic repair of large (>3 cm) rotator cuff tears involving two tendons.[16] Twenty-two patients were randomized to GraftJacket® augmentation and 20 patients to no augmentation. At a mean follow-up of 24 months (range 12-38 months), the American Shoulder and Elbow Surgeons score improved from 48.5 to 98.9 in the GraftJacket® group and from 46.0 to 94.8 in the control group (p=0.035). The Constant score improved from 41 to 91.9 in the GraftJacket® group and from 45.8 to 85.3 in the control group (p=0.008). The University of California, Los Angeles score did not differ significantly between groups. Gadolinium-enhanced MRI scans showed intact cuffs in 85% of repairs in the GraftJacket® group and 40% of repairs in the control group. However, no correlation was found between MRI findings and clinical outcomes. Rotator cuff retears occurred in three (14%) patients in the GraftJacket® group and nine (45%) patients in the control group.

Rashid (2020) reported disruption of the native extracellular matrix with either GraftJacket® or Permacol™ (porcine acellular dermis) as a patch overlay for rotator cuff repair in a small controlled study with 13 patients.[17] The disruption was greater in the Permacol™ group and there was an immune response in one of three patients following use of the xenograft.

SURGICAL REPAIR OF HERNIAS OR PARASTOMAL REINFORCEMENT

A systematic review by Bellows (2013) evaluated the clinical effectiveness of acellular collagen-based scaffolds for the repair of incisional hernias.[18] The bioprosthetic materials could be harvested from bovine pericardium, human cadaveric dermis, porcine small intestine mucosa, porcine dermal collagen, or bovine dermal collagen. Products included in the search were Surgisis®, Tutomesh®, Veritas®, AlloDerm®, FlexHD®, AlloMax™, CollaMend™, Permacol™, Strattice™, FortaGen®, ACell, DermaMatrix™, XenMatrix™, and SurgiMend®. Sixty publications with 1,212 repairs were identified and included in the review, although meta-analysis could not be performed. There were four level III studies (two AlloDerm®, two Permacol™); the remainder was level IV or V. The largest number of publications were on AlloDerm® (n=27) and Permacol™ (n=18). No publications on incisional hernia repair were identified for AlloMax®, FortaGen®, DermaMatrix™, or ACell. The overall incidence of a surgical site occurrence (e.g., postoperative infection, seroma/hematoma, pain, bulging, dehiscence, fistula, mechanical failure) was 82.6% for porcine small intestine mucosa, 50.7%
for xenogenic dermis, 48.3% for human dermis, and 6.3% for xenogenic pericardium. No comparative data were identified that could establish superiority to permanent synthetic meshes.

**AlloDerm® as an Overlay**

Espinosa-de-los-Monteros (2007) retrospectively reviewed 39 abdominal wall reconstructions with AlloDerm® performed in 37 patients and compared them with 39 randomly selected cases.[19] They reported a significant decrease in recurrence rates when human cadaveric acellular dermis was added as an overlay to primary closure plus rectus muscle advancement and imbrication in patients with medium-sized hernias. However, no differences were observed when adding human cadaveric acellular dermis as an overlay to patients with large-size hernias treated with underlay mesh.

**Comparisons Between Products**

**AlloDerm® Versus Surgisis® Gold**

Gupta (2006) compared the efficacy and complications associated with use of AlloDerm® and Surgisis® bioactive mesh in 74 patients who underwent ventral hernia repair.[20] The first 41 procedures were performed using Surgisis® Gold 8-ply mesh formed from porcine small intestine submucosa, and the remaining 33 patients had ventral hernia repair with AlloDerm®. Patients were seen 7 to 10 days after discharge from the hospital and at six weeks. Any signs of wound infection, diastasis, hernia recurrence, changes in bowel habits, and seroma formation were evaluated. The use of the AlloDerm® mesh resulted in eight (24%) hernia recurrences. Fifteen (45%) of the AlloDerm® patients developed a diastasis or bulging at the repair site. Seroma formation was only a problem in two patients.

**AlloDerm® Versus FlexHD®**

A study by Bochicchio (2013) compared AlloDerm® with FlexHD® for complicated hernia surgery.[21] From 2005 to 2007, AlloDerm® was used to repair large (>200 cm²), symptomatic, complicated ventral hernias that resulted from trauma or emergency surgery (n=55). From 2008 to 2010, FlexHD® was used to repair large, complicated ventral hernias in patients meeting the same criteria (n=40). The two groups were comparable at baseline. At one-year follow-up, all AlloDerm® patients were diagnosed with hernia recurrence (abdominal laxity, functional recurrence, true recurrence) requiring a second repair. Eleven (31%) patients in the FlexHD® group required a second repair. This comparative study is limited by the use of nonconcurrent comparisons, which is prone to selection bias and does not control for temporal trends in outcomes.

**FlexHD® Versus Strattice™**

Roth (2017) reported on a prospective study assessing clinical and quality-of-life outcomes following complex hernia repair with a human (FlexHD®) or porcine (Strattice™) ADM.[22] The study was funded by the Musculoskeletal Transplant Foundation, which prepares and supplies FlexHD®. Patients were enrolled if they had a hernia at least 6 cm in the transverse dimension, active or prior infection of the abdominal wall, and/or enterocutaneous fistula requiring mesh removal. Eighteen (51%) of the 35 patients had undergone a previous hernia repair. After abdominal wall repair with the ADM, 20 (57%) patients had a surgical site occurrence, and nearly one-third had hospital readmission. The type of biologic material did
not impact hernia outcomes. There was no comparison with synthetic mesh in this study, limiting interpretation.

**Strattice™ Versus Synthetic Mesh**

Bellows (2014) reported early results of an industry-sponsored multicenter RCT that compared Strattice™ (non-cross-linked porcine ADM, n=84) with a standard synthetic mesh (n=88) for the repair of inguinal hernias.[23] The trial was designed by the surgeons and was patient- and assessor-blinded to reduce risk of bias. Blinding continued through two years of follow-up. The primary outcome was resumption of activities of daily living at one year. Secondary outcomes included complications, recurrences, or chronic pain (i.e., pain that did not disappear by three months postsurgery). At three-month follow-up, there were no significant differences in either the occurrence or type of wound events (relative risk 0.98, 95% CI 0.52 to 1.86). Pain was reduced from one to three days postoperative in the group treated with Strattice™, but at three-month follow-up pain scores did not differ significantly between groups.

A double-blind RCT by Brunbjerg (2020) compared Strattice™ to synthetic mesh (Prolene®) to prevent hernia or bulging in 29 patients admitted to a single center in Denmark for pedicled transverse rectus abdominis musculocutaneous flap surgery.[24] At two-years post-surgery, bulging frequency was higher in the Strattice™ group (35.7%) than in the synthetic mesh group (6.7%), but the difference was not statistically significant (p = 0.11). Two Strattice™ patients developed a hernia, while none of the mesh patients did. No differences were found for abdominal muscle strength between baseline and two-year measurements.

**Strattice™ Versus No Reinforcement**

Also in 2014, the Parastomal Reinforcement With Strattice™ (PRISM) Study Group reported a multicenter, double-blinded, randomized trial of Strattice™ for parastomal reinforcement in patients undergoing surgery for permanent abdominal wall ostomies.[25] Patients were randomized to standard stoma construction with no reinforcement (n=58) or stoma construction with Strattice™ as parastomal reinforcement (n=55). At 24-month follow-up (n=75), the incidence of parastomal hernias was similar for the two groups (13.2% of controls, 12.2% of study group).

**Adverse Events**

Permacol™ (porcine acellular dermal matrix) was reported in a case series of 13 patients to result in recurrent intestinal fistulation and intestinal failure when used for abdominal reconstructive surgery.[26]

**DIABETIC LOWER-EXTREMITY ULCERS**

**Systematic Reviews**

A 2016 Cochrane review evaluated skin substitutes for the treatment of diabetic foot ulcers.[27] Seventeen trials (total n=1,655 participants) were included in the meta-analysis. Most trials identified were industry-sponsored, and an asymmetric funnel plot indicated publication bias. Pooled results of published trials found that skin substitutes increased the likelihood of achieving complete ulcer closure compared with standard of care (SOC) alone (relative risk 1.55, 95% CI 1.30 to 1.85). Use of skin substitutes also led to a statistically significant reduction in amputations (relative risk 0.43, 95% CI 0.23 to 0.81), although the absolute risk difference was small. Analysis by individual products found a statistically significant benefit on
ulcer closure for Apligraf®, EpiFix®, and Hyalograft-3D™. The products that did not show a statistically significant benefit for ulcer closure were Dermagraft®, GraftJacket®, Kaloderm®, and OrCel®.

Martinson and Martinson (2016) conducted an industry-sponsored analysis of Medicare claims data (13,193 treatment episodes) to compare efficacy and cost of skin substitutes for the management of diabetic foot ulcers.[28] Included in the analysis were treatment episodes with Apligraf® (37%), Dermagraft® (42%), Oasis® (19%), and Cytal® (MatriStem®, 2%). The mean number of applications was 3.24 for Apligraf®, 4.48 for Oasis®, 5.53 for Cytal®, and 5.96 for Dermagraft®. All comparisons were statistically significant. Healing at 90 days was modestly but statistically higher for Oasis® (63%) and Cytal® (62%) than for Apligraf® (58%) or Dermagraft® (58%). Amputation rates were similar after treatment with the four products, ranging from 1.3% for Oasis® to 2.1% for Cytal®.

A systematic review by Lakmal (2021) included eight RCTs, two prospective studies and two retrospective studies that evaluated the use of amniotic membrane allografts for the treatment of diabetic foot ulcers.[29] Generally, the studies reported that better wound closure rates were seen with the amniotic membrane products than with standard care, but a meta-analysis was not possible due to study heterogeneity.

Guo (2017) reported a systematic review of ADM for the treatment of diabetic foot ulcers.[30] Most data were from an RCT of Integra® Dermal Regeneration Template, which is a bilayer product with the outer layer composed of a thin silicone film and not a pure ADM.

Amniotic Membranes

At least seven RCTs have evaluated rates of healing with amniotic membrane grafts or placental membrane graft compared to SOC or an advanced wound therapy in patients with chronic diabetic foot ulcers (see Table 4). The number of patients in these studies ranged from 25 to 155. Human amniotic membrane (HAM) or placental membrane grafts improved healing compared to SOC by 22% (EpiCord® vs. alginate dressing) to 60% (EpiFix®) in the intention-to-treat (ITT) analysis (see Table 5). In a 2018 trial, the cryopreserved placental membrane Grafix® was found to be non-inferior to an advanced fibroblast-derived wound therapy (Dermagraft®).[31]

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Intervention</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serena (2020)</td>
<td>76 patients with chronic (&gt;4 weeks) non-healing diabetic foot ulcers unresponsive to SOC and extending into dermis, subcutaneous tissue, muscle, or tendon</td>
<td>n=38, Affinity</td>
<td>n=38, SOC</td>
</tr>
<tr>
<td>Ananian (2018)</td>
<td>75 patients with chronic (&gt;4 weeks) non-healing diabetic foot ulcers between 1 cm² and 15 cm²</td>
<td>n=38, Grafix® weekly for up to 8 weeks</td>
<td>n=37, Dermagraft® (fibroblast-derived) weekly for up to 8 weeks</td>
</tr>
</tbody>
</table>

Table 4. Summary of Key RCT Characteristics
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Intervention</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tettelbach (2019)[33]</td>
<td>155 patients with chronic (&gt; 4 weeks) non-healing diabetic foot ulcers</td>
<td>n=101 EpiCord® plus SOC</td>
<td>n=54 SOC with alginate dressing</td>
</tr>
<tr>
<td>DiDomenico (2018)[34]</td>
<td>80 patients with non-healing (4 weeks) diabetic foot ulcers</td>
<td>AmnioBand® Membrane plus SOC</td>
<td>SOC</td>
</tr>
<tr>
<td>Snyder (2016)[35]</td>
<td>29 patients with non-healing diabetic foot ulcers</td>
<td>AmnioExcel® plus SOC</td>
<td>SOC</td>
</tr>
<tr>
<td>Zelen (2015, 2016)[36, 37]</td>
<td>60 patients with less than 20% wound healing in a 2-week run-in period</td>
<td>EpiFix®</td>
<td>Apligraf® or SOC with collagen-alginate dressing</td>
</tr>
<tr>
<td>Tettelbach (2019)[38]</td>
<td>110 patients with non-healing (4 weeks) lower extremity ulcers</td>
<td>EpiFix®</td>
<td>SOC with alginate dressing</td>
</tr>
<tr>
<td>Lavery (2014)[39]</td>
<td>97 patients with chronic diabetic foot ulcers</td>
<td>Grafix® Weekly</td>
<td>SOC</td>
</tr>
</tbody>
</table>

RCT: randomized controlled trial; SOC: standard of care including debridement, nonadherent dressing, moisture dressing, a compression dressing and offloading.

Table 5. Summary of Key RCT Results

<table>
<thead>
<tr>
<th>Study</th>
<th>Wounds Healed</th>
<th>Time to Complete Healing</th>
<th>Adverse Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serena (2020)[32]</td>
<td>16 Weeks (ITT)</td>
<td>Median</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>76</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Affinity</td>
<td>58%</td>
<td>11 weeks</td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>29%</td>
<td>not attained by 16 weeks</td>
<td></td>
</tr>
<tr>
<td>HR (95% CI), p-value</td>
<td>1.75 (1.16 to 2.70), p=0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ananian (2018)[31]</td>
<td>8 Weeks (PP) n (%)</td>
<td></td>
<td>Patients with Index Ulcer Related Adverse Events n (%)</td>
</tr>
<tr>
<td>N</td>
<td>62</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Grafix®</td>
<td>15 (48.4%)</td>
<td>1 (5.9%)</td>
<td></td>
</tr>
<tr>
<td>Dermagraft®</td>
<td>12 (38.7%)</td>
<td>4 (16.7%)</td>
<td></td>
</tr>
<tr>
<td>Diff (95% CI), Lower bound for non-inferiority</td>
<td>9.68% (~-10.7 to 28.9), -15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tettelbach (2018)[33]</td>
<td>12 Weeks (ITT) n (%)</td>
<td></td>
<td>Patients with Adverse Events (% of total)</td>
</tr>
<tr>
<td>N</td>
<td>155</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td>EpiCord®</td>
<td>71 (70%)</td>
<td>42 (42%)</td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>26 (48%)</td>
<td>33 (61%)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiDomenico (2018)[34]</td>
<td>12 weeks (ITT) n (%)</td>
<td></td>
<td>Mean Days (95% CI)</td>
</tr>
<tr>
<td>N</td>
<td>80</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Wounds Healed</td>
<td>Time to Complete Healing</td>
<td>Adverse Events</td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>--------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Amnioband®</td>
<td>34 (85)</td>
<td>37.0 (29.5 to 44.4)</td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>13 (33)</td>
<td>67.3 (59.0 to 79.6)</td>
<td></td>
</tr>
<tr>
<td>HR (95% CI)</td>
<td></td>
<td>4.25 (0.44 to 0.79), p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Snyder (2016)</td>
<td>6 Weeks (PP) Mean (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmnioExcel®</td>
<td>45.5% (32.9% to 58.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zelen (2015, 2016)</td>
<td>Wounds Healed at 12 Weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EpiFix®</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apligraf®</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HR (95% CI)</td>
<td>5.66; (3.03 to 10.57), p&lt;0.001 vs. SOC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tettelbach (2019)</td>
<td>Wounds Healed at 12 Weeks (ITT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EpiFix®</td>
<td>70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lavery (2014)</td>
<td>Wounds Healed at 12 Weeks</td>
<td>Patients with Adverse Events</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>97</td>
<td>97</td>
<td>97</td>
</tr>
<tr>
<td>Grafix®</td>
<td>62.0%</td>
<td>42.0</td>
<td>44.0%</td>
</tr>
<tr>
<td>SOC</td>
<td>21.3%</td>
<td>69.5</td>
<td>66.0%</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.001</td>
<td>0.019</td>
<td>0.031</td>
</tr>
</tbody>
</table>

CI: confidence interval; DIFF: difference; HR: hazard ratio; ITT: intention-to-treat; NR: not reported; PP: per-protocol; RCT: randomized controlled trial; SOC: standard of care.

Many of these studies had methodologic limitations, including a lack of blinding and loss of patients to follow-up.

Smiehl (2015) reported on an industry-sponsored, multicenter registry study of Biovance® d-HAM for the treatment of various chronic wound types; about a third (n=47) were diabetic foot wounds. Of those treated, 28 ulcers had failed prior treatment with advanced biologic therapies. For all wound types, 41.6% closed within a mean time of eight weeks and a mean of 2.4 amniotic membrane applications.

Frykberg (2017) reported treatment of complex chronic wounds (exposed tendon or bone) with Grafix®. With the cryopreserved placental membrane applied weekly for up to 16 weeks, 59% of wounds closed with a mean time to closure of nine weeks.

Apligraf®
Veves (2001) reported on a randomized prospective trial on the effectiveness of Apligraf® (previously called Graftskin), a living skin equivalent, in treating noninfected nonischemic chronic plantar diabetic foot ulcers. The trial involved 24 centers in the United States; 208 patients were randomized to ulcer treatment with Apligraf® (112 patients) or saline-moistened gauze (96 patients, control group). Standard state-of-the-art adjunctive therapy, including extensive surgical débridement and adequate foot off-loading, was provided in both groups. Apligraf® was applied at the beginning of the study and weekly thereafter for a maximum of four weeks (maximum of five applications) or earlier if complete healing occurred. At the 12-week follow-up visit, 63 (56%) Apligraf®-treated patients achieved complete wound healing compared with 36 (38%) in the control group (p=0.004). The Kaplan-Meier method median time to complete closure was 65 days for Apligraf®, which was significantly lower than the 90 days observed in the control group (p=0.003). The rates of adverse reactions were similar between groups, except osteomyelitis and lower-limb amputations, both of which were less frequent in the Apligraf® group. Trialists concluded that application of Apligraf® for a maximum of four weeks resulted in higher healing rates than state-of-the-art treatment and was not associated with any significant adverse events. This trial was reviewed in a 2001 TEC Assessment, which concluded that Apligraf®, in conjunction with good local wound care, met the TEC criteria for the treatment of diabetic ulcers that fail to respond to conservative management.

Steinberg (2010) reported on a study of 72 subjects from Europe and Australia that assessed the safety and efficacy of Apligraf® in the treatment of noninfected diabetic foot ulcers. The study design and patient population were similar to the 208-subject U.S. study (previously described), which led to FDA approval of Apligraf® for the treatment of diabetic foot ulcers. For these studies, subjects with noninfected neuropathic diabetic foot ulcers present for at least two weeks were enrolled in prospective, multicenter, open-label RCTs that compared Apligraf® use plus standard therapy (sharp débridement, standard wound care, off-loading) with standard therapy alone. Pooling of data was performed because of the similarity and consistency of the two studies. Efficacy and safety results were consistent across studies independent of mean ulcer duration, which was significantly longer in the European study (21 months vs.10 months in the U.S. study). Reported adverse events by 12 weeks were comparable across treatment groups in the two studies. Efficacy measures demonstrated superiority of Apligraf® treatment over control-treated groups in both studies. Combining the data from both studies, 55.2% (80/145) of Apligraf® subjects had complete wound closure by 12 weeks, compared with 34.3% (46/134) of control subjects (p<0.001), and Apligraf® subjects had a significantly shorter time to complete wound closure (p<0.001). The authors concluded that both the EU and U.S. studies exhibited superior efficacy and comparable safety for subjects treated with Apligraf® compared with control subjects and that the studies provided evidence of the benefit of Apligraf® in treating diabetic foot ulcer.

Kirsner (2010) analyzed 2,517 patients with diabetic neuropathic foot ulcers treated between 2001 and 2004. This retrospective analysis used a wound care database; the patients received advanced biologic therapy, specifically, Apligraf® (446 patients), Regranex®, or Procuren®. The analysis found that advanced biologic therapy was used, on average, within 28 days from the first wound clinic visit and was associated with a median time to healing of 100 days. Wounds treated with engineered skin (Apligraf®) as the first advanced biologic therapy were 31% more likely to heal than wounds first treated with topical recombinant growth factor (p<0.001) and 40% more likely to heal than those first treated with platelet releasate (p=0.01). Wound size, wound grade, duration of wound, and time to initiation of advanced biologic therapy affected the time to healing.
**Dermagraft®**

A 2003 pivotal multicenter FDA-regulated trial randomized 314 patients with chronic diabetic ulcers to Dermagraft® (human-derived fibroblasts cultured on mesh) or control.[46] Over the 12-week study, patients received up to eight applications of Dermagraft®. All patients received pressure-reducing footwear and were encouraged to stay off their study foot as much as possible. At 12 weeks, the median percent wound closure for the Dermagraft® group was 91% compared with 78% for the control group. Ulcers treated with Dermagraft® closed significantly faster than ulcers treated with conventional therapy. No serious adverse events were attributed to Dermagraft®. Ulcer infections developed in 10.4% of the Dermagraft® patients compared with 17.9% of the control patients. Together, there was a lower rate of infection, cellulitis, and osteomyelitis in the Dermagraft®-treated group (19% vs. 32.5%). A 2015 retrospective analysis of the trial data found a significant reduction in amputation/bone resection rates with Dermagraft® (5.5% vs. 12.6%, p=0.031).[47] Of the 28 cases of amputation/bone resection, 27 were preceded by ulcer-related infection.

**AlloPatch®**

AlloPatch® Pliable human reticular acellular dermis was compared with SOC in an industry-sponsored multicenter trial by Zelen (2017, 2018).[48, 49] The initial trial with 20 patients per group was extended to determine the percent healing at six weeks with 40 patients per group. Healing was evaluated by the site investigator and confirmed by an independent panel. At six weeks, 68% (27/40) of wounds treated using AlloPatch® had healed compared with 15% (6/40) in the SOC-alone group (p<0.001). At 12 weeks, 80% (32/40) of patients in the AlloPatch® group had healed compared to 30% (12/40) in the control group. Mean time to heal within 12 weeks was 38 days (95% CI 29 to 47 days) for the HR-ADM group and 72 days (95% CI 66 to 78 days) for the SOC group (p<0.001).

**Integra® Omnigraft Dermal Regeneration Template or Integra® Flowable Wound Matrix**

Integra® Dermal Regeneration Template is a biosynthetic skin substitute that is FDA-approved for life-threatening thermal injury. The FOUNDER (Foot Ulcer New Dermal Replacement) multicenter study (32 sites) assessed Integra® Dermal Regeneration Template (marketed as Omnigraft™) for chronic nonhealing diabetic foot ulcers under an FDA-regulated investigational device exemption.[50] A total of 307 patients with at least one chronic diabetic foot ulcer were randomized to treatment with the Integra® Template or a control condition (sodium chloride gel 0.9%). Treatment was given for 16 weeks or until wound closure. There was a modest increase in wound closure with the Integra® Template (51% vs. 32%, p=0.001) and a shorter median time to closure (43 days vs. 78 days, p=0.001). There was a strong correlation between investigator-assessed and computerized planimetry assessment of wound healing (r=0.97). Kaplan-Meier analysis showed the greatest difference between groups in wound closure up to 10 weeks, with diminishing differences after 10 weeks. Trial strengths included adequate power to detect an increase in wound healing of 18%, which was considered to be clinically significant, secondary outcomes of wound closure and time to wound closure by computerized planimetry, and intention-to-treat (ITT) analysis.

Integra® Flowable Wound Matrix is composed of a porous matrix of cross-linked bovine tendon collagen and glycosaminoglycan. It is supplied as a granular product that is mixed with saline. Campitiello (2017) published an RCT that compared the flowable matrix with wet dressing in 46 patients who had Wagner grade 3 diabetic foot ulcers.[51] The ulcers had developed over 39 weeks. Complete healing at six weeks was achieved in significantly more
patients in the Integra® Flowable Wound Matrix group than in the control group, while the risk of rehospitalization and major amputation was reduced with Integra® Flowable Wound Matrix (see Table 6).

Table 6. Probability of Wound Healing with IFWM Versus SOC

<table>
<thead>
<tr>
<th>Study</th>
<th>Complete Wound Healing</th>
<th>Rehospitalization</th>
<th>Major Amputation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campitiello (2017)[51]</td>
<td>IFWM, n (%)</td>
<td>20 (86.95)</td>
<td>2 (6.69)</td>
</tr>
<tr>
<td></td>
<td>SOC, n (%)</td>
<td>12 (52.17)</td>
<td>10 (43.47)</td>
</tr>
<tr>
<td>RR (95% CI)</td>
<td>1.67 (1.09 to 2.54)</td>
<td>0.10 (0.01 to 0.72)</td>
<td>0.16 (0.02 to 1.17)</td>
</tr>
<tr>
<td>p</td>
<td>0.010</td>
<td>0.001</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Cl: confidence interval; IFWM: Integra® Flowable Wound Matrix; RR: relative risk; SOC: standard of care.

GraftJacket® Regenerative Tissue Matrix

Brigido (2004) reported a small (n=40) randomized pilot study comparing GraftJacket® with conventional treatment for chronic nonhealing diabetic foot ulcers.[52] Control patients received conventional therapy with débridement, wound gel with gauze dressing, and off-loading. GraftJacket® patients received surgical application of the scaffold using skin staples or sutures and moistened compressive dressing. A second graft application was necessary after the initial application for all patients in the GraftJacket® group. Preliminary one-month results showed that, after a single treatment, ulcers treated with GraftJacket® healed at a faster rate than conventional treatment. There were significantly greater decreases in wound length (51% vs. 15%), width (50% vs. 23%), area (73% vs. 34%), and depth (89% vs. 25%), respectively. With follow-up to four weeks, no data were reported on the proportion with complete closure or the mean time to heal. All grafts were incorporated into the host tissue.

Reyzelman (2009) reported an industry-sponsored multicenter randomized study that compared a single application of GraftJacket® with SOC in 86 patients with diabetic foot ulcers.[53] Eight patients, six in the study group and two in the control group, did not complete the trial. At 12 weeks, complete healing was observed in 69.6% of the GraftJacket® group and 46.2% of controls. After adjusting for ulcer size at presentation, a statistically significant difference in nonhealing rate was calculated, with odds of healing 2.0 times higher in the study group. Mean healing time was 5.7 weeks for the GraftJacket® group versus 6.8 weeks for the control group. The authors did not report whether this difference was statistically significant. Median time to healing was 4.5 weeks for GraftJacket® (range 1-12 weeks) and 7.0 weeks for control (range 2-12 weeks). Kaplan-Meier method survivorship analysis for time to complete healing at 12 weeks showed a significantly lower nonhealing rate for the study group (30.4%) than for the control group (53.9%). The authors commented that a single application of GraftJacket®, as used in this study, was often sufficient for complete healing. Conclusions drawn from this study are limited by the small study population and differences in ulcer size at baseline. Questions also remain about whether the difference in mean time to healing is statistically or clinically significant.

Reyzelman and Bazarov (2015)[54] reported the results of an industry-sponsored meta-analysis of GraftJacket® for diabetic foot ulcers, which included the two studies described above and a third RCT by Brigido (2006)[55] (total n=154 patients). The time to heal was estimated for the Brigido (2004) study,[52] based on the average wound reduction per week. The estimated difference in time to heal was larger for Brigido’s (2004) study (~4.30 weeks) than for the other two studies that measured the difference in time to heal (~1.58 weeks and ~1.10 weeks).
Analysis of the proportion of wounds that healed included Brigido (2006) and Reyzelman (2009). The odds ratio in the smaller study by Brigido (2006) was considerably larger, with a lack of precision in the estimate (odds ratio, 15.0, 95% CI 2.26 to 99.64), and the combined odds (3.75, 95% CI 1.72 to 8.19) was not significant when analyzed using a random-effects model. Potential sources of bias included publication and reporting biases, study selection biases, incomplete data selection, post hoc manipulation of data, and subjective choice of analytic methods.

**DermACELL® Versus GraftJacket® Regenerative Tissue Matrix or Standard of Care**

DermACELL® and GraftJacket® are both composed of human ADM. Walters (2016) reported on a multicenter randomized comparison of DermACELL®, GraftJacket®, or SOC (2:1:2 ratio) in 168 patients with diabetic foot ulcers.[56] The study was sponsored by LifeNet Health, a nonprofit organ procurement association and processor for DermACELL®. At 16 weeks, the proportion of completely healed ulcers was 67.9% for DermACELL®, 47.8% for GraftJacket®, and 48.1% for SOC. The 20% difference in completely healed ulcers was statistically significant for DermACELL® versus SOC (p=0.039). The mean time to complete wound closure did not differ significantly for DermACELL® (8.6 weeks), GraftJacket® (8.6 weeks), and SOC (8.7 weeks).

A second report from this study was published by Cazzell (2017).[57] This analysis compared DermACELL® with SOC and did not include the GraftJacket® arm. The authors reported that either one or two applications DermACELL® led to a greater proportion of wounds healed compared with SOC in per-protocol analysis, but there was no significant difference between DermACELL® (one or two applications) and SOC when analyzed by ITT. For the group of patients who received only a single application, the percentage of patients who achieved complete wound healing was significantly higher than SOC at 16 and 24 weeks, but not at 12 weeks. Although reported as ITT analysis, results were analyzed only for the group who received a single application of DermACELL®. This would not typically be considered ITT unless the number of DermACELL® applications was prespecified.

**Theraskin® Versus Standard of Care**

An industry funded retrospective study by Gurtner (2020) was a matched comparison of Theraskin® to standard of care alone in 3,994 lower extremity wounds of multiple etiologies.[58] Data were collected from electronic medical records from 644 wound care centers that were managed by a single large wound management company. Patients were matched for eight characteristics including wound size, severity, duration, comorbidities and body mass index. Diabetic wounds comprised 42% of the total cases and venous ulcers 29%. The next most frequent etiologies were pressure ulcers (~8%), surgical wounds (~9%), and trauma (~8%). Patients were excluded from analysis if they had greater than 50% wound closure during a four-week run-in period. The overall healing rate was 68.3% in the allograft group and 60.3% for standard of care (p<0.001). Diabetic wounds were treated with an average of 2.8 allografts prior to closure with a difference in closure rates of approximately 12% (67.5% vs 55.1%). A limitation of this retrospective analysis is that although the groups were well matched on a number of variables, the application of the Theraskin® allograft was at the investigators discretion and not standardized.

**Theraskin® Versus Dermagraft®**
Sanders (2014) reported on a small (n=23) industry-funded randomized comparison of Theraskin® (cryopreserved human skin allograft with living fibroblasts and keratinocytes) and Dermagraft® for diabetic foot ulcers. Wound size at baseline ranged from 0.5 to 18.02 cm²; the average wound size was about 5 cm² and was similar for the two groups (p=0.51). Grafts were applied according to manufacturers’ instructions over the first 12 weeks of the study until healing, with an average of 4.4 Theraskin® grafts (every two weeks) compared with 8.9 Dermagraft® applications (every week). At week 12, complete wound healing was observed in 63.6% of ulcers treated with Theraskin® and 33.3% of ulcers treated with Dermagraft® (p<0.049). At 20 weeks, complete wound healing was observed in 90.9% of the Theraskin®-treated ulcers compared with 66.7% of the Dermagraft® group (p=0.428).

**Theraskin® Versus Apligraf®**

DiDomenico (2011) compared Theraskin® with Apligraf® for the treatment of diabetic foot ulcers in a small (n=29) RCT. The risk of bias in this study is uncertain because reporting did not include a description of power analysis, statistical analysis, method of randomization, or blinding. The percentage of wounds closed at 12 weeks was 41.3% in the Apligraf® group and 66.7% in the Theraskin® group. Results at 20 weeks were not substantially changed from those at 12 weeks, with 47.1% of wounds closed in the Apligraf® group and 66.7% closed in the Theraskin® group. The percentage healed in the Apligraf® group was lower than expected based on prior studies. The average number of grafts applied was similar for both groups (1.53 for Apligraf®, 1.38 for Theraskin®). The low number of dressing changes may have influenced results, with little change in the percentage of wounds closed between 12 and 20 weeks. An adequately powered trial with blinded evaluation of wound healing and a standard treatment regimen would permit greater certainty on the efficacy of this product.

**Cytal® (MatriStem) Versus Dermagraft®**

Frykberg (2016) reported a prespecified interim analysis of an industry-funded multicenter noninferiority trial of Cytal® (a porcine urinary bladder-derived extracellular matrix) versus Dermagraft® in 56 patients with diabetic foot ulcers. The mean duration of ulcers before treatment was 263 days (range, 30-1095 days). The primary outcome was the percent wound closure with up to eight weeks of treatment using blinded evaluation of photographs. ITT analysis found complete wound closure in five (18.5%) wounds treated with Cytal® compared with two (6.9%) wounds treated with Dermagraft® (not statistically significant). Quality of life, measured by the Diabetic Foot Ulcer Scale, improved from 181.56 to 151.11 in the Cytal® group and from 184.46 to 195.73 in the Dermagraft® group (p=0.074). It should be noted that this scale is a subjective measure and patients were not blinded to treatment.

**PriMatrix®**

Lantis (2021) reported on a multicenter RCT comparing PriMatrix® plus standard of care to PriMatrix® alone in 226 patients with diabetic foot ulcers. Study subjects underwent a two-week run-in period of SOC treatment and were excluded if they had a wound reduction of 30% or more. Patients randomized to the SOC group received weekly treatment at the study site identical to the SOC treatment applied during the screening period. In addition, control group patients performed daily dressing changes, which consisted of wound cleaning, application of saline gel and secondary dressings. The primary endpoint was the percentage of subjects with complete wound closure, defined as 100% re-epithelialization without drainage during the 12-week treatment phase. Significantly more patients in the PriMatrix® group experienced complete wound closure at 12 weeks (45.6% vs. 27.9%, p=0.008). It is unclear if this difference...
(17.7%) is clinically significant; the study was powered to detect a 20% difference between groups. The time to complete healing did not differ between groups for the wounds that healed. Major study limitations include lack of blinding, limited generalizability, and insufficient duration of follow-up to assess wound recurrence.

Kavros (2014) reported a prospective multicenter study of PriMatrix® (a xenograft fetal bovine dermal collagen matrix) for the treatment of chronic diabetic foot ulcers in 55 patients.\(^{[63]}\) Average duration of ulcers before treatment was 286 days, and average wound area was 4.34 cm\(^2\). Of the 46 patients who completed the study, 76% healed by 12 weeks with an average of two applications of PriMatrix®. For the ITT population, 64% of wounds healed by 12 weeks.

Karr (2011) published a retrospective comparison of PriMatrix® and Apligraf® in 40 diabetic foot ulcers.\(^{[64]}\) The first 20 diabetic foot ulcers matching the inclusion and exclusion criteria for each graft were compared. The criteria were: diabetic foot ulcers of four weeks in duration; ulcer of at least 1 cm\(^2\) in diameter and to the depth of subcutaneous tissue; healthy tissue at the ulcer; adequate arterial perfusion to heal; and ability to off-load the diabetic ulcer. The time to complete healing for PriMatrix® was 38 days with 1.5 applications compared with 87 days with two applications for Apligraf®. Although promising, additional study with a larger number of subjects is needed to compare the efficacy of PriMatrix® with current SOC or advanced wound therapies.

### Oasis® Wound Matrix Versus Regranex Gel

Niezgoda (2005) compared healing rates at 12 weeks for full-thickness diabetic foot ulcers treated with Oasis® Wound Matrix (a porcine acellular wound care product) to Regranex Gel.\(^{[65]}\) This industry-sponsored, multicenter RCT was conducted at nine outpatient wound care clinics and involved 73 patients with at least one diabetic foot ulcer. Patients were randomized to receive either Oasis® Wound Matrix (n=37) or Regranex Gel (n=36) and secondary dressing. Wounds were cleaned and débrided, if needed, at a weekly visit. The maximum treatment period for each patient was 12 weeks. After 12 weeks, 18 (49%) Oasis®-treated patients had complete wound closure compared with 10 (28%) Regranex-treated patients. Oasis® treatment met the noninferiority margin but did not demonstrate that healing in the Oasis® group was statistically superior (p=0.055). Post hoc subgroup analysis showed no significant difference in incidence of healing in patients with type 1 diabetes (33% vs. 25%) but showed a significant improvement in patients with type 2 diabetes (63% vs. 29%). There was also increased healing of plantar ulcers in the Oasis® group (52% vs. 14%). These post hoc findings are considered hypothesis-generating. Additional study with a larger number of subjects is needed to compare the effect of Oasis® treatment to current SOC.

### Autologous Grafting on HYAFF Scaffolds

Uccioli (2011) reported a multicenter RCT of cultured expanded fibroblasts and keratinocytes grown on an HYAFF scaffold (benzyl ester of hyaluronic acid) compared with paraffin gauze for difficult diabetic foot ulcers.\(^{[66]}\) A total of 180 patients were randomized. At 12 weeks, complete ulcer healing was similar for the two groups (24% treated vs. 21% controls). At 20 weeks, complete ulcer healing was achieved in a similar proportion of the treatment group (50%) and the control group (43%, log-rank test = 0.344). Subgroup analysis, adjusted for baseline factors and possibly post-hoc, found a statistically significant benefit of treatment on dorsal ulcers but not plantar ulcers.

### Omega3 Wound

**MED170 | 23**
Lullove (2021) reported results of a RCT of the Kerecis™ Omega3 Wound plus standard wound care compared to standard care alone in 49 patients with diabetic lower extremity skin ulcers.[67] The primary outcome was healing at 12 weeks. Complete ulcer healing was based on the site investigator’s assessment, as evidenced by complete (100%) re-epithelialization without drainage and need of dressing. An independent panel of wound care experts who were blinded to the patient allocation process and the principal investigator’s assessment reviewed all study-related decisions made by the site investigators and confirmed healing status. Secondary outcomes were time to heal and wound area reduction by percentage at 12 weeks. Patients underwent a two-week run-in period prior to randomization. If the ulcer reduced in area by 20% or more after 14 days of standard care, the patient was excluded as a screening failure. If the wound area was reduced by less than 20%, the patient was randomized and enrolled in the study. Of 58 patients screened, 49 were eligible to enroll and were randomized. Significantly more patients in the Omega3 Wound group experienced healing at 12 weeks (67% vs. 32%, p=0.015), but time to healing was similar between groups for wounds that healed completely. Adverse events were not reported.

LOWER-EXTREMITY ULCERS DUE TO VENOUS INSUFFICIENCY

EpiFix®

Two RCTs evaluated the use of EpiFix® for venous leg ulcers. Serena (2014) reported on an industry-sponsored multicenter open-label RCT that compared EpiFix® d-HAM plus compression therapy with compression therapy alone for venous leg ulcers.[68] The primary outcome in this trial was the proportion of patients with 40% wound closure at four weeks, which was achieved by about twice as many patients in the combined EpiFix® group compared with the control group. However, a similar percentage of patients in the combined EpiFix® group and the control group achieved complete wound closure during the four-week study. There was no significant difference in healing for wounds given one versus two applications of amniotic membrane (62% vs. 63%, respectively). Strengths of this trial included adequate power and ITT analysis with last observation carried forward. Limitations included the lack of blinding for wound evaluation and use of 40% closure rather than complete closure. A 2015 retrospective study of 44 patients from this RCT (31 treated with amniotic membrane) found that wounds with at least 40% closure at four weeks (n=20) had a closure rate of 80% by 24 weeks; however, this analysis did not account for additional treatments after the four-week randomized trial period.

A second industry-sponsored multicenter open-label RCT, reported by Bianchi (2018, 2019), evaluated the time to complete ulcer healing following weekly treatment with EpiFix® d-HAM plus compression therapy or compression wound therapy alone.[69, 70] Patients treated with EpiFix® had a higher probability of complete healing by 12 weeks, as adjudicated by blinded outcome assessors (hazard ratio 2.26, 95% CI 1.25 to 4.10, p=0.01), and improved time to complete healing, as assessed by Kaplan-Meier analysis. In per-protocol analysis, healing within 12 weeks was reported for 60% of patients in the EpiFix® group and 35% of patients in the control group (p<0.013). Intent-to-treat analysis found complete healing in 50% of patients in the EpiFix® group compared to 31% of patients in the control group (p=0.0473). There were several limitations of this trial. In the per-protocol analysis, 19 (15%) patients were excluded from the analysis, and the proportion of patients excluded differed between groups (19% from the EpiFix® group vs. 11% from the control group). There was also a difference between the groups in how treatment failures at eight weeks were handled. Patients in the control group who did not have a 40% decrease in wound area at eight weeks were considered study...
failures and treated with advanced wound therapies. The ITT analysis used last-observation-carried-forward for these patients and sensitivity analysis was not performed to determine how alternative methods of handling the missing data would affect results. Kaplan-Meier analysis suggested a modest improvement in the time to heal when measured by ITT analysis, but may be subject to the same methodological limitations.

**Biovance**

As described above, Smiell (2015) reported on an industry-sponsored, multicenter registry study of Biovance d-HAM for the treatment of various chronic wound types; about half (n=89) were venous ulcers. Of the 179 treated, 28 (16%) ulcers had failed prior treatment with advanced biologic therapies. For all wound types, 41.6% closed within a mean time of eight weeks and a mean of 2.4 amniotic membrane applications. However, without a control group, the percentage of wounds that would have healed with SOC is unknown.

**Apligraf®**

Falanga (1998) reported on a multicenter randomized trial of Apligraf® living cell therapy. A total of 293 patients with venous insufficiency and clinical signs of venous ulceration were randomized to compression therapy alone or to compression therapy and treatment with Apligraf®. Apligraf® was applied up to a maximum of five (mean, 3.3) times per patient during the initial three weeks. The primary endpoints were the percentage of patients with complete healing by six months after initiation of treatment and the time required for complete healing. At six-month follow-up, the percentage of patients healed was higher with Apligraf® (63% vs. 49%), and the median time to complete wound closure was shorter (61 days vs. 181 days). Treatment with Apligraf® was superior to compression therapy in healing larger (>1000 mm²) and deeper ulcers and ulcers of more than six months in duration. There were no symptoms or signs of rejection, and the occurrence of adverse events was similar in both groups. This study was reviewed in a 2001 TEC Assessment, which concluded that Apligraf® (Graftskin), in conjunction with good local wound care, met TEC criteria for the treatment of venous ulcers that fail to respond to conservative management.

**Oasis® Wound Matrix**

Mostow (2005) reported on an industry-sponsored multicenter (12 sites) randomized trial that compared weekly treatment using Oasis® Wound Matrix (xenogenic collagen scaffold from porcine small intestinal mucosa) with SOC in 120 patients who had chronic ulcers due to venous insufficiency that had not adequately responded to conventional therapy. Healing was assessed weekly for up to 12 weeks, with follow-up performed after six months to assess recurrence. After 12 weeks of treatment, there was a significant improvement in the percentage of wounds healed in the Oasis® group (55% vs. 34%). After adjusting for baseline ulcer size, patients in the Oasis® group were 3 times more likely to heal than those in the group receiving SOC. Patients in the SOC group whose wounds did not heal by week 12 were allowed to cross over to Oasis® treatment. None of the healed patients treated with Oasis® wound matrix who was seen for the 6-month follow-up experienced ulcer recurrence.

A research group in Europe has described two comparative studies of the Oasis® matrix for mixed arteriovenous ulcers. In a quasi-randomized study, Romanelli (2007) compared the efficacy of two extracellular matrix-based products, Oasis® and Hyaloskin® (extracellular matrix with hyaluronic acid). Fifty-four patients with mixed arteriovenous leg ulcers were assigned to the two arms based on order of entry into the study; 50 patients completed the
study. Patients were followed twice weekly, and dressings changed more than once a week, only when necessary. After 16 weeks of treatment, complete wound closure was achieved in 82.6% of Oasis®-treated ulcers compared with 46.2% of Hyaloskin®-treated ulcers. Oasis® treatment significantly increased the time to dressing change (mean, 6.4 days vs. 2.4 days), reduced pain on a 10-point scale (3.7 vs. 6.2), and improved patient comfort (2.5 vs. 6.7).

Romanelli (2010) compared Oasis® with a moist wound dressing (SOC) in 23 patients with mixed arteriovenous ulcers and 27 patients with venous ulcers. The trial was described as randomized, but the method of randomization was not described. After the eight-week study period, patients were followed monthly for six months to assess wound closure. Complete wound closure was achieved in 80% of the Oasis®-treated ulcers at eight weeks compared with 65% of the SOC group. On average, Oasis®-treated ulcers achieved complete healing in 5.4 weeks compared with 8.3 weeks for the SOC group. Treatment with Oasis® also increased the time to dressing change (5.2 days vs. 2.1 days) and the percentage of granulation tissue formed (65% vs. 38%).

Dermagraft®

Dermagraft® living cell therapy has been approved by the FDA for repair of diabetic foot ulcers. Use of Dermagraft® for venous ulcers is an off-label indication. Harding (2013) reported an open-label multicenter RCT that compared Dermagraft® plus compression therapy (n=186) with compression therapy alone (n=180). The trial had numerous inclusion and exclusion criteria that restricted the population to patients who had nonhealing ulcers with compression therapy but had the capacity to heal. ITT analysis revealed no significant difference between the two groups in the primary outcome measure, the proportion of patients with completely healed ulcers by 12 weeks (34% Dermagraft® vs. 31% control). Prespecified subgroup analysis revealed a significant improvement in the percentage of wounds healed for ulcers of 12 months or less in duration (52% vs. 37%) and for ulcers of 10 cm or less in diameter (47% vs. 39%). There were no significant differences in the secondary outcomes of time to healing, complete healing by week 24, and percent reduction in ulcer area.

PriMatrix®

Karr (2011) published a retrospective comparison of PriMatrix® (xenogenic ADM) and Apligraf® in 28 venous stasis ulcers. The first 14 venous stasis ulcers matching the inclusion and exclusion criteria for each graft were compared. Criteria were venous stasis ulcers of four weeks in duration, at least 1 cm² in diameter, and to a depth of subcutaneous tissue, with healthy tissue at the ulcer edge, adequate arterial perfusion to heal, and ability to tolerate compression therapy. The time to complete healing for PriMatrix® was 32 days with 1.3 applications compared with 63 days with 1.7 applications for Apligraf®. Although promising, additional study with a larger number of subjects is needed to assess the effect of PriMatrix® treatment in compared with current SOC.

DermACELL®

Cazzell (2019) published an RCT on DermACELL® ADM for venous leg ulcers in 18 patients. This was part of a larger study of the acellular dermal matrix for chronic wounds of the lower extremity in 202 patients; the component on diabetic lower extremity ulcers was previously reported by Cazzell (2017) and is described above. When including patients who required more than one application of the ADM, the percent of wounds closed at 24 weeks was...
29.4% with DermACELL® and 33.3% with SOC, suggesting no benefit DermACELL® for the treatment of venous ulcers in this small substudy.

Theraskin® Versus Standard of Care

In the propensity matched study by Gurtner (2020) described above, Theraskin® did not improve the healing rate of venous ulcers (66.1%) compared to SOC (70.1%).[58]

DEEP DERMAL BURNS

Epicel®

One case series from 2000 has described the treatment of 30 severely burned patients with Epicel®.[77] The cultured epithelial autografts were applied to a mean of 37% of total body surface area (TBSA). Epicel® achieved permanent coverage of a mean of 26% of TBSA, an area similar to that covered by conventional autografts (mean 25%). Survival was 90% in these severely burned patients.

Integra® Dermal Regeneration Template

A 2013 study compared Integra® with split-thickness skin graft and with viscose cellulose sponge (Cellonex), using three, 10 x 5 cm test sites on each of 10 burn patients.[78] The surrounding burn area was covered with meshed autograft. Biopsies were taken from each site on days 3, 7, 14, and 21, and at months 3 and 12. The tissue samples were stained and examined for markers of inflammation and proliferation. The Vancouver Scar Scale was used to assess scars. At 12-month follow-up, the three methods resulted in similar clinical appearance, along with similar histologic and immunohistochemical findings.

Branski (2007) reported on a randomized trial that compared Integra® with a standard autograft-allograft technique in 20 children with an average burn size of 73% TBSA (71% full-thickness burns).[79] Once vascularized (about 14-21 days), the Silastic epidermis was stripped and replaced with thin (0.05-0.13 mm) epidermal autograft. There were no significant differences between the Integra® group and controls in burn size (70% vs. 74% TBSA), mortality (40% vs. 30%), and hospital length of stay (41 vs. 39 days), all respectively. Long-term follow-up revealed a significant increase in bone mineral content and density (24 months) and improved scarring in terms of height, thickness, vascularity, and pigmentation (at 12 months and 18-24 months) in the Integra® group. No differences were observed between groups in the time to first reconstructive procedure, cumulative reconstructive procedures required during two years, and cumulative operating room time required for these procedures. The authors concluded that Integra® can be used for immediate wound coverage in children with severe burns without the associated risks of cadaver skin.

Heimbach (2003) reported on a multicenter (13 U.S. burn care facilities) post-approval study involving 222 burn injury patients (36.5% TBSA, range 1%-95%) who were treated with Integra® Dermal Regeneration Template.[80] Within two to three weeks, the dermal layer regenerated, and a thin epidermal autograft was placed over the wound. The incidence of infection was 16.3%. Mean take rate (absence of graft failure) of Integra® was 76.2%; the median take rate was 98%. The mean take rate of epidermal autograft placed over Integra® was 87.7%; the median take rate was 95%.

Hicks (2019) conducted a systematic review of Integra® dermal regeneration template for the treatment of acute full thickness burns and burn reconstruction.[81] A total of 72 studies with
1,084 patients (four RCTs, four comparative studies, five cohort studies, two case control studies, 24 case series, and 33 case reports) were included in the review. The majority of patients (74%) were treated with Integra® for acute burns, and the remainder (26%) for burn reconstruction. The take of the skin substitute was 86% (range 0-100%) for acute burn injuries and 95% (range 0-100%) for reconstruction. The take of the split-thickness skin graft over the template was 90% for acute burn injuries and 93% for reconstruction. There was high variability in reporting of outcomes, but studies generally supported satisfactory cosmetic results in patients who have insufficient autograft and improvement in range of motion in patients who were treated with Integra® for burn reconstruction. There was an overall complication rate of 13%, primarily due to infection, graft loss, hematoma formation, and contracture.

An infection rate of 18% was noted in a systematic review of complication rates in 10 studies that used Integra® dermal regeneration template for burns.[82]

TransCyte®

TransCyte® is no longer commercially available.

Earlier studies included a report by Lukish (2001) that found improved healing in 20 consecutive cases of pediatric burns greater than 7% TBSA that underwent wound closure using TransCyte® compared to the previous 20 consecutive burn cases greater than 7% TBSA that received standard therapy.[83] Amani (2006) found significant improvement in healing in 110 consecutive patients who had deep partial-thickness burns treated with TransCyte® as compared to results from the American Burn Association Patient Registry for similar burns.[84]

DYSTROPHIC EPIDERMOLYSIS BULLOSA

OrCel® was approved under a humanitarian device exemption (HDE) for use in patients with dystrophic epidermolysis bullosa undergoing hand reconstruction surgery, to close and heal wounds created by the surgery, including those at donor sites. HDE status has been withdrawn for Dermagraft® for this indication.

Fivenson (2003) reported the off-label use of Apligraf® in five patients with recessive dystrophic epidermolysis bullosa who underwent syndactyly release.[85]

HUMAN AMNIOTIC MEMBRANE FOR OPHTHALMOLOGIC CONDITIONS

Sutured HAM transplant has been used for many years for the treatment of ophthalmic conditions. Many of these conditions are rare, leading to difficulty in conducting RCTs. The rarity, severity, and variability of the ophthalmic condition was taken into consideration in evaluating the evidence.

Liu (2019) conducted a systematic review of 17 studies (390 eyes) of amniotic membrane for corneal ulcers.[86] All but one of the studies was conducted outside of the U.S. There was one RCT with 30 patients, the remainder of the studies were prospective or retrospective case series. Corneal healing was obtained in 97% (95% CI 0.94 to 0.99, p=0.089) of patients evaluated. In the 12 studies (222 eyes) that reported on vision, the vision improvement rate was improved in 113 eyes (53%, 95% CI 0.42 to 0.65, p<0.001).

Khokhar (2005) reported on an RCT of 30 patients (30 eyes) with refractory neurotrophic corneal ulcers who were randomized to HAM transplantation (n=15) or conventional treatment
with tarsorrhaphy or bandage contact lens. At the three-month follow-up, 11 (73%) of 15 patients in the HAM group showed complete epithelialization compared with 10 (67%) of 15 patients in the conventional group. This difference was not significantly significant.

Suri (2013) published a series of 35 eyes of 33 patients who were treated with the self-retained Prokera® HAM for a variety of ocular surface disorders. Nine of the eyes had non-healing corneal ulcers. Complete or partial success was seen in two of nine (22%) patients with this indication. This study also reported on 11 eyes of 11 patients with neurotrophic keratopathy that had not responded to conventional treatment. The mean duration of treatment prior to Prokera® insertion was 51 days. Five of the 11 patients (45.5%) were considered to have had a successful outcome.

Dos Santos Paris (2013) published an RCT that compared fresh HAM with stromal puncture for the management of pain in patients with bullous keratopathy. Forty patients with pain from bullous keratopathy who were either waiting for a corneal transplant or had no potential for sight in the affected eye were randomized to the two treatments. Symptoms had been present for approximately two years. HAM resulted in a more regular epithelial surface at up to 180 days follow-up, but there was no difference between the treatments related to the presence of bullae or the severity or duration of pain. Because of the similar effects on pain, the authors recommended initial use of the simpler stromal puncture procedure, with use of HAM only if the pain did not resolve.

John (2017) reported on an RCT with 20 patients with moderate-to-severe dry eye disease who were treated with Prokera® c-HAM or maximal conventional treatment. The c-HAM was applied for an average of 3.4 days (range 3-5 days), while the control group continued treatment with artificial tears, cyclosporine A, serum tears, antibiotics, steroids, and nonsteroidal anti-inflammatory medications. The primary outcome was an increase in corneal nerve density. Signs and symptoms of dry eye disease improved at both one-month and three-month follow-ups in the c-HAM group but not in the conventional treatment group. For example, pain scores decreased from 7.1 at baseline to 2.2 at one month and 1.0 at three months in the c-HAM group. In vivo confocal microscopy, reviewed by masked readers, showed a significant increase in corneal nerve density in the study group at three months, with no change in nerve density in the controls. Corneal sensitivity was similarly increased in the c-HAM group but not in controls.

The DRy Eye Amniotic Membrane (DREAM) study, reported by McDonald (2018), was a retrospective series of 84 patients (97 eyes) with severe dry eye despite maximal medical therapy who were treated with Prokera® self-retained c-HAM. A majority of patients (86%) had superficial punctate keratitis. Other patients had filamentary keratitis (13%), exposure keratitis (19%), neurotrophic keratitis (2%), and corneal epithelial defect (7%). Treatment with Prokera® for a mean of 5.4 days (range 2-11) resulted in an improved ocular surface and reduction in the DEWS score from 3.25 at baseline to 1.44 at one week, 1.45 at one month and 1.47 at three months (p=0.001). Ten percent of eyes required repeated treatment. There was no significant difference in the number of topical medications following c-HAM treatment.

MISCELLANEOUS

Punch Biopsy Wounds

Baldursson (2015) reported a double-blinded RCT with 81 patients (162 punch biopsy wounds) that compared Kerecis™ Omega3 Wound (derived from fish skin) with Oasis® SIS ECM
The primary outcome (the percentage of wounds healed at 28 days) was similar for the fish skin ADM (95%) and the porcine SIS ECM (96.3%). The rate of healing was faster with Kerecis™ Omega3 (p=0.041). At 21 days, 72.5% of the fish skin ADM group had healed compared with 56% of the porcine SIS ECM group.

A similar RCT by Kirsner (2020) included 85 patients and compared the Kerecis™ Omega3 to a dehydrated human amnion/chorion membrane product.[93] This study also reported faster healing in the Kerecis™ Omega3 group (hazard ratio 2.37, 95% CI 1.75 to 3.21, p=0.0014). Interpretation of these studies is limited because they did not include an accepted control condition for this indication.

Split-Thickness Donor Sites

There is limited evidence to support the efficacy of OrCel® compared with SOC for the treatment of split-thickness donor sites in burn patients. Still (2003) (examined the safety and efficacy of bilayered OrCel® to facilitate wound closure of split-thickness donor sites in 82 severely burned patients.[94] Each patient had two designated donor sites that were randomized to a single treatment of OrCel® or standard dressing (Biobrane-L). The healing time for OrCel® sites was significantly shorter than for sites treated with a standard dressing, enabling earlier recropping. OrCel® sites also exhibited a nonsignificant trend for reduced scarring. Additional studies are needed to evaluate the effect of this product on health outcomes.

Pressure Ulcers

Brown-Etris (2019) reported an RCT of 130 patients with stage 3 or stage 4 pressure ulcers who were treated with Oasis® Wound Matrix (extracellular collagen matrix derived from porcine small intestinal submucosa) plus SOC or SOC alone.[95] At 12 weeks, the proportion of wounds healed in the collagen matrix group was 40% compared to 29% in the SOC group. This was not statistically significant (p=0.111). There was a statistical difference in the proportion of patients who achieved 90% wound healing (55% vs. 38% p=0.037), but complete wound healing is the preferred and most reliable measure. It is possible that longer follow-up may have identified a significant improvement in the percent of wounds healed. The study did include six-month follow-up, but there was high loss to follow-up and an insufficient number of patients at this time point for statistical comparison.

In the propensity matched study by Gurtner (2020) described above, Theraskin® improved the healing rate of pressure ulcers by 20% (66.7% vs 46.8%).[58]

Plantar Fasciitis

A 2016 network meta-analysis of 22 RCTs (total n=1,216 patients) compared injection therapies for plantar fasciitis.[96] In addition to c-HAM and micronized d-HAM/chorionic membrane, treatments included corticosteroids, botulinum toxin type A, autologous whole blood, platelet-rich plasma, nonsteroidal anti-inflammatory drugs, dry needling, dextrose prolotherapy, and polydeoxyribonucleotide. Placebo arms included normal saline, local anesthetic, sham dry needling, and tibial nerve block. Analysis indicated d-HAM had the highest probability for improvement in pain and composite outcomes in the short-term, however, this finding was based only on a single RCT. Outcomes at two to six months (seven RCTs) favored botulinum toxin for pain and patient recovery plan for composite outcomes.
An RCT by Cazzell (2018) enrolled 145 patients and reported three-month follow-up.[97] In this trial, amniotic membrane injection led to greater improvements in the Visual Analog Scale (VAS) for pain and the Foot Functional Index between baseline and three months compared to controls. VAS at three months had decreased to 17.1 in the AmnioFix® group compared to 38.8 in the placebo control group, which would be considered a clinically significant difference. The major limitation of the study is the short-term follow-up.

**Osteoarthritis**

In 2016, a feasibility study (n=6) was reported of ReNu™ cryopreserved human amniotic membrane (c-HAM) suspension with amniotic fluid-derived cells for the treatment of knee osteoarthritis.[98] A single intra-articular injection of the suspension was used, with follow-up at one and two weeks and at 3, 6, and 12 months posttreatment. Outcomes included the Knee Injury and Osteoarthritis Outcome Score, International Knee Documentation Committee scale, and a numeric pain scale. Statistical analyses were not performed for this small sample. No adverse events, aside from a transient increase in pain, were noted.

**Repair Following Mohs Micrographic Surgery**

Lu (2022) published a systematic review of skin substitutes for management of Mohs micrographic surgery wounds.[99] Of the 40 studies that met inclusion criteria, there were 23 case series, 14 case reports, two cohort studies, and one RCT. The most frequently used substitutes were porcine collagen (57.5%), bovine collagen (11.3%), Integra (7.7%), hyaluronic acid-derived products (6.2%), amnion/chorion-derived products (5.8%), and allogeneic epidermal-dermal composite grafts (5.8%). Follow-up in these studies ranged from one week to 21 months. The authors noted a lack of high-quality evidence and a need for blinded RCTs comparing the performance of skin substitutes with traditional methods.

Toman (2022) conducted an observational study that compared repair using a dehydrated human amnion/chorion membrane product (EpiFix®) with surgical repair using autologous tissue in patients who underwent same-day repair following Mohs microsurgery for removal of skin cancer on the face, head, or neck.[100] Propensity-score matching using retrospective data from medical records was used to identify 143 matched pairs. The primary endpoint was the incidence of postoperative morbidity, including the rate of infection, bleeding/hematoma, dehiscence, surgical reintervention, or development of a nonhealing wound. Postoperative cosmetic outcomes were assessed at nine months or later and included documentation of suboptimal scarring, scar revision, treatment, and patient satisfaction. A greater proportion of patients who received EpiFix® repair experienced zero complications (97.9% vs. 71.3%, p<0.0001, relative risk 13.67, 95% CI 4.33 to 43.12). Placental allograft reconstructions developed less infection (p=0.004) and were less likely to experience poor scar cosmesis (p<0.0001). Confidence in these findings is limited, however, by the study's retrospective design and potential for bias due to missing data. Additionally, the study's relevance is limited due to a lack of diversity in the study population and no comparison to non-surgical treatment options.

**Other Indications**

In addition to indications previously reviewed, off-label uses of bioengineered skin substitutes have included inflammatory ulcers (e.g., pyoderma gangrenosum, vasculitis), scleroderma digital ulcers, post-keloid removal wounds, genetic conditions, and variety of other conditions.[101] Products that have been FDA-approved or -cleared for one indication (e.g.,
lower-extremity ulcers) have also been used off-label in place of other FDA-approved or -cleared products (e.g., for burns).[102] No controlled trials were identified for these indications.

PRACTICE GUIDELINE SUMMARY

Wound Healing Society

In 2016, the Wound Healing Society updated their guidelines on diabetic foot ulcer treatment.[103] The Society concluded that there was level 1 evidence that cellular and acellular skin equivalents improve diabetic foot ulcer healing, noting that, “healthy living skin cells assist in healing DFUs [diabetic foot ulcers] by releasing therapeutic amounts of growth factors, cytokines, and other proteins that stimulate the wound bed.” References from two randomized controlled trials on dehydrated amniotic membrane were included with references on living and acellular bioengineered skin substitutes.

Society for Vascular Surgery

In 2016, the Society for Vascular Surgery in collaboration with the American Podiatric Medical Association and the Society for Vascular Medicine made the following recommendation:[104] "For DFUs [diabetic foot ulcers] that fail to demonstrate improvement (>50% wound area reduction) after a minimum of 4 weeks of standard wound therapy, we recommend adjunctive wound therapy options. These include negative pressure therapy, biologics (platelet-derived growth factor [PDGF], living cellular therapy, extracellular matrix products, amnionic membrane products), and hyperbaric oxygen therapy. Choice of adjuvant therapy is based on clinical findings, availability of therapy, and cost-effectiveness; there is no recommendation on ordering of therapy choice."

SUMMARY

BREAST RECONSTRUCTION

There is enough evidence to show that some allogeneic acellular dermal matrix (ADM) products can improve health outcomes for individuals who are undergoing medically necessary breast reconstruction. A systematic review found no difference in overall complication rates with ADM allograft compared with standard procedures for breast reconstruction. Reconstructions with ADM have been reported to have higher seroma, infection, and necrosis rates than reconstructions without ADM, however, capsular contracture and malposition of implants may be reduced. Therefore, the use of AlloDerm®, AlloMend®, Cortiva® (AlloMax™), DermACELL®, DermaMatrix™, FlexHD®, FlexHD® Pliable™, or GraftJacket® may be considered medically necessary for breast reconstruction.

There is not enough evidence to show that other amniotic products or bioengineered skin or soft tissue substitutes can improve health outcomes for patients undergoing breast reconstruction. Therefore, the use of products other than AlloDerm®, AlloMend®, Cortiva® (AlloMax™), DermACELL®, DermaMatrix™, FlexHD®, FlexHD® Pliable™, or GraftJacket® is considered investigational for this indication.

DIABETIC LOWER-EXTREMITY ULCERS

There is enough research to show that certain skin substitutes can improve health outcomes for certain patients who have diabetic lower-extremity ulcers that have not responded to...
conventional treatment. Randomized controlled trials have demonstrated that these products may improve ulcer healing compared with the standard of care. In addition, clinical practice guidelines for diabetic wound care recommend the use of skin substitutes in some cases. Therefore, the use of Affinity®, AlloPatch®, AmnioBand® Membrane, Apligraf®, Biovance®, Dermagraft®, EpiCord®, EpiFix®, Grafix®, Integra® Omnigraft™, or Integra® Flowable Wound Matrix may be considered medically necessary for the treatment of non-healing diabetic lower-extremity ulcers that have not responded to a 1-month period of conventional ulcer therapy. Treatment of diabetic lower-extremity ulcers with skin substitutes prior to 1-month of conventional ulcer therapy is considered not medically necessary.

There is not enough evidence to show that other amniotic products or bioengineered skin or soft tissue substitutes can improve health outcomes for patients with nonhealing diabetic lower-extremity ulcers. Therefore, the use of products other than Affinity®, AlloPatch®, AmnioBand® Membrane, Apligraf®, Biovance®, Dermagraft®, EpiCord®, EpiFix®, Grafix®, Integra® Omnigraft™, or Integra® Flowable Wound Matrix is considered investigational for this indication.

LOWER-EXTREMITY ULCERS DUE TO VENOUS INSUFFICIENCY

There is enough evidence to show that the use of Apligraf® or Oasis® Wound Matrix can improve health outcomes for individuals who have nonhealing lower-extremity ulcers due to venous insufficiency. Randomized controlled trials have demonstrated that these products can improve the healing of these wounds compared with the standard of care. Therefore, Apligraf® or Oasis® Wound Matrix may be considered medically necessary for the treatment of ulcers that have not responded to 1-month period of conventional ulcer therapy. Treatment of lower-extremity ulcers due to venous insufficiency with skin substitutes prior to 1-month of conventional ulcer therapy is considered not medically necessary.

There is not enough evidence to show that other amniotic products or bioengineered skin or soft tissue substitutes can improve health outcomes for patients with lower-extremity ulcers due to venous insufficiency. Therefore, the use of products other than Apligraf® or Oasis® Wound Matrix is considered investigational for this indication.

DYSTROPHIC EPIDERMOLOYSIS BULLOSA

OrCel® was approved by the FDA under a humanitarian drug exemption for use in patients with dystrophic epidermolysis bullosa undergoing hand reconstruction surgery, to close and heal wounds created by the surgery, including those at donor sites. Therefore, OrCel® may be considered medically necessary for this indication.

There is not enough evidence to show that other amniotic products or bioengineered skin or soft tissue substitutes can improve health outcomes for patients with dystrophic epidermolysis bullosa, and only OrCel® has received a humanitarian drug exemption for this condition. Therefore, the use of products other than OrCel® is considered investigational for dystrophic epidermolysis bullosa.

DEEP DERMAL BURNS

There is enough evidence to show that Epicel® and Integra® Dermal Regeneration Template may improve health outcomes for individuals who have deep dermal burns. Epicel® has received FDA approval under a humanitarian device exemption for the
treatment of deep dermal or full-thickness burns comprising a total body surface area of 30% or more. Comparative studies have demonstrated improved outcomes for Integra® Dermal Regeneration Template for the treatment of burns. Therefore, Epitel® or Integra® Dermal Regeneration Template may be considered medically necessary for the treatment of second- or third-degree burns.

There is not enough evidence to show that products other than Epitel® or Integra® Dermal Regeneration Template can improve health outcomes for patients with second- or third-degree burns. Therefore, the use of other amniotic products or bioengineered skin substitutes is considered investigational for this indication.

OPHTHALMIC INDICATIONS

There is limited evidence to show that human amniotic membrane products can improve health outcomes for patients with ophthalmologic indications, however these disorders are rare, and randomized controlled trials are unlikely. The use of certain amniotic products has become standard of care for the treatment of corneal injuries or as a component of corneal or conjunctival surgical repair, and therefore human amniotic membranes for ocular use, including but not limited to Prokera®, AmbioDisk™, or AmnioGraft® may be considered medically necessary for these indications.

SURGICAL REPAIR OF HERNIAS OR PARASTOMAL REINFORCEMENT

There is enough evidence to show that bioengineered skin substitutes do not improve health outcomes for individuals who are undergoing surgical repair of hernias or parastomal reinforcement. Several comparative studies including RCTs have shown no difference in outcomes between tissue-engineered skin substitutes and either standard synthetic mesh or no reinforcement. Therefore, the use of bioengineered skin substitutes is considered not medically necessary for these indications.

TENDON REPAIR

There is not enough research to show that skin substitutes or amniotic products can improve health outcomes for individuals who are undergoing tendon repair. A single trial found improved outcomes with the GraftJacket® allograft for rotator cuff repair. Although these results were positive, additional study with a larger number of patients is needed to evaluate the consistency of the effect. Therefore, the use of skin substitutes or amniotic products for tendon repair is considered investigational.

OTHER INDICATIONS

There is not enough research to show that skin substitutes or amniotic products can improve health outcomes for patients with disorders other than those listed in the medical necessity criteria. Off-label uses of bioengineered skin substitutes have included inflammatory ulcers, scleroderma digital ulcers, post-keloid removal wounds, genetic conditions, and variety of other conditions, however there is a lack of controlled trials for these uses. Therefore, the use of skin substitutes or amniotic products for other indications is considered investigational.

REFERENCES


64. JC Karr. Retrospective comparison of diabetic foot ulcer and venous stasis ulcer healing outcome between a dermal repair scaffold (PriMatrix) and a bilayered living cell therapy (Apligraf). *Advances in skin & wound care*. 2011;24(3):119-25. PMID: 21326023


### CODES

**NOTE:** While codes for skin substitute application (15271-15278, 15777) do not have pre-authorization requirements, they may be denied when used for the application of a product that does not meet medical necessity criteria.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>15272</td>
<td>Application of skin substitute graft to trunk, arms, legs, total wound surface area up to 100 sq cm; first 25 sq cm or less wound surface area</td>
<td></td>
</tr>
<tr>
<td>15273</td>
<td>Application of skin substitute graft to trunk, arms, legs, total wound surface area greater than or equal to 100 sq cm; first 100 sq cm wound surface area, or 1% of body area of infants and children</td>
<td></td>
</tr>
<tr>
<td>15274</td>
<td>; each additional 100 sq cm wound surface area, or part thereof, or each additional 1% of body area of infants and children, or part thereof (List separately in addition to code for primary procedure)</td>
<td></td>
</tr>
<tr>
<td>15275</td>
<td>Application of skin substitute graft to face, scalp, eyelids, mouth, neck, ears, orbits, genitalia, hands, feet, and/or multiple digits, total wound surface area up to 100 sq cm; first 25 sq cm or less wound surface area</td>
<td></td>
</tr>
</tbody>
</table>

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>15276</td>
<td></td>
<td>; total wound surface area up to 100 sq cm; each additional 25 sq cm wound surface area, or part thereof (List separately in addition to code for primary procedure)</td>
</tr>
<tr>
<td>15277</td>
<td></td>
<td>Application of skin substitute graft to face, scalp, eyelids, mouth, neck, ears, orbits, genitalia, hands, feet, and/or multiple digits, total wound surface area greater than or equal to 100 sq cm; first 100 sq cm wound surface area, or 1% of body area of infants and children</td>
</tr>
<tr>
<td>15278</td>
<td></td>
<td>; each additional 100 sq cm wound surface area, or part thereof, or each additional 1% of body area of infants and children, or part thereof (List separately in addition to code for primary procedure)</td>
</tr>
<tr>
<td>15777</td>
<td></td>
<td>Implantation of biologic implant (eg, acellular dermal matrix) for soft tissue reinforcement (ie, breast, trunk) (List separately in addition to code for primary procedure)</td>
</tr>
</tbody>
</table>

**HCPCS**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2001</td>
<td>Innovamatrix ac, per square centimeter</td>
</tr>
<tr>
<td>A2002</td>
<td>Mirragen advanced wound matrix, per square centimeter</td>
</tr>
<tr>
<td>A2004</td>
<td>Xcellistem, 1 mg</td>
</tr>
<tr>
<td>A2005</td>
<td>Microlyte matrix, per square centimeter</td>
</tr>
<tr>
<td>A2006</td>
<td>Novosorb synpath dermal matrix, per square centimeter</td>
</tr>
<tr>
<td>A2007</td>
<td>Restrata, per square centimeter</td>
</tr>
<tr>
<td>A2008</td>
<td>Theragenesis, per square centimeter</td>
</tr>
<tr>
<td>A2009</td>
<td>Symphony, per square centimeter</td>
</tr>
<tr>
<td>A2010</td>
<td>Apis, per square centimeter</td>
</tr>
<tr>
<td>A2011</td>
<td>Supra sdm, per square centimeter</td>
</tr>
<tr>
<td>A2012</td>
<td>Suprathel, per square centimeter</td>
</tr>
<tr>
<td>A2013</td>
<td>Innovamatrix fs, per square centimeter</td>
</tr>
<tr>
<td>A4100</td>
<td>Skin substitute, fda cleared as a device, not otherwise specified</td>
</tr>
<tr>
<td>A6460</td>
<td>Synthetic resorbable wound dressing, sterile, pad size 16 sq in or less, without adhesive border, each dressing</td>
</tr>
<tr>
<td>A6461</td>
<td>Synthetic resorbable wound dressing, sterile, pad size more than 16 sq in but less than or equal to 48 sq in, without adhesive border, each dressing</td>
</tr>
<tr>
<td>C1849</td>
<td>Skin substitute, synthetic, resorbable, per sq cm</td>
</tr>
<tr>
<td>C9356</td>
<td>Tendon, porous matrix of cross-linked collagen and glycosaminoglycan matrix (TenoGlide Tendon Protector Sheet), per sq cm</td>
</tr>
<tr>
<td>C9358</td>
<td>Dermal substitute, native, non-denatured collagen, fetal bovine origin (SurgiMend Collagen Matrix), per 0.5 square centimeters</td>
</tr>
<tr>
<td>C9360</td>
<td>Dermal substitute, native, non-denatured collagen, neonatal bovine origin (SurgiMend Collagen Matrix), per 0.5 square centimeters</td>
</tr>
<tr>
<td>C9363</td>
<td>Skin substitute, Integra Meshed Bilayer Wound Matrix, per square centimeter</td>
</tr>
<tr>
<td>C9364</td>
<td>Porcine implant, Permacol, per square centimeter</td>
</tr>
<tr>
<td>Q4100</td>
<td>Skin substitute, not otherwise specified</td>
</tr>
<tr>
<td>Q4101</td>
<td>Apligraf, per square centimeter</td>
</tr>
<tr>
<td>Q4102</td>
<td>Oasis Wound Matrix, per square centimeter</td>
</tr>
<tr>
<td>Q4103</td>
<td>Oasis Burn Matrix, per square centimeter</td>
</tr>
<tr>
<td>Q4104</td>
<td>Integra Bilayer Matrix Wound Dressing (BMWD), per square centimeter</td>
</tr>
<tr>
<td>Q4105</td>
<td>Integra Dermal Regeneration Template (DRT) or Integra Omnigraft dermal regeneration matrix, per square centimeter</td>
</tr>
<tr>
<td>Q4106</td>
<td>Dermagraft, per square centimeter</td>
</tr>
<tr>
<td>Q4107</td>
<td>Graftjacket, per square centimeter</td>
</tr>
<tr>
<td>Q4108</td>
<td>Integra Matrix, per square centimeter</td>
</tr>
<tr>
<td>Q4110</td>
<td>PriMatrix, per square centimeter</td>
</tr>
<tr>
<td>Q4111</td>
<td>GammaGraft, per square centimeter</td>
</tr>
<tr>
<td>Codes</td>
<td>Number</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>Q4112</td>
<td></td>
</tr>
<tr>
<td>Q4113</td>
<td></td>
</tr>
<tr>
<td>Q4114</td>
<td></td>
</tr>
<tr>
<td>Q4115</td>
<td></td>
</tr>
<tr>
<td>Q4116</td>
<td></td>
</tr>
<tr>
<td>Q4117</td>
<td></td>
</tr>
<tr>
<td>Q4118</td>
<td></td>
</tr>
<tr>
<td>Q4121</td>
<td></td>
</tr>
<tr>
<td>Q4122</td>
<td></td>
</tr>
<tr>
<td>Q4123</td>
<td></td>
</tr>
<tr>
<td>Q4124</td>
<td></td>
</tr>
<tr>
<td>Q4125</td>
<td></td>
</tr>
<tr>
<td>Q4126</td>
<td></td>
</tr>
<tr>
<td>Q4127</td>
<td></td>
</tr>
<tr>
<td>Q4128</td>
<td></td>
</tr>
<tr>
<td>Q4130</td>
<td></td>
</tr>
<tr>
<td>Q4132</td>
<td></td>
</tr>
<tr>
<td>Q4133</td>
<td></td>
</tr>
<tr>
<td>Q4134</td>
<td></td>
</tr>
<tr>
<td>Q4135</td>
<td></td>
</tr>
<tr>
<td>Q4136</td>
<td></td>
</tr>
<tr>
<td>Q4137</td>
<td></td>
</tr>
<tr>
<td>Q4138</td>
<td></td>
</tr>
<tr>
<td>Q4139</td>
<td></td>
</tr>
<tr>
<td>Q4140</td>
<td></td>
</tr>
<tr>
<td>Q4141</td>
<td></td>
</tr>
<tr>
<td>Q4142</td>
<td></td>
</tr>
<tr>
<td>Q4143</td>
<td></td>
</tr>
<tr>
<td>Q4145</td>
<td></td>
</tr>
<tr>
<td>Q4146</td>
<td></td>
</tr>
<tr>
<td>Q4147</td>
<td></td>
</tr>
<tr>
<td>Q4148</td>
<td></td>
</tr>
<tr>
<td>Q4149</td>
<td></td>
</tr>
<tr>
<td>Q4150</td>
<td></td>
</tr>
<tr>
<td>Q4151</td>
<td></td>
</tr>
<tr>
<td>Q4152</td>
<td></td>
</tr>
<tr>
<td>Q4153</td>
<td></td>
</tr>
<tr>
<td>Q4154</td>
<td></td>
</tr>
<tr>
<td>Q4155</td>
<td></td>
</tr>
<tr>
<td>Q4156</td>
<td></td>
</tr>
<tr>
<td>Q4157</td>
<td></td>
</tr>
<tr>
<td>Q4158</td>
<td></td>
</tr>
<tr>
<td>Q4159</td>
<td></td>
</tr>
<tr>
<td>Q4160</td>
<td></td>
</tr>
<tr>
<td>Q4161</td>
<td></td>
</tr>
</tbody>
</table>

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q4162</td>
<td>WoundEx Flow, BioSkin Flow, 0.5 cc</td>
<td></td>
</tr>
<tr>
<td>Q4163</td>
<td>WoundEx, BioSkin, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4164</td>
<td>Helicoll, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4165</td>
<td>Keramatrix, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4166</td>
<td>Cytal, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4167</td>
<td>Truskin, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4168</td>
<td>Amnioband, 1 mg</td>
<td></td>
</tr>
<tr>
<td>Q4169</td>
<td>Artacent wound, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4170</td>
<td>Cygnus, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4171</td>
<td>Interfyl, 1 mg</td>
<td></td>
</tr>
<tr>
<td>Q4172</td>
<td>Puraply or puraply am, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4173</td>
<td>Palingen or palingen xplus, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4174</td>
<td>Palingen or promatrix, 0.36 mg per 0.25 cc</td>
<td></td>
</tr>
<tr>
<td>Q4175</td>
<td>Miroderm, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4176</td>
<td>Neopatch, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4177</td>
<td>Floweramnioflno, 0.1 cc</td>
<td></td>
</tr>
<tr>
<td>Q4178</td>
<td>Floweramniopatch, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4179</td>
<td>Flowerderm, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4180</td>
<td>Revita, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4181</td>
<td>Amnio wound, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4182</td>
<td>Transcyte, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4183</td>
<td>Surgigraft, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4184</td>
<td>Cellesta or cellesta duo, per square centimeter (revised 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4185</td>
<td>Cellesta flowable amnion (25 mg per cc); per 0.5 cc</td>
<td></td>
</tr>
<tr>
<td>Q4186</td>
<td>Epifix, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4187</td>
<td>Epicord, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4188</td>
<td>Amnioarmor, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4189</td>
<td>Artacent ac, 1 mg</td>
<td></td>
</tr>
<tr>
<td>Q4190</td>
<td>Artacent ac, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4191</td>
<td>Restorigin, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4192</td>
<td>Restorigin, 1 cc</td>
<td></td>
</tr>
<tr>
<td>Q4193</td>
<td>Coll-e-derm, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4194</td>
<td>Novachor, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4195</td>
<td>Puraply, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4196</td>
<td>Puraply am, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4197</td>
<td>Puraply xt, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4198</td>
<td>Genesis amniotic membrane, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4199</td>
<td>Cygnus matrix, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4200</td>
<td>Skin te, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4201</td>
<td>Matrion, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4202</td>
<td>Keroxx (2.5g/cc), 1cc</td>
<td></td>
</tr>
<tr>
<td>Q4203</td>
<td>Derma-gide, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4204</td>
<td>Xwrap, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4205</td>
<td>Membrane graft or membrane wrap, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4206</td>
<td>Fluid flow or fluid GF, 1 cc (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4208</td>
<td>Novafix, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4209</td>
<td>Surgraft, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4210</td>
<td>Axolotl graft or axolotl dualgraft, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4211</td>
<td>Amnion bio or Axobiomembrane, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4212</td>
<td>Allogen, per cc (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4213</td>
<td>Ascent, 0.5 mg (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4214</td>
<td>Cellesta cord, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Codes</td>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>Q4215</td>
<td>Axolotl ambient or axolotl cryo, 0.1 mg (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4216</td>
<td>Artacent cord, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4217</td>
<td>Woundfix, BioWound, Woundfix Plus, BioWound Plus, Woundfix Xplus or BioWound Xplus, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4218</td>
<td>Surgicord, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4219</td>
<td>Surgigraft-dual, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4220</td>
<td>BellaCell HD or Surederm, per square centimeter (eff 10/01/19)</td>
<td></td>
</tr>
<tr>
<td>Q4221</td>
<td>Amniowrap2, per square centimeter (new eff 10/1/19)</td>
<td></td>
</tr>
<tr>
<td>Q4222</td>
<td>Progenamatrix, per square centimeter (eff 10/01/19)</td>
<td></td>
</tr>
<tr>
<td>Q4224</td>
<td>Human health factor 10 amniotic patch (hhf10-p), per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4225</td>
<td>Amniobind, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4226</td>
<td>MyOwn skin, includes harvesting and preparation procedures, per square centimeter (eff 10/01/19)</td>
<td></td>
</tr>
<tr>
<td>Q4227</td>
<td>Amniocore, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4228</td>
<td>BioNextPATCH, per square centimeter (new eff 7/1/20) (Deleted 10/01/2021)</td>
<td></td>
</tr>
<tr>
<td>Q4229</td>
<td>Cogenex amniotic membrane, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4230</td>
<td>Cogenex flowable amnion, per 0.5 cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4231</td>
<td>Corplex P, per cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4232</td>
<td>Corplex, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4233</td>
<td>Surfactor or Nudyn, per 0.5 cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4234</td>
<td>Xcellerate, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4235</td>
<td>Amniorepair or altiply, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4236</td>
<td>carePATCH, per square centimeter (new eff 7/1/20) (Deleted 10/01/2021)</td>
<td></td>
</tr>
<tr>
<td>Q4237</td>
<td>Cryo-cord, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4238</td>
<td>Derm-maxx, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4239</td>
<td>Amnio-maxx or Amnio-maxx lite, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4240</td>
<td>Corecyte, for topical use only, per 0.5 cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4241</td>
<td>Polycyte, for topical use only, per 0.5 cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4242</td>
<td>Amniocyte plus, per 0.5 cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4244</td>
<td>Procenta, per 200 mg (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4245</td>
<td>Amniotext, per cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4246</td>
<td>Coretext or Protext, per cc (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4247</td>
<td>Amniotext patch, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4248</td>
<td>Dermacyte Amniotic Membrane Allograft, per square centimeter (new eff 7/1/20)</td>
<td></td>
</tr>
<tr>
<td>Q4249</td>
<td>AMNIPLY, for topical use only, per sq cm</td>
<td></td>
</tr>
<tr>
<td>Q4250</td>
<td>AmnioAmp-MP, per sq cm</td>
<td></td>
</tr>
<tr>
<td>Q4251</td>
<td>Vim, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4252</td>
<td>Vendaje, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4253</td>
<td>Zenith amniotic membrane, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4254</td>
<td>Novafix DL, per sq cm</td>
<td></td>
</tr>
<tr>
<td>Q4255</td>
<td>REGUaRD, for topical use only, per sq cm</td>
<td></td>
</tr>
<tr>
<td>Q4256</td>
<td>Mlg-complete, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4257</td>
<td>Relese, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4258</td>
<td>Enverse, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4259</td>
<td>Celera dual layer or celera dual membrane, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4260</td>
<td>Signature apatch, per square centimeter</td>
<td></td>
</tr>
<tr>
<td>Q4261</td>
<td>Tag, per square centimeter</td>
<td></td>
</tr>
</tbody>
</table>

*Date of Origin: December 2018*
IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Digital health products are technologies, platforms, and systems that engage consumers for lifestyle, wellness, and health-related purposes. Digital therapeutic products differ from digital health products in that they are practitioner-prescribed software that delivers evidence-based therapeutic interventions to prevent, manage, or treat a medical disorder or disease.

MEDICAL POLICY CRITERIA

Note:

- Member contracts for covered services vary. Member contract language takes precedence over medical policy.
- This policy does not address:
  - Software that is used for the function or control of an FDA-cleared or approved stand-alone medical device (e.g., external insulin pump or pacemaker. See Cross References).
  - Applications operated by a health care practitioner for remote health monitoring.
  - Products for which coverage is required by a particular health plan under state or federal law (see Policy Guidelines).
The following general Criteria are applied to digital health products, including digital therapeutic products, not already addressed in any other Medical Policy (see Cross References).

I. The use of a digital health product in the treatment or prevention of any health condition is considered **medically necessary** when all of the following Criteria are met:

   A. The digital health product has been prescribed by a healthcare practitioner providing medical oversight; and

   B. The digital health product has been Food and Drug Administration (FDA) approved for the requested indication; and

   C. High-quality evidence demonstrates the digital health product improves clinically meaningful net health outcomes as much or more than an established alternative; and

   D. The improved net health outcome provided by the digital health product is attainable outside of investigational settings.

II. The use of a digital health product in the treatment or prevention of any health condition is considered **investigational** when Criterion I. is not met, including but not limited to general wellness and fitness applications, which are not considered therapeutic in nature (see Policy Guidelines).

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

**POLICY GUIDELINE**

When a digital health product is a software that delivers evidence-based therapeutic intervention to prevent, manage, or treat a medical disorder or disease, it may be considered as a digital therapeutic product. Digital therapeutics are distinguished from digital medicine and digital health products, such as mobile health products or wearable devices, in that clinical evidence of effectiveness and regulatory oversight are required for digital therapeutic products.[1, 2]

How Digital Therapeutics Differ From Digital Health[1]

<table>
<thead>
<tr>
<th>Definition</th>
<th>Digital Health</th>
<th>Digital Medicine</th>
<th>Digital Therapeutics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>Technologies, platforms, and systems that engage consumers for lifestyle, wellness, and health-related purposes; capture, store or transmit health data; and/or support life science and clinical operations.</td>
<td>Evidence-based software and/or hardware products that measure and/or intervene in the service of human health.</td>
<td>Delivers evidence-based therapeutic interventions to prevent, manage, or treat a medical disorder or disease.</td>
</tr>
</tbody>
</table>
Digital Health

<table>
<thead>
<tr>
<th>Clinical Evidence Required?</th>
<th>Digital Medicine</th>
<th>Digital Therapeutics</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Real-World Outcomes required?</th>
<th>Digital Medicine</th>
<th>Digital Therapeutics</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regulatory Oversight Required?</th>
<th>Digital Medicine</th>
<th>Digital Therapeutics</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>Varies</td>
<td>YES</td>
</tr>
</tbody>
</table>

**Digital Health**

Do not meet the regulatory definition of a medical device.

As required to support product claims of risk, efficacy, and intended use.

**Health Resources and Services Administration Women’s Preventive Services Guidelines (HRSA Guidelines)** ensure women’s access to the full range of FDA-approved contraceptive methods including, but not limited to barrier methods, hormonal methods, and implanted devices, as well as patient education and counseling, as prescribed by a health care provider.[3]

**LIST OF INFORMATION NEEDED FOR REVIEW**

**REQUIRED DOCUMENTATION:**

The information below must be submitted for review to determine whether policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

1. Name and manufacturer of the digital health product
2. Indication for which the digital health product is being prescribed
3. Relevant billing codes
4. Brief description of how the digital health product will improve net health outcomes for the patient
5. Medical records related to this request, including but not limited to history and physical exam, conventional testing and outcomes, and treatment provided, if any.

**CROSS REFERENCES**

1. Digital Health Products for Attention Deficit Hyperactivity Disorder, Medicine, Policy No. 175.01
2. Digital Health Products for Substance Use Disorders, Medicine, Policy No. 175.02
3. Insulin Infusion Pumps, Automated Insulin Delivery and Artificial Pancreas Device Systems, DME, Policy No. 77

**BACKGROUND**

In 2020 alone, more than 90,000 new digital health applications, an average of more than 250 apps per day, were introduced, bringing the total number of digital health applications available to consumers to over 350,000. Among these applications almost half (47%) focus on the management of a specific disease or health condition.[4] Examples of digital health products currently available include applications that purport to perform cognitive behavior therapy,
support weight loss goals, distinguish normal cardiac sinus rhythm and potentially dangerous arrhythmias, and to identify a suspicious mole. In addition, over 80% of adults in the United States own a smartphone.[6] The ability to utilize a personal mobile device, such as a smartphone, as a medical device has substantial potential to impact clinical care and to promote general health and wellness. However, despite the rapid influx of digital health products into the market, there remains no widely accepted framework for the evaluation of efficacy of these products. The use of a digital health product to prevent, manage, or treat a medical disorder or disease must be evaluated in the setting of existing evidence frameworks, as discussed below.

DEFINING DIGITAL THERAPEUTICS

The field of digital health is broad and rapidly changing. Digital therapeutic products fall under the umbrella term of digital health, however, digital therapeutic products are distinguished from digital medicine and digital health products in that clinical evidence of effectiveness and regulatory oversight are required for digital therapeutic products.[1]

The Academy of Managed Care Pharmacy (AMCP) published a review in 2020 which provides the definition of digital therapeutics as: software that delivers a clinical mechanism of action, either alone or in combination with other standard-of-care treatments to improve outcomes.[2]

This review also states, “Digital therapeutics represents one segment of digital health products and can be distinguished from other products, such as mobile health products or wearable devices, specifically by their demonstrated impact on measurable clinical outcomes.” The AMCP provides examples of products that do not meet the definition of a digital therapeutic product, as summarized in Table 1.

Table 1. Products Not Considered Digital Therapeutics[2]

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Mobile Health                 | The practice of medicine and public health supported by mobile devices       | • Clinician-facing: Apps that are for displaying, storing, analyzing, or transmitting patient-specific medical device data  
|                               |                                                                            | • Consumer-facing: Lifestyle, fitness tracking, nutrition, and medication adherence apps          |
| Health Information Technology | • Information technology applied to health and health care                   | • Electronic Medical records                                                                      |
|                               | • Supports health information management across computerized systems and the secure exchange of health information | • Electronic prescribing systems                                                               |
|                               |                                                                            | • Consumer health interface (e.g., MyChart)                                                       |
The World Health Organization has developed a classification system to define various types of digital health products. While this system categorizes the different ways digital and mobile technologies are used to support health system needs, it does not provide a definition of therapeutic digital health products, specifically.

The Digital Therapeutics Alliance (DTA) is a global non-profit trade association of industry leaders and stakeholders engaged in the evidence-driven advancement of digital therapeutics. The DTA provides the following definition of digital therapeutics:

Digital therapeutics deliver clinical-grade therapeutic interventions to patients. Digital therapeutic products may be used independently or in tandem with in-person or remote clinician-delivered therapy to optimize patient outcomes.

Digital therapeutics undergo clinical trials, collect real world outcomes, and are based on patient-centered core principles and product development best practices, including product design, usability, data security, and privacy standards.

Table 2. How Digital Therapeutics Differ From Digital Health.

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devices, sensors, wearables</td>
<td>• Devices that can be worn, attached on skin, or ingested to continuously and closely monitor an individual’s activities</td>
<td>• Wearable and wireless devices, (e.g., Fitbit, Apple Watch)</td>
</tr>
<tr>
<td></td>
<td>• Supported by embedded technology for data communication and sensors to interact with both internal and external objects and the environment</td>
<td>• Biometric sensors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Diagnostic products</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Proprietary algorithms that control the function of physical devices, such as insulin pumps</td>
</tr>
<tr>
<td>Telehealth</td>
<td>Provision of health care remotely</td>
<td>• Telemedicine, telehealth platforms</td>
</tr>
</tbody>
</table>

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
### REGULATORY STATUS

The US Food and Drug Administration (FDA) defines Software as a Medical Device (SaMD) as, “intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device.”[7]

The FDA notes the following regarding SaMD:

- SaMD is a medical device and includes in-vitro diagnostic (IVD) medical device
- SaMD is capable of running on general purpose (non-medical purpose) computing platforms
- “Without being part of” means software not necessary for a hardware medical device to achieve its intended medical purpose
- Software does not meet the definition of SaMD if its intended purpose is to drive a hardware medical device
- SaMD may be used in combination (e.g., as a module) with other products including medical devices
- SaMD may be interfaced with other medical devices, including hardware medical devices and other SaMD software, as well as general purpose software
- Mobile apps that meet the definition above are considered SaMD.

SaMD are reviewed by the FDA under existing 510(k) and DeNovo pathways established for the review of medical devices.

- A 510(k) is a premarket submission made to FDA to demonstrate that the device to be marketed is as safe and effective, that is, substantially equivalent, to a legally marketed device.[8]
- The De Novo process provides a pathway to classify novel medical devices for which general controls alone, or general and special controls, provide reasonable assurance of safety and effectiveness for the intended use, but for which there is no legally marketed predicate device. De Novo classification is a risk-based classification process. Devices that are classified into class I or class II through a De Novo classification request (De Novo request) may be marketed and used as predicates for future premarket notification [510(k)] submissions.[9]

The Digital Health Center of Excellence (DHCoE) is a resource under the Center for Devices and Radiological Health (CDRH) of the FDA. The DHCoE’s stated aim is to align and coordinate digital health efforts across the FDA and to serve as a resource for external stakeholders to promote digital health technologies.[10] The DHCoE focuses on a range of

---

Adapted from the Digital Therapeutics Alliance[1]
technologies including mobile health devices, SaMD, and wearables when used as a medical device. The DHCoE’s objectives include:

- Connect and build partnerships to accelerate digital health advancements; and
- Share knowledge to increase awareness and understanding, drive synergy, and advance best practices; and
- Innovate regulatory approaches to provide efficient and least burdensome oversight while meeting the FDA standards for safe and effective products; and

In 2017, the FDA announced the Software Pre-Cert Pilot Program as part of the Digital Health Innovation Action Plan “to develop a new regulatory paradigm that would focus first on the assessment of organizations that perform high-quality software design, testing, and monitoring.”[11] In January 2019, the FDA released a Test Plan for the Pre-Cert program as well as a Regulatory Framework for conducting the pilot program.[12, 13] The FDA currently is continuing testing the Pre-Cert program to determine if the results align with the results of the current approval pathways and satisfy the FDA’s established regulatory requirements for safety and effectiveness.

### PRACTICE GUIDELINE AND POSITION STATEMENT SUMMARY

At this time, no single framework has been adopted by medical or regulatory bodies for evaluation of digital therapeutic products. However, several organizations, both global and national, have initiated efforts to develop a framework for evaluation of digital health products, including those summarized below.

#### DIGITAL THERAPEUTICS ALLIANCE

The Digital Therapeutics Alliance (DTA), a global non-profit trade association of industry leaders and stakeholders, provides a summary of Industry Core Principals to which all digital therapeutic products should adhere “to demonstrate product integrity and ensure patient safety.”[14] These Principals include the statements that digital therapeutic products should:

- prevent, manage, or treat a medical disorder or disease; and
- produce a medical intervention that is driven by software; and
- publish trial results inclusive of clinically-meaningful outcomes in peer-reviewed journals; and
- be reviewed and cleared or certified by regulatory bodies as required to support product claims of risk, efficacy, and intended use; and
- make claims appropriate to clinical evaluation and regulatory status; and
- collect, analyze, and apply real world evidence and/or product performance data.

#### NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

The National Institute for Health and Care Excellence (NICE) published an Evidence Standards Framework for Digital Health Technologies (DHTs) in 2019.[15] The framework provides standards for evidence that should be available or developed for DHTs to demonstrate their value in the UK health and care system, specifically. Framework is broken into evidence tiers (minimum evidence standards) based on the functional classification of the technology. Per the definitions above, digital therapeutics would fit primarily in the highest evidence tier, 3b.
Minimum evidence for effectiveness standards for tier 3b DHTs include the following:

High quality intervention study (experimental or quasi-experimental design) showing improvements in relevant outcomes, such as:

- diagnostic accuracy
- patient-reported outcomes (preferably using validated tools) including symptom severity or quality of life
- other clinical measures of disease severity or disability
- healthy behaviors
- physiological measures
- user satisfaction and engagement.

Generic outcome measures may also be useful when reported alongside condition-specific outcomes. The comparator should be a care option that is reflective of the current care pathway, such as a commonly used active intervention.

**XCERTIA**

Xcertia, founded in December 2016 by representatives from groups including the American Medical Association, American Heart Association, Healthcare Information and Management Systems Society and digital health nonprofit DHX Group, published a guideline in 2019 that addressed “key areas of guidance to ensure mHealth apps deliver true value in a trusted environment.”[16]

These guidelines include the following statement regarding documentation of evidence for the app:

- The app’s public description should clearly state which type of research has been performed to validate its content. These can include the following levels of research:
  - I. Systematic review or meta-analysis of randomized control trials
  - II. Randomized control trial/s (number of trials if more than one)
  - III. Quasi-experimental study
  - IV. Case-control or cohort studies
  - V. Systematic reviews of descriptive and qualitative studies
  - VI. Single descriptive or qualitative study
  - VII. Expert medical or academic opinion
- If level of research performed on the opinion [sic] is based on expert or academic opinion (VII) or no study, the app’s public description should clearly state, “The effectiveness of the app has not been studied.”

**THE AMERICAN MEDICAL ASSOCIATION**

The American Medical Association Policy on Integration of Mobile Health Applications and Devices into Practice (2017) states; “Our AMA supports the establishment of coverage, payment and financial incentive mechanisms to support the use of mobile health applications (mHealth apps) and associated devices, trackers and sensors by patients, physicians and other providers that:

- support the establishment or continuation of a valid patient-physician relationship
• have a high-quality clinical evidence base to support their use in order to ensure mHealth app safety and effectiveness
• follow evidence-based practice guidelines, especially those developed and produced by national medical specialty societies and based on systematic reviews, to ensure patient safety, quality of care and positive health outcomes
• support care delivery that is patient-centered, promotes care coordination and facilitates team-based communication
• support data portability and interoperability in order to promote care coordination through medical home and accountable care models
• abide by state licensure laws and state medical practice laws and requirements in the state in which the patient receives services facilitated by the app
• require that physicians and other health practitioners delivering services through the app be licensed in the state where the patient receives services, or be providing these services as otherwise authorized by that state's medical board
• ensure that the delivery of any services via the app be consistent with state scope of practice laws.”[17]

REFERENCES

2. AMCP Partnership Forum: Digital Therapeutics-What Are They and Where Do They Fit in Pharmacy and Medical Benefits? J Manag Care Spec Pharm. 2020;26(5):674-81. PMID: 32175784


### CODES

**NOTE:** Not all digital health products will have a specific code. These are examples of codes that may be relevant.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0702T</td>
<td>Remote therapeutic monitoring of a standardized online digital cognitive behavioral therapy program ordered by a physician or other qualified health care professional; supply and technical support, per 30 days</td>
</tr>
<tr>
<td></td>
<td>0703T</td>
<td>Management services by physician or other qualified health care professional, per calendar month</td>
</tr>
<tr>
<td>99199</td>
<td></td>
<td>Unlisted special service, procedure or report [when specified as a digital health management software application]</td>
</tr>
<tr>
<td>HCPCS</td>
<td>A9291</td>
<td>Prescription digital behavioral therapy, FDA cleared, per course of treatment</td>
</tr>
<tr>
<td></td>
<td>A9300</td>
<td>Exercise Equipment [when specified as a digital health management software application]</td>
</tr>
<tr>
<td></td>
<td>E1399</td>
<td>Durable medical equipment, miscellaneous [when specified as a digital health management software application]</td>
</tr>
</tbody>
</table>

**Date of Origin:** September 2021
Digital Health Products for Attention Deficit Hyperactivity Disorder

Effective: April 1, 2022

Next Review: September 2022
Last Review: March 2022

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Digital health products are technologies, platforms, and systems that engage consumers for lifestyle, wellness, and health-related purposes. Digital therapeutic products differ from digital health products in that they are practitioner-prescribed software that delivers evidence-based therapeutic interventions to prevent, manage, or treat a medical disorder or disease. Digital therapeutic products have been proposed to supplement or replace established treatments for attention-deficit/hyperactivity disorder (ADHD).

MEDICAL POLICY CRITERIA

Note:

- Member contracts for covered services vary. Member contract language takes precedence over medical policy.
- This policy addresses the use of practitioner-prescribed software applications for therapeutic intervention.
- This policy does not address:
  - Software that is used for the function or control of an FDA-cleared or approved stand-alone medical device (e.g., external insulin pump or pacemaker).
Applications operated by a health care practitioner for remote health monitoring.

The use of a digital health product (including digital therapeutics) for the treatment of attention-deficit/hyperactivity disorder (ADHD), either as a stand-alone treatment or as an adjunct to standard treatment, is considered **investigational**, including but not limited to EndeavorRx® (AKL-T01).

*NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.*

### CROSS REFERENCES

1. Digital Health Products, Medicine, Policy No. 175

### BACKGROUND

#### ATTENTION-DEFICIT/HYPERACTIVITY DISORDER

Attention-deficit/hyperactivity disorder (ADHD) is a chronic condition characterized by core symptoms of hyperactivity, impulsivity, and inattention, which are considered excessive for the person’s age. Both the International Classification of Mental and Behavioral Disorders 10th edition (ICD-10) and the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) require that the symptoms are reported or observed in several settings and that the symptoms of ADHD affect psychological, social, and/or educational/occupational functioning. Prevalence estimates for ADHD vary from 7.2% to 15.5% of children.[1]

For children younger than 17 years of age, the DSM-5 requires at least six symptoms of hyperactivity-impulsivity or at least six symptoms of inattention. The combined type requires a minimum of six symptoms of hyperactivity-impulsivity plus at least six symptoms of inattention. The symptoms must 1) occur often, 2) be present in more than one setting, 3) persist for at least six months, 4) be present before 12 years of age, 5) impair function in academic, social, or occupational activities, and 6) be excessive for the developmental level of the child.

#### Treatment

Established treatments for ADHD in children include educational, environmental, psychological, and behavioral interventions, and medication. Almost two-thirds of children with ADHD take medication, and about one half receive behavioral treatment.[1]

- Educational intervention involves discussion with parents about symptoms and access to services, environmental modifications such as seating arrangements, changes to lighting and noise, reducing distractions, and the benefit of having movement breaks and teaching assistants at school.
- Parent-child behavioral therapy teaches parenting techniques within the principles of behavior therapy. The therapy programs typically last two to three months and includes rewarding positive behavior, identifying unintentional reinforcement of negative behaviors, limiting choices, and using calm discipline.
- Medication with stimulants, such as methylphenidate, are considered first-line therapy for ADHD in school-age children. However, adverse effects of stimulants may include sleep disturbance, decreased appetite, and weight changes. Combination therapy with
medication and behavioral interventions can improve both core ADHD symptoms and non-ADHD symptoms such as social skills and parent-child relations.

REGULATORY STATUS

In April 2020, EndeavorRx® (Akili Interactive Labs) received marketing clearance by the U.S. Food and Drug Administration (FDA) through the De Novo premarket review process (DEN200026).[2] EndeavorRx® is a prescription device that is indicated to “improve attention function as measured by computer-based testing in children ages 8 to 12 years old with primarily inattentive or combined type ADHD, who have a demonstrated attention issue. Patients who engage with EndeavorRx® demonstrate improvements in a digitally assessed measure Test of Variables of Attention (TOVA) of sustained and selective attention and may not display benefits in typical behavioral symptoms, such as hyperactivity.” EndeavorRx® is intended to be used as part of a therapeutic program that may include clinician-directed therapy, medication, and/or educational programs. EndeavorRx® was referred to as “ProjectEvo” and in later evaluations as “AKL-T01.”

EVIDENCE SUMMARY

Evidence reviews assess the clinical evidence to determine whether the use of a technology improves the net health outcome. Broadly defined, health outcomes are length of life, quality of life, and ability to function including benefits and harms. Every clinical condition has specific outcomes that are important to patients and to managing the course of that condition. Validated outcome measures are necessary to ascertain whether a condition improves or worsens; and whether the magnitude of that change is clinically significant. The net health outcome is a balance of benefits and harms.

To assess whether the evidence is sufficient to draw conclusions about the net health outcome of a technology, two domains are examined: the relevance and the quality and credibility. To be relevant, studies must represent one or more intended clinical use of the technology in the intended population and compare an effective and appropriate alternative at a comparable intensity. For some conditions, the alternative will be supportive care or surveillance. The quality and credibility of the evidence depend on study design and conduct, minimizing bias and confounding that can generate incorrect findings. The randomized controlled trial (RCT) is preferred to assess efficacy; however, in some circumstances, nonrandomized studies may be adequate. Randomized controlled trials are rarely large enough or long enough to capture less common adverse events and long-term effects. Other types of studies can be used for these purposes and to assess generalizability to broader clinical populations and settings of clinical practice.

DIGITAL THERAPIES FOR ATTENTION-DEFICIT/HYPERACTIVITY DISORDER

Clinical Context and Therapy Purpose

The purpose of digital therapeutic products is to provide a treatment option that is an alternative to or an improvement on existing therapies for patients with attention-deficit/hyperactivity disorder (ADHD).

Attention-deficit/hyperactivity disorder is a syndrome that can include hyperactivity, impulsivity, and/or inattention, which in turn can affect cognitive, academic, behavioral, emotional, and social functioning. The symptoms of the hyperactive-impulsive presentation typically occur
together and are characterized by the inability to sit still or inhibit behavior. The inattentive presentation is characterized by reduced ability to focus attention and reduced speed of cognitive processing, which is exhibited by difficulty with maintaining attention, lack of follow through and organization, distraction, and forgetfulness. The combined presentation includes symptoms of both the hyperactive-impulsive presentation and the inattentive presentation.

Treatment may include environmental adjustments, behavioral and psychological interventions, and medications. In some children, these treatments do not sufficiently address symptoms. In others, there may be resistance by the parents to treat children with medications, or there may be other barriers to obtaining established therapies. EndeavorRx® is proposed to address these barriers with improved access to care and minimal side effects. The therapy is based on research showing that impairments in attention and cognitive control are associated with lower activation of frontal, frontoparietal, and ventral attention networks. Previously, a game-like intervention was shown to improve cognitive performance and alter the electroencephalogram in the prefrontal cortex in older adults. The similarity between cognitive control in older adults and attention deficits in ADHD led to the development of EndeavorRx® for the treatment of ADHD in children.

ADHD-specific rating scales are described in Table 1.

### Table 1. ADHD Rating Scales

<table>
<thead>
<tr>
<th>Rating Scale</th>
<th>Description</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD Rating Scale (ADHD-RS-IV)⁴</td>
<td>The ADHD-RS-IV is an 18-item, clinician-administered questionnaire for which a parent respondent rates the frequency of occurrence of ADHD symptoms and behaviors as defined by criteria outlined for ADHD in the DSM-IV. Each item is scored on a 4-point scale ranging from 0 (rarely or never) to 3 (very often) with total scores ranging from 0 to 54. The 18 items are grouped into 2 subscales: hyperactivity/impulsivity and inattentiveness.</td>
<td>Each subscale produces a subscale score ranging from 0 to 27. A higher score indicates more severe ADHD symptoms and behaviors and a negative change in total score indicates improvement.</td>
</tr>
<tr>
<td>The Clinical Global Impression Scale – Improvement (CGI-I)⁵</td>
<td>The CGI-I is a clinician's comparison of the participant's overall clinical condition at follow-up to the overall clinical condition at baseline. It includes an assessment of the change from the initiation of treatment with a rating from 1 to 7.</td>
<td>The 7-point scale is: 1 = Very much improved, 2=Much improved, 3=Minimally improved, 4=No change, 5=Minimally worse, 6=Much worse, and 7=Very much worse. A score of 1, 2, or 3 would indicate overall improvement of ADHD severity.</td>
</tr>
<tr>
<td>Conners Comprehensive Behavior Rating Scales⁶</td>
<td>Parent and teacher forms are available in full (90-item, 59-item) and abbreviated (27-item, 28-item) versions.</td>
<td>Normative values are provided separately by gender and age.</td>
</tr>
<tr>
<td>The Vanderbilt Assessment Scales for parents and teachers⁷,⁸</td>
<td>The Vanderbilt Assessment Scales are based on DSM-IV scales. The scale for parents has 55 questions that rate symptoms and their impact on family and school. The teacher scale includes 43</td>
<td>Normative data and percentile ranks are provided for each subscale by grade and gender.</td>
</tr>
</tbody>
</table>
Follow-up after the treatment period (1 to 3 months), at six months, and annually for three years is of interest to monitor outcomes of the effect of EndeavorRx®.

**STUDY SELECTION CRITERIA**

Methodologically credible studies were selected using the following principles:

- To assess efficacy outcomes, comparative controlled prospective trials were sought, with a preference for randomized controlled trials (RCTs);
- In the absence of such trials, comparative observational studies were sought, with a preference for prospective studies.
- To assess long-term outcomes and adverse events, single-arm studies that capture longer periods of follow-up and/or larger populations were sought.
- Studies with duplicative or overlapping populations were excluded.

**REVIEW OF EVIDENCE**

**Randomized Controlled Trials**

Key RCT characteristics and results are described in Tables 2 and 3. Limitations in study relevance and study design and conduct are described in Tables 4 and 5.

Kollins (2020) reported results of the STARS-ADHD (Software Treatment for Actively Reducing Severity of ADHD) randomized double blind trial, which compared treatment with EndeavorRx® (AKL-T01) to a digital control (EVO Words) that targets cognitive domains other than those targeted by AKL-T01.[10] AKL-T01 is a digital game played on a mobile device as described above. EVO Words requires the child to spell as many words as possible by connecting letters in a grid in a fixed amount of time. Parents and children were informed that the study was evaluating two different investigational interventions for ADHD, and only the study coordinator was aware of which video game that the children received. Compliance was monitored by study coordinators, who notified parents by email if the game was not played for more than 48 hours. After four weeks, patients were reassessed for attentional functioning, ADHD symptoms, and impairment. The primary outcome was the change in the computerized test of variable of attention, attention performance index (TOVA API). Secondary outcomes
included a number of clinician and parent reported measures such as the ADHD rating scale, Impairment Rating Scale, and Clinical Global Impressions-Improvement. Out of 348 patients who were randomly assigned, five were lost to follow-up, four were withdrawn by the parent or investigator, and 10 had invalid test results, resulting in a final sample of 329 children for the primary outcome measure. The two children who received the incorrect allocation were included in the intention-to-treat population. The mean change from baseline on the TOVA API was 0.93 in the AKL-T01 group and 0.03 in the control group (p<.05). However, there were no between-group differences for secondary measures, which included the clinician and parent ratings of ADHD symptoms; both groups showed improvement in ADHD ratings from baseline to post-treatment. Treatment-related adverse events AKL-T01 group included frustration (5 [3%] of 180) and headache (3 [2%] of 180) with a mean number of completed sessions of 83%, compared to 96% compliance in the EVO Words group. The study was well-designed and conducted, but there are a number of limitations in study relevance due to the limited age range, limited follow-up, and most importantly the uncertainty of the association of computerized tests with observable behavior. There are also questions regarding the most effective treatment schedule and characteristics of patients who might benefit from this intervention. The trial authors conclude "the results of the current trial are not sufficient to suggest that AKL-T01 should be used as an alternative to established and recommended treatments for ADHD." This study was funded by Akili Interactive Labs and multiple study authors have a financial interest in the funding company.

Table 2. Summary of Key RCT Characteristics

<table>
<thead>
<tr>
<th>Study; Trial</th>
<th>Countries</th>
<th>Sites</th>
<th>Dates</th>
<th>Participants</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kollins (2020); STARS-ADHD[10]</td>
<td>US</td>
<td>20</td>
<td>2016 to 2017</td>
<td>348 pediatric patients aged 8 to 12 years, with confirmed ADHD, TOVA API scores -1.8 and below, without or with washout of disorder-related medication.</td>
<td>AKL-T01 (EndeavorRx®) for 25 min a day on 5 days per week for 4 weeks (n=180)</td>
</tr>
</tbody>
</table>

ADHD: attention-deficit/hyperactivity disorder; RCT: randomized controlled trial; STARS-ADHD: Software Treatment for Actively Reducing Severity of ADHD; TOVA API: test of variables of attention, attention performance index.

Table 3. Summary of Key RCT Results

<table>
<thead>
<tr>
<th>Study</th>
<th>TOVA API mean improvement (SD)</th>
<th>TOVA API Improvement &gt;1.4 points n/N (%)</th>
<th>ADHD-Rating Scale Improvement &gt;2 points n/N (%)</th>
<th>Impairment Rating Scale n/N (%)</th>
<th>Clinical Global Impressions ≤2 n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>329</td>
<td>329</td>
<td>337</td>
<td>332</td>
<td>339</td>
</tr>
<tr>
<td>AKL-T01</td>
<td>0.93 (3.15)</td>
<td>79/169 (47%)</td>
<td>128/173 (74%)</td>
<td>82/171 (48%)</td>
<td>29/175 (17%)</td>
</tr>
<tr>
<td>EVO Words</td>
<td>0.03 (3.16)</td>
<td>51/160 (32%)</td>
<td>119/164 (73%)</td>
<td>60/161 (37%)</td>
<td>26/164 (16%)</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.05</td>
<td>0.006</td>
<td>0.77</td>
<td>0.049</td>
<td>0.86</td>
</tr>
</tbody>
</table>
ADHD: attention deficit/hyperactivity disorder; RCT: randomized controlled trial; SD: standard deviation; STARS-ADHD: Software Treatment for Actively Reducing Severity of ADHD; TOVA API: test of variables of attention, attention performance index.

Tables 4 and 5 display notable limitations identified in each study.

**Table 4. Title**

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes</th>
<th>Duration of Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kollins (2020)</td>
<td>4. The study population was limited to children 8 to 12 years of age.</td>
<td></td>
<td></td>
<td>6. Improvement on computerized tests of attention is weakly associated with classroom attention.</td>
<td>1. There was no follow-up after the 4 week intervention period.</td>
</tr>
</tbody>
</table>

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

| a Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use. |
| b Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest. |
| c Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3. Delivery not similar intensity as intervention; 4. Not delivered effectively. |
| d Outcomes key: 1. Key health outcomes not addressed; 2. Physiologic measures, not validated surrogates; 3. No CONSORT reporting of harms; 4. Not establish and validated measurements; 5. Clinical significant difference not prespecified; 6. Clinical significant difference not supported. |
| e Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms. |

**Table 5. Title**

<table>
<thead>
<tr>
<th>Study</th>
<th>Allocation</th>
<th>Blinding</th>
<th>Selective Reporting</th>
<th>Data Completeness</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kollins (2020)</td>
<td></td>
<td></td>
<td></td>
<td>2. Missing data was not included in the intention-to-treat analysis.</td>
<td></td>
</tr>
</tbody>
</table>

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

| d Data Completeness key: 1. High loss to follow-up or missing data; 2. Inadequate handling of missing data; 3. High number of crossovers; 4. Inadequate handling of crossovers; 5. Inappropriate exclusions; 6. Not intent to treat analysis (per protocol for noninferiority trials). |
| e Power key: 1. Power calculations not reported; 2. Power not calculated for primary outcome; 3. Power not based on clinically important difference. |
| f Statistical key: 1. Analysis is not appropriate for outcome type: (a) continuous; (b) binary; (c) time to event; 2. Analysis is not appropriate for multiple observations per patient; 3. Confidence intervals and/or p values not reported; 4. Comparative treatment effects not calculated. |

**Nonrandomized Studies**

In 2021, Kollins published the results of an additional open-label study of the effectiveness of EndeavorRx® as an adjunct to pharmacotherapy in 8 to 14-year-old children with ADHD on stimulant medication (n = 130) or not on any ADHD medication (n = 76). Study participants were instructed to use the EndeavorRx® (approximately 25 min/day, five days/week) followed by a treatment break of four weeks and a second treatment period of four weeks. The primary study outcome was change in ADHD-related impairment as assessed by the Impairment Rating Scale (IRS) after four weeks. Secondary outcomes included changes in IRS, ADHD Rating Scale (ADHD-RS) and Clinical Global Impressions Scale—Improvement (CGI-I) on
Significantly improved ADHD-related impairment as measured by clinician-rated IRS was found after the first 4-week treatment in both cohorts; mean changes from Baseline to Day 28 in IRS overall severity score was −0.7 (95% confidence interval (CI): [−0.86, −0.50]; DOF: 127; Cohen’s d: .65; p < 0.001) in the On Stimulants cohort and −0.5 (95% CI: [−0.73, −0.32]; DOF: 73; Cohen’s d: .59; p < 0.001) in the No Stimulants cohort.

Participants with an improvement of ≥1 point on the IRS total score from Baseline to Day 28 were considered responders, and 55.5% of the On Stimulants cohort and 40.5% of the No Stimulants cohort were IRS responders. Significant improvement also was found in both cohorts for all secondary endpoints. Mean change from baseline to Day 56 in IRS overall severity score, ADHD-RS total score, and Inattention and Hyperactivity-Impulsivity subscale scores remained significantly improved for participants in both cohorts (all p < 0.001), indicating stability of treatment effects over this timeframe. While this study provides valuable information regarding longer-term treatment effects and observations in an expanded population not available from the pivotal trial discussed above, there are considerable limitations to the study. This study was conducted without randomization did not include a blinded control condition, which precludes evaluation of a possible placebo effects. The manufacturer of the application, Akili Interactive Labs, provided research support and was involved in trial conceptualization. Multiple study authors have a financial interest in the study product. There was no clear effort to mitigate the potential for bias resulting from these possible conflicts of interest.

SUMMARY OF EVIDENCE

For individuals with ADHD who receive a prescription digital therapy, the evidence includes an RCT and an open-label, uncontrolled study. Relevant outcomes are symptoms, functional outcomes, quality of life, and treatment-related morbidity. The single RCT that has been identified compared outcomes of the predecessor of the FDA-cleared EndeavorRx® (AKL-T01) to a word game that targeted different cognitive abilities. Although the experimental treatment group had significantly greater improvement on a computerized test of attention, both the experimental and control groups improved to a similar extent on parent and clinician assessments. The clinical significance of an improvement in a computerized test of attention without a detectable improvement in behavior by parents and clinicians is uncertain. A single-arm, open-label study evaluating the EndeavorRx® in patients with ADHD with and without current pharmaceutical intervention provided additional information regarding the effectiveness of the intervention in a broader population. However, the lack of a control group or randomization limit interpretation of study findings. Several questions remain concerning the efficacy of this treatment. At this time, the digital therapy is not recommended as an alternative or adjunct to established treatments. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

PRACTICE GUIDELINE SUMMARY

AMERICAN ACADEMY OF PEDIATRICS

In 2019, the American Academy of Pediatrics (AAP) updated their 2011 clinical practice guideline on the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder (ADHD) in children and adolescents.[1]

The guidelines were based on a systematic evidence review by the Agency for Healthcare Research and Quality. The AAP gave strong recommendations based on level A evidence for medications and training and behavioral treatment for ADHD implemented with the family and school.
In 2020, the Society for Developmental and Behavioral Pediatrics published a clinical practice guideline for the assessment and treatment of children and adolescents with complex ADHD.[11] Complex ADHD is defined by age (<4 years or presentation >12 years), presence of coexisting conditions, moderate to severe functional impairment, diagnostic uncertainty, or inadequate response to treatment. The society gave a strong recommendation based on grade B evidence for psychoeducation and evidence-based behavioral and educational interventions (e.g., parent training, classroom management, behavioral peer interventions, organizational skills training). The society gave a recommendation based on grade C to B evidence for the frequent need to combine behavioral approaches with pharmacological treatments, and that "treatment should focus on areas of functional impairment and not just symptom reduction, by incorporating developmentally appropriate strategies for self-management, skill building, and prevention of adverse outcomes."

**SUMMARY**

There is not enough research to show that digital health products for the treatment of attention-deficit/hyperactivity disorder (ADHD) improves net health outcomes. No clinical guidelines based on research recommend digital health products for the treatment of attention-deficit/hyperactivity disorder (ADHD). Therefore, digital health products (including digital therapeutics) for the treatment of attention-deficit/hyperactivity disorder (ADHD) are considered investigational.

**REFERENCES**


**CODES**

**NOTE:** Not all digital health products will have a specific code. These are examples of codes that may be relevant.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0702T</td>
<td>Remote therapeutic monitoring of a standardized online digital cognitive behavioral therapy program ordered by a physician or other qualified health care professional; supply and technical support, per 30 days</td>
</tr>
<tr>
<td></td>
<td>0703T</td>
<td>Management services by physician or other qualified health care professional, per calendar month</td>
</tr>
<tr>
<td></td>
<td>99199</td>
<td>Unlisted special service, procedure or report [when specified as a digital health management software application]</td>
</tr>
<tr>
<td>HCPCS</td>
<td>A9291</td>
<td>Prescription digital behavioral therapy, FDA cleared, per course of treatment</td>
</tr>
<tr>
<td></td>
<td>E1399</td>
<td>Durable medical equipment, miscellaneous [when specified as a digital health management software application]</td>
</tr>
</tbody>
</table>

*Date of Origin: September 2021*
IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Digital health products are technologies, platforms, and systems that engage consumers for lifestyle, wellness, and health-related purposes. Digital therapeutic products differ from digital health products in that they are practitioner-prescribed software that delivers evidence-based therapeutic interventions to prevent, manage, or treat a medical disorder or disease. Digital therapeutic products have been proposed to supplement or replace individual or group therapy and/or to deliver cognitive-behavioral therapy for the treatment of substance use disorders.

MEDICAL POLICY CRITERIA

Note:
- Member contracts for covered services vary. Member contract language takes precedence over medical policy.
- This policy does not address:
  - Software that is used for the function or control of an FDA-cleared or approved stand-alone medical device (e.g. external insulin pump or pacemaker).
  - Applications operated by a health care practitioner for remote health monitoring.
The use of a digital health product (including digital therapeutics) for the treatment of a substance use disorder, either as a stand-alone treatment or as an adjunct to standard treatment, is considered investigational, including but not limited to reSET® and reSET-O®.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES
1. Digital Health Products, Medicine, Policy No. 175

BACKGROUND

SUBSTANCE USE DISORDER

The American Psychiatric Association (APA) defines substance use disorder (SUD) as a complex condition “in which there is uncontrolled use of a substance despite harmful consequence. People with SUD have an intense focus on using a certain substance(s) such as alcohol, tobacco, or illicit drugs, to the point where the person’s ability to function in day-to-day life becomes impaired.”[1] The APA notes that individuals can become addicted to several substances including alcohol, marijuana, PCP, LSD and other hallucinogens, inhalants, opioids, sedatives, hypnotics, anxiolytics, cocaine, methamphetamine and other stimulants, and tobacco. The Diagnostic and Statistical Manual of Mental Disorders (DSM) details 11 problematic patterns of use that lead to clinically significant impairment or distress. Mild substance use disorder (SUD) is defined as meeting 2 to 3 criteria, moderate as 4 to 5 criteria, and severe as 6 or more criteria.

1. Often taken in larger amounts or over a longer period than was intended.
2. A persistent desire or unsuccessful efforts to cut down or control use.
3. A great deal of time is spent in activities necessary to obtain, use, or recover from the substance’s effects.
4. Craving or a strong desire or urge to use the substance.
5. Recurrent use resulting in a failure to fulfill major role obligations at work, school, or home.
6. Continued use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by its effects.
7. Important social, occupational, or recreational activities are given up or reduced because of use.
8. Recurrent use in situations in which it is physically hazardous.
9. Continued use despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the substance.
10. Tolerance.

TREATMENT

Treatments for drug addiction include behavioral counseling, skills training, medication, treatment for withdrawal symptoms, treatment for co-occurring mental health issues, and long-
term follow-up to prevent relapse. For patients with primary opioid use disorder (OUD),
medication-assisted treatment is the most common approach. U.S. Food and Drug
Administration (FDA) approved drugs for opioid use treatment include a full opioid agonist
(methadone), a partial opioid agonist (buprenorphine), and an opioid antagonist (naltrexone).
These are used to suppress withdrawal symptoms and reduce cravings and may be used in
combination with counseling and behavioral therapies.

One common psychosocial intervention is cognitive-behavioral therapy (CBT). CBT is an
established therapy based on social learning theory that addresses a patient’s thinking and
behavior. CBT has proven positive effects for the treatment of SUD.[2] There are two main
goals of CBT: first, recognize thoughts and behaviors that are associated with substance
abuse, and second, expand the repertoire of effective coping responses. Specific goals for
SUD and OUD include a better understanding of risk factors for use, more accurate attributions
of cause and effect, increased belief in the ability to address problems, and coping skills.
Specific skills may include motivation, drink/drug refusal skills, communication, coping with
anger and depression, dealing with interpersonal problems, and managing stress.

The community reinforcement approach is a form of CBT that has a goal of making abstinence
more rewarding than continued use. Community reinforcement approach increases non-drug
reinforcement by teaching skills and encouraging behaviors that help improve employment
status, family/social relations and recreational activities. Community reinforcement approach
was originally developed for alcohol dependence and cocaine use, and has been shown to be
more effective than usual care in reducing the number of substance use days.

Contingency management may also be a component of addiction treatment. Contingency
management, also known as motivational incentives, provides immediate positive
reinforcement to encourage abstinence and attendance. Positive reinforcement may range
from a verbal/text acknowledgement of completion of a task to monetary payment for drug-
negative urine specimens. Contingency management is based on the principles of operant
conditioning as formulated by B.F. Skinner, which posits that rewarding a behavior will
increase the frequency of that behavior. Contingency management is typically used to
augment a psychosocial treatment such as community reinforcement approach.

The combination of community reinforcement approach plus contingency management was
shown in a 2018 network meta-analysis of 50 RCTs to be the most efficacious and accepted
intervention among 12 structured psychosocial interventions, including contingency
management alone, in individuals with cocaine or amphetamine addiction.[3] Positive
reinforcement with voucher draws (eg, from a fishbowl) of variable worth that range from a
congratulatory message to an occasional high dollar value are as effective as constant
monetary vouchers. Studies conducted by the National Drug Abuse Treatment Clinical Trials
Network have shown that intermittent reinforcement with incentives totaling $250 to $300 over
8 to 12 weeks both increases retention in a treatment program and reduces stimulant drug use
during treatment.[4]

SOFTWARE AS A MEDICAL DEVICE

The International Medical Device Regulators Forum, a consortium of medical device regulators
from around the world which is led by the FDA, distinguishes between 1) software in a medical
device and 2) software as a medical device (SaMD). The Forum defines SaMD as "software
that is intended to be used for one or more medical purposes that perform those purposes
without being part of a hardware medical device".[5]
FDA's Center for Devices and Radiological Health is taking a risk-based approach to regulating SaMD. Medical software that "supports administrative functions, encourages a healthy lifestyle, serves as electronic patient records, assists in displaying or storing data, or provides limited clinical decision support, is no longer considered to be and regulated as a medical device".[6]

Regulatory review will focus on mobile medical apps that present a higher risk to patients.

- Notably, FDA will not enforce compliance for lower risk mobile apps such as those that address general wellness.
- FDA will also not address technologies that receive, transmit, store, or display data from medical devices.

The agency has launched a software pre-cert pilot program for SaMD that entered its test phase in 2019. Key features of the regulatory model include the approval of manufacturers prior to evaluation of a product, which is based on a standardized "Excellence Appraisal" of an organization, and its commitment to monitor product performance after introduction to the U.S. market. Criteria include excelling in software design, development, and validation. Companies that obtain pre-certification participate in a streamlined pre-market review of the SaMD. Pre-certified organizations might also be able to market lower-risk devices without additional review. In 2017, FDA selected 9 companies to participate in the pilot program, including Pear Therapeutics.

REGULATORY STATUS

In 2017, reSET® (Pear Therapeutics), received De Novo marketing clearance from the FDA to provide CBT as an adjunct to contingency management, for patients with SUD who are enrolled in outpatient treatment under the supervision of a clinician (DEN160018). This is the first prescription digital therapeutic to be approved by the FDA. FDA product code: PWE

In 2018, reSET-O® (Pear Therapeutics) was cleared for marketing by the FDA through the 510(k) pathway as a prescription-only digital therapeutic to “increase retention of patients with opioid use disorder (OUD) in outpatient treatment by providing cognitive behavioral therapy, as an adjunct to outpatient treatment that includes transmucosal buprenorphine and contingency management” (K173681). FDA determined that this device was substantially equivalent to existing devices. The predicate device was reSET®.

Vorvida® and Modia® (Orexo) provide support for individuals with problematic drinking and OUD. These digital technologies have not received marketing clearance by U.S. Food and Drug Administration and are not reviewed here. They are currently available in the U.S. through the Enforcement Policy for Digital Health Devices for Treating Psychiatric Disorders During COVID19.[7]

EVIDENCE SUMMARY

Evidence reviews assess the clinical evidence to determine whether the use of technology improves the net health outcome. Broadly defined, health outcomes are the length of life, quality of life, and ability to function including benefits and harms. Every clinical condition has specific outcomes that are important to patients and managing the course of that condition. Validated outcome measures are necessary to ascertain whether a condition improves or worsens; and whether the magnitude of that change is clinically significant. The net health outcome is a balance of benefits and harms.
To assess whether the evidence is sufficient to draw conclusions about the net health outcome of a technology, two domains are examined: the relevance, and quality and credibility. To be relevant, studies must represent one or more intended clinical use of the technology in the intended population and compare an effective and appropriate alternative at a comparable intensity. For some conditions, the alternative will be supportive care or surveillance. The quality and credibility of the evidence depend on study design and conduct, minimizing bias and confounding that can generate incorrect findings. The randomized controlled trial (RCT) is preferred to assess efficacy; however, in some circumstances, nonrandomized studies may be adequate. RCTs are rarely large enough or long enough to capture less common adverse events and long-term effects. Other types of studies can be used for these purposes and to assess generalizability to broader clinical populations and settings of clinical practice.

DIGITAL HEALTH TECHNOLOGIES FOR SUBSTANCE USE DISORDER

Clinical Context and Therapy Purpose

Substance abuse is a serious health problem in the U.S. A 2019 survey from the Substance Abuse and Mental Health Services Administration found that 20.4 million people age 12 or older in the U.S., or 7.4 percent of the U.S. population, had substance use disorder (SUD), but only 1.5 million people were enrolled in substance use treatment.[8] The most common substances reported in the survey are alcohol, followed by tobacco and marijuana. Illicit drug use and prescription drug misuse occur in a lower percentage of the population.

Significant barriers to treatment exist for patients with SUD and opioid use disorder (OUD). There are an insufficient number of clinicians who are trained for substance abuse treatment, particularly in rural areas, and access to outpatient programs may be difficult and time consuming. Patients typically present to their primary care provider, who may not have sufficient time or training to treat patients with substance abuse. In addition, the stigma of substance abuse may prevent individuals from seeking treatment. Digital technologies could potentially increase access to specialty care for patients who may not otherwise be able to attend a treatment program. Digital technologies might also reduce the need for attendance in a clinic for patients who avoid treatment due to stigma or other factors.

A computer-delivered cognitive-behavioral therapy (CBT) program named CBT4CBT (Computer-Based Training for Cognitive Behavioral Therapy) has been developed to provide therapy for patients with substance abuse. The program includes seven core CBT skills delivered by on-screen narration, graphic animation, quizzes, and interactive exercises. In a 2018 RCT, both clinician and computer delivery of CBT reduced the frequency of substance use more than treatment as usual.[9] In addition, patients who received the computer-based CBT with minimal monitoring had the best treatment retention, learning of CBT concepts, and six-month outcomes compared to either clinician-delivered CBT or treatment as usual. A computer-based community reinforcement approach (CRA) plus vouchers was reported in a 2008 study to lead to similar levels of abstinence as patients who received clinician-guided CRA plus vouchers.[10] These results suggest that computerized CRA (CCRA) could potentially substitute for clinician-guided therapy and increase access to treatment.

In 2017 and 2018, the first prescription mobile apps (ie, reSET® and reSET-O®) were cleared for marketing by the U.S. Food and Drug Administration (FDA). These have the potential to increase access to substance abuse treatments in patients who have SUD or OUD. These two apps are intended to provide CCRA as an addition to traditional therapy in the context of an outpatient program.
Evaluation of clinically meaningful outcomes

The outcome which is most frequently cited as the most important outcome for patients is abstinence from the substance of abuse. \textsuperscript{[11]} This primary outcome should be measured during therapy, at the end of therapy, and at longer-term (e.g., 3, 6, and 12 months) follow-up to assess the durability of the treatment.

Other outcomes that have been reported as important to patients are drug craving, employment, and stable relationships. A semi-structured assessment of seven potential problem areas in substance-abusing patients is the Addiction Severity Index. \textsuperscript{[12]} The domains are medical status, employment and support, drug use, alcohol use, legal status, family/social status, and psychiatric status. The Addiction Severity Index provides severity ratings of the client’s need for treatment and composite scores which measure problem severity during the prior 30 days.

The Maudsley Addiction Profile is a brief standardized interview that assesses treatment outcomes in domains of substance abuse, health risk behavior, physical and psychological health, and personal social functioning. \textsuperscript{[13]}

Retention in a treatment program is commonly used in addiction research but is an indirect measure of treatment success. Observational data from the Drug Abuse Treatment Outcome Studies suggest that most addicted individuals need at least three months in treatment to significantly reduce or stop their drug use and that the best outcomes occur in patients who participate in longer treatment. \textsuperscript{[14, 15]}

REVIEW OF EVIDENCE

Randomized Controlled Trials

The two pivotal RCTs for the prescription digital apps for substance use disorder (SUD) (resSET) and opioid use disorder (OUD) (reSET-O®) are described below and in Tables 1 and 2. The technology was developed by the National Institute of Drug Abuse-funded Center for Technology and Behavioral Health as the Therapeutic Education System, which was subsequently submitted to the FDA for a mobile platform by Pear Therapeutics.

Campbell (2014) reported the pivotal multicenter trial for reSET®, in which patients with SUD or OUD completed 20 to 30 minute multimedia modules on a desktop while in the clinic or at home. \textsuperscript{[16, 17]} The active treatment was the Therapeutic Education System, which combined CCRA plus contingency management, and was compared to treatment as usual (therapy alone) at 10 community-based outpatient treatment programs as part of the National Drug Abuse Clinical Trials Network. Clinicians were able to access reports on computer activity and discussed module completion in the individual therapy sessions. Contingency management consisted of random selection of vouchers, which ranged from a congratulatory message to $100 cash, for module completion and negative urine drug results. The mean dollar earned was $277 (SD $226) over the 12 weeks. Although the study was intended to replace some of the hours of therapy, the Therapeutic Education System group received the same number of therapy session as the control group, so the combined program was effectively in addition to counseling alone.

The co-primary outcomes were abstinence from drug/heavy alcohol use in the last four weeks of treatment and retention in the treatment program. In the analysis by Campbell (2014), \textsuperscript{[16] the}
Therapeutic Education System reduced drop-out from the treatment program (hazard ratio=0.72 [95% CI: 0.57 to 0.92], p=.010), and the odds of achieving abstinence was 1.62 fold greater in the group with CCRA and contingency management group (p=.010). However, the beneficial effect of the Therapeutic Education System was observed only in patients who were not abstinent at baseline. For patients who were abstinent at baseline, the Therapeutic Education System did not increase abstinence, and at three- and six-months follow-up, the effect of Therapeutic Education System was no longer significant. Subsequent analyses of the trial found that the Therapeutic Education System was not associated with improvements in social functioning compared to standard outpatient care.[18]

In the FDA analyses of the trial,[17] results were analyzed for the entire cohort and for cohorts that excluded patients who reported opioid use. Abstinence during weeks 9 to 12 and total abstinence with CCRA plus contingency management was significantly greater in the cohort as a whole and more so in the analyses that excluded primary opioid users. For example, abstinence during weeks 9 to 12 was 40.3% in the SUD subgroup who received CCRA plus vouchers compared to 17.6% in the group who received only therapy (p<.001). Total abstinence, defined as the number of half weeks with a negative urine drug test, was 11.9 half weeks in the SUD subgroup who received the experimental treatment and 8.8 half weeks in controls (p=0.003).

In the pivotal study reported by Christensen (2014), CCRA was added to treatment as usual in patients who had opioids as the primary substance of abuse.[19, 20] Treatment as usual in this second trial included clinic visits three times per week with a reward for a negative urine drug screen (maximum of $997.50), sublingual buprenorphine/naloxone, and a clinician visit every two weeks. Patients who did not show up for any of the thrice weekly clinic visits were considered to have a positive drug screen and were considered drop-outs if they missed three visits in a row. The primary outcomes were the longest continuous abstinence and total abstinence. The study was powered to detect a three-week difference between groups in mean weeks of continuous abstinence. In the 84-day treatment program there were 9.7 more days of abstinence in the CCRA group (67.1 days) than in the control group (57.4 days, p=0.01). The trial did not meet one of the primary outcomes of a significant difference between the two groups in the longest abstinence (5.5 days p=0.214). The group using the computerized therapy had an increase in medication Addiction Severity Index scores (p=0.04), but did not show a significant improvement on the overall Addiction Severity Index (p>0.16). The data on abstinence and Addiction Severity Index was not reported in the 510(K) Summary for the U.S. FDA.[20]

Both trials reported a significant increase in retention during the 12-week programs. The SUD subgroup had a 23.8% drop out rate compared to 36.8% in the control group (p=.004). The addition of CCRA to treatment as usual in patients with OUD also increased retention, with a hazard ratio for dropping out of treatment of 0.47 (0.26 to 0.85).

The purpose of the limitations tables (Tables 3 and 4) is to display notable limitations identified in each study. This information is synthesized as a summary of the body of evidence following each table and provides the conclusions on the sufficiency of evidence supporting the position statement. Both trials had limitations in relevance and in design and conduct that preclude determination of the effect of the intervention on relevant health outcomes.
• A major limitation for the reSET® and reSET-O® mobile apps is regarding the generalizability of results from these trials. Both studies were conducted with desktop computers, used primarily during clinic visits. In the study by Christensen (2014) CCRA was only available in the clinic to avoid confounding the efficacy of the program with compliance issues. Regular use of a mobile app without close supervision and outside of the constraints of a trial setting is unknown. Although a proposed benefit of digital technology is to increase access to evidence-based treatments, particularly in rural areas or where there are other limitations to specialist care, consistent use of a mobile device in the home and the resources and expertise of local providers to supervise addiction treatment is uncertain.

• An additional major concern in the study of Campbell (2014) is the experimental group received both the web-based CCRA and a reward for a negative drug test. The trial was designed to assess the combined treatment approach, and not specifically the CCRA program. Because a reward for a negative drug screen is known by itself to increase both retention and abstinence during a trial,[4] the contribution of the digital technology to the increase in abstinence in patients with SUD cannot be determined. Notably, abstinence was not improved at the three and six-month follow-up, raising further questions about whether the increase in abstinence during the trial was due to contingency management rather than the CCRA.

• A limitation of both trials is the choice (eg, retention) and timing (eg, during treatment) of the outcome measures. Abstinence after a treatment program is a main objective of therapy. In the study published by Christensen (2014), the main effect of the technology was on retention, and there was no follow-up after 12 weeks. In the study published by Campbell (2014), abstinence was greater during the trial, but not improved at the three and six-month follow-up.

• An additional limitation is the potential for performance bias among the volunteers in these unblinded studies. Nearly half of patients who qualified for the Campbell (2014) study chose not to participate. There may have been greater motivation to use the new technology in patients who agreed to participate in the study. While acknowledging the difficulty of blinding with this type of intervention, providing a control intervention of similar intensity, such as computer time that is not based on CRA, is feasible.

Given all of these limitations, further study in well-designed trials is needed to determine the effects of the technology on addiction.
Table 1. Summary of Key RCT Characteristics

<table>
<thead>
<tr>
<th>Study; Trial</th>
<th>Countries</th>
<th>Sites</th>
<th>Participants</th>
<th>Interventions</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campbell (2014) FDA Summary DEN160018[^16, ^17]</td>
<td>U.S.</td>
<td>10</td>
<td>507 adult patients with self-report of drug use, with a subset of 305 who did not have primary use of opioids treated at community health centers</td>
<td>12 weeks of treatment as usual + CCRA (62 modules on a desktop) + contingency management for module completion and negative drug screen (n=255)</td>
<td>12 weeks of treatment as usual consisting ≥ 2 individual or group therapy sessions per week (n=252)</td>
</tr>
<tr>
<td>Christensen (2014) FDA summary K173681[^19, ^20]</td>
<td>U.S.</td>
<td>1</td>
<td>170 opioid-dependent adults</td>
<td>12 weeks of CCRA (69 modules on a desktop in the clinic) + contingency management + buprenorphine/ naloxone (n=92)</td>
<td>12 weeks of contingency management + buprenorphine/ naloxone (n=78)</td>
</tr>
</tbody>
</table>

CCRA: computer-based community reinforcement approach; RCT: randomized controlled trial.

[^16]: 20 to 30 min multimedia computer modules. Patients completed a mean of 36.6 (standard deviation, 18.1) out of 62 total CCRA modules in the study by Campbell et al.
[^17]: There were a total of 69 CCRA modules in the study by Christensen et al.

Table 2. Summary of Key RCT Results

<table>
<thead>
<tr>
<th>Study</th>
<th>Abstinence</th>
<th>Total Abstinence</th>
<th>Retention</th>
<th>Dropping Out of Treatment</th>
<th>ASI overall</th>
<th>ASI Medication Subscale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campbell (2014) FDA Submission DEN160018[^16, ^17]</td>
<td>Rate During Weeks 9-12</td>
<td>Half weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire Cohort (n=507)</td>
<td>Excluding Primary Opioid Abusers (n=399)</td>
<td>Entire Cohort (n=507)</td>
<td>Excluding Primary Opioid Abusers (n=399)</td>
<td>Entire Cohort (n=507)</td>
<td>Excluding Primary Opioid Abusers (n=399)</td>
<td>Entire Cohort (n=507)</td>
</tr>
<tr>
<td>Treatment as usual + CCRA + contingency management</td>
<td>29.7%</td>
<td>40.3%</td>
<td>10.9</td>
<td>11.9</td>
<td>72.2%</td>
<td>76.2%</td>
</tr>
<tr>
<td>Treatment as usual</td>
<td>16.0%</td>
<td>17.6%</td>
<td>8.6</td>
<td>8.8</td>
<td>63.5%</td>
<td>63.2%</td>
</tr>
<tr>
<td>Study</td>
<td>Abstinence</td>
<td>Total Abstinence</td>
<td>Retention</td>
<td>Dropping Out of Treatment</td>
<td>ASI overall</td>
<td>ASI Medication Subscale</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>---------------------------</td>
<td>------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>p</td>
<td>0.008</td>
<td>&lt;0.001</td>
<td>0.002</td>
<td>0.003</td>
<td>0.03</td>
<td>0.004</td>
</tr>
<tr>
<td>Christensen (2014) K173681¹¹⁹, ²⁰</td>
<td>Longest Abstinence in Days (+ SD)</td>
<td>Total Days + SD</td>
<td>Treatment Completion</td>
<td>0.001</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>CRA + contingency management</td>
<td>55</td>
<td>67.1 + 19.3</td>
<td>80.4%</td>
<td>17.6%</td>
<td>0.214</td>
<td>0.011</td>
</tr>
<tr>
<td>Contingency management</td>
<td>49.5</td>
<td>57.4 + 28.0</td>
<td>64.1%</td>
<td>31.6%</td>
<td>Diff: 9.7 (2.3 to 17.2)</td>
<td>OR: 2.30 (1.15 to 4.60)</td>
</tr>
<tr>
<td>HR/Diff/OR (95% CI)</td>
<td>Diff: 5.5</td>
<td>9.7 (2.3 to 17.2)</td>
<td>0.47 (0.26 to 0.85)</td>
<td>0.0224</td>
<td>&gt;0.24</td>
<td>0.04</td>
</tr>
</tbody>
</table>

ASI: Addiction Severity Index; CI: confidence interval; (C)CRA: (computer-based) community reinforcement approach; HR: hazard ratio; OR: odds ratio; RCT: randomized controlled trial; SD: standard deviation.

### Table 3. Study Relevance Limitations

<table>
<thead>
<tr>
<th>Study</th>
<th>Population a</th>
<th>Intervention b</th>
<th>Comparator c</th>
<th>Outcomes d</th>
<th>Follow-Up e</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campbell (2014); FDA Submission DEN16001¹¹⁶, ¹²⁷</td>
<td>4. The study volunteers may not be representative of the general population with substance use disorder.</td>
<td>2. Was an earlier desktop technology and was conducted mostly in the clinic</td>
<td>3. The comparator did not include contingency management with vouchers. Delivery was not a similar intensity as the intervention.</td>
<td>1. Uncertain significance of retention as an outcome. 5. The minimal clinically important difference for abstinence was not pre-specified</td>
<td></td>
</tr>
<tr>
<td>Christensen (2014) K173681¹¹⁹, ²⁰</td>
<td>2. Was an earlier desktop technology and was conducted in the clinic</td>
<td>3. Delivery was not a similar intensity as the intervention.</td>
<td>1. Uncertain significance of retention as an outcome. 5. The minimal clinically important difference for abstinence was not pre-specified.</td>
<td>1. The study did not extend after 12 week treatment period, limiting inferences on efficacy for abstinence.</td>
<td></td>
</tr>
</tbody>
</table>

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

a Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
Table 4. Study Design and Conduct Limitations

<table>
<thead>
<tr>
<th>Study</th>
<th>Allocationa</th>
<th>Blindingb</th>
<th>Selective Reportingc</th>
<th>Data Completenessd</th>
<th>Powere</th>
<th>Statisticalf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campbell (2014); FDA Submission DEN160018(^{[16, 17]})</td>
<td>1. Participants and investigators were not blinded to treatment assignment.</td>
<td>2. Subgroup analyses in the FDA Summary were not pre-specified</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christensen (2014) K173681(^{[19, 20]})</td>
<td>1. Participants and investigators were not blinded to treatment assignment.</td>
<td>2. Data on abstinence was not included in the FDA Summary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.


\(^{b}\) Blinding key: 1. Not blinded to treatment assignment; 2. Not blinded outcome assessment; 3. Outcome assessed by treating physician.

\(^{c}\) Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.

\(^{d}\) Data Completeness key: 1. High loss to follow-up or missing data; 2. Inadequate handling of missing data; 3. High number of crossovers; 4. Inadequate handling of crossovers; 5. Inappropriate exclusions; 6. Not intent to treat analysis (per protocol for noninferiority trials).

\(^{e}\) Power key: 1. Power calculations not reported; 2. Power not calculated for primary outcome; 3. Power not based on clinically important difference.

\(^{f}\) Statistical key: 1. Analysis is not appropriate for outcome type: (a) continuous; (b) binary; (c) time to event; 2. Analysis is not appropriate for multiple observations per patient; 3. Confidence intervals and/or p values not reported; 4. Comparative treatment effects not calculated.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
OBSERVATIONAL STUDIES

Study design and results of observational studies are shown in Tables 5 and 6.

Marichich (2021) performed an industry-funded analysis of reSET-O® data from 3144 patients with OUD who had filled a 12-week prescription of the software.[21] Participants were instructed to complete at least 4 modules per week with a total possible of 31 core modules and 36 supplemental modules. Analysis of the software's data showed that about half of the patients completed all 31 modules, 66% completed half of the modules, and 74% of patients actively participated through 12 weeks. Use decreased from 100% in the first week to 55% of individuals completing 4 modules in week 12. (Retention in the pivotal study by Christensen was 80% for the software compared to 64% for contingency management alone).18,19, Abstinence during the last 4 weeks of treatment was determined by either urine drug screening or self-report recorded on reSET-O®. With a conservative estimate of missing data considered to be a positive drug screen, 66% of patients were estimated to be abstinent during the last 4 weeks of the prescription. For patients who completed 3 to 5 modules in the first week, abstinence in the final 4 weeks ranged from 83% to 89%. A limitation of this study is that patients who completed more modules in the first week may have been more motivated to remain abstinent, and cause and effect cannot be determined from this non-comparative observational study.

Table 5. Observational Study Characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Participants</th>
<th>Treatment</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marichich (2021)[21]</td>
<td>U.S.</td>
<td>3144 patients with buprenorphine medication for OUD who were under the care of a clinician and filled a 12-week prescription for reSET-O®</td>
<td>Four 30 min modules per week for a total of 31 core modules and 36 supplemental modules on a mobile device</td>
<td>12 weeks</td>
</tr>
</tbody>
</table>

OUD: opioid use disorder

Table 6. Title

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment</th>
<th>Participants Completing all Modules</th>
<th>Participants Completing Half of Modules</th>
<th>Retention Through 12 Weeks</th>
<th>Abstinence During Weeks 9-12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marichich (2021)[21]</td>
<td>31 modules of reSET-O®</td>
<td>49%</td>
<td>66%</td>
<td>74.2%</td>
<td>66%</td>
</tr>
</tbody>
</table>

SUMMARY OF EVIDENCE

For individuals with SUD who receive a prescription digital therapeutic, the evidence includes two pivotal RCTs and analysis of data of more than 3000 patients from the mobile app. Relevant outcomes are symptoms, morbid events, change in disease status, quality of life, and medication use. Mobile digital technology is proposed to increase access to evidence-based cognitive-behavioral treatments, particularly in rural areas or where there are other limitations to specialist care. Although it is theoretically appealing to replace in-person counseling with computer-based therapy, there are a number of limitations in the current evidence base that limit any conclusions regarding efficacy. Specifically, one of the trials assessed the combined intervention of computer-based learning and a reward for abstinence. Since reward for abstinence alone has been shown to increase both abstinence and retention, the contribution of the web-based program to the overall treatment effect cannot be determined. The treatment effect on abstinence was not observed at follow-up, raising further questions about the relative...
effects of the rewards and the web program. The second trial, conducted in patients with primary opioid use disorder, did not meet a primary objective of longest days of abstinence. While both RCTs reported a positive effect on the intermediate outcome of retention, the relationship between retention and relevant health outcomes in these trials is uncertain. In addition, both trials were conducted with an earlier technology (a desktop in a clinic) and were unblinded, so there are issues of possible performance bias and questions about generalizability of these results. The observational study shows that participants who complete more modules with the mobile app have greater abstinence during weeks 9 to 12, but the lack of a control group with comparable motivation limits interpretation of results. Given all of these limitations, further study in well-designed trials is needed to determine the effects of prescription digital therapeutics on relevant outcomes in patients with SUD and OUD. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

PRACTICE GUIDELINE SUMMARY

NATIONAL INSTITUTE ON DRUG ABUSE

The 2018 Principles of Drug Addiction and Treatment from the National Institute on Drug Abuse describes evidence-based approaches to drug addiction treatment.[15] Behavioral therapies include cognitive-behavioral therapy (alcohol, marijuana, cocaine, methamphetamine, nicotine), contingency management (alcohol, stimulants, opioids, marijuana, nicotine), community reinforcement approach plus vouchers (alcohol, cocaine, opioids), motivational enhancement therapy (alcohol, marijuana, nicotine), the matrix model (stimulants), 12-step facilitation therapy (alcohol, stimulants, opiates) and family behavior therapy.

SUMMARY

There is not enough research to show that digital health products for the treatment of substance use disorders improves net health outcomes. No clinical guidelines based on research recommend digital health products for the treatment of substance use disorders. Therefore, digital health products (including digital therapeutics) for the treatment of substance use disorders are considered investigational.

REFERENCES


### CODES

**NOTE:** Not all digital health products will have a specific code. These are examples of codes that may be relevant.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0702T</td>
<td>Remote therapeutic monitoring of a standardized online digital cognitive behavioral therapy program ordered by a physician or other qualified health care professional; supply and technical support, per 30 days</td>
</tr>
<tr>
<td></td>
<td>0703T</td>
<td>Management services by physician or other qualified health care professional, per calendar month</td>
</tr>
<tr>
<td>HCPCS</td>
<td>99199</td>
<td>Unlisted special service, procedure or report [when specified as a digital health management software application]</td>
</tr>
<tr>
<td></td>
<td>A9291</td>
<td>Prescription digital behavioral therapy, FDA cleared, per course of treatment</td>
</tr>
<tr>
<td></td>
<td>E1399</td>
<td>Durable medical equipment, miscellaneous [when specified as a digital health management software application]</td>
</tr>
</tbody>
</table>

*Date of Origin: September 2021*
IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Laser interstitial thermal therapy (LITT) involves the introduction of a laser fiber probe to deliver thermal energy for the targeted ablation of diseased tissue. The goal of therapy is selective thermal injury through the maintenance of a sharp thermal border, as monitored via the parallel use of real-time magnetic resonance (MR) thermography and controlled with the use of actively cooled applicators. In neurological applications, LITT involves the creation of a transcranial burr hole for the placement of the laser probe at the target brain tissue. Probe position, ablation time, and intensity are controlled under MRI guidance. LITT has been proposed as a less invasive treatment option for patients with neurological conditions compared to surgery.

MEDICAL POLICY CRITERIA

I. Laser interstitial thermal therapy (LITT) may be considered medically necessary for the treatment of refractory epilepsy when both of the following Criteria (A. and B.) are met:
   
   A. There is documentation of disabling seizures despite use of two or more antiepileptic drug regimens (i.e., medically refractory epilepsy), and
B. There is a well-defined epileptogenic focus of seizure propagation in the temporal lobe or hypothalamus accessible by LITT.

II. Laser interstitial thermal therapy (LITT) is considered **investigational** for all other neurological indications, including but not limited to the treatment of refractory epilepsy when Criterion I. is not met and for the treatment of primary or metastatic brain tumors or radiation necrosis.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

---

**LIST OF INFORMATION NEEDED FOR REVIEW**

**REQUIRED DOCUMENTATION:**

The information below **must** be submitted for review to determine whether policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

1. Medical records related to:
   - History and physical/chart notes including those documenting disabling seizures
   - Conservative treatment provided, including documentation of two or more antiepileptic drug regimens
   - Documentation of well-defined epileptogenic focus of seizure propagation in the temporal lobe or hypothalamus **that is accessible by LITT.**

---

**CROSS REFERENCES**

1. Stereotactic Radiosurgery and Stereotactic Body Radiation Therapy of Intracranial, Skull Base, and Orbital Sites, Surgery, Policy No. 213
2. Focal Laser Ablation of Prostate Cancer, Surgery, Policy No. 222

---

**BACKGROUND**

**LASER INTERSTITIAL THERMAL THERAPY**

Laser interstitial thermal therapy (LITT) involves the introduction of a laser fiber probe to deliver thermal energy for the targeted ablation of diseased tissue. Thermal destruction of tissue is mediated via DNA damage, necrosis, protein denaturation, membrane dissolution, vessel sclerosis, and coagulative necrosis.[1] The goal of therapy is selective thermal injury through the maintenance of a sharp thermal border, as monitored via the parallel use of real-time magnetic resonance (MR) thermography and controlled with the use of actively cooled applicators.[2] In neurological applications, LITT involves the creation of a transcranial burr hole for the placement of the laser probe at the target brain tissue. Probe position, ablation time, and intensity are controlled under MRI guidance.

The majority of neurological LITT indications described in the literature involve the ablation of primary and metastatic brain tumors, epileptogenic foci, and radiation necrosis in surgically inaccessible or eloquent brain areas.[2] LITT may offer a minimally invasive treatment option for patients with a high risk of morbidity with traditional surgical approaches. The most common complications following LITT are transient and permanent weakness, cerebral edema, hemorrhage, seizures, and hyponatremia.[3] Delayed neurological deficits due to brain edema...
are temporary and typically resolve after corticosteroid therapy. Contraindications to MRI are also applicable to the administration of LITT.

REGULATORY STATUS

In August 2007, the Visualase™ Thermal Therapy System (Medtronic; formerly Biotex, Inc.) received initial marketing clearance by the FDA through the 510(k) pathway (K071328). As of March 2019, the system is indicated for use “to necrotize or coagulate soft tissue through interstitial irradiation or thermal therapy under magnetic resonance imaging (MRI) guidance in medicine and surgery in cardiovascular thoracic surgery (excluding the heart and vessels in the pericardial sac), dermatology, ear-nose-throat surgery, gastroenterology, general surgery, gynecology, head and neck surgery, neurosurgery, plastic surgery, orthopedics, pulmonology, radiology, and urology, for wavelengths 800 nm through 1064 nm” (K181859). Data from compatible MRI sequences can be processed via proton resonance-frequency shift analysis and image subtraction to relate imaging changes to relative changes in tissue temperature during therapy. The Visualase™ cooling applicator utilizes saline.

In April 2013, the NeuroBlate® System (Monteris Medical) received initial clearance for marketing by the FDA through the 510(k) pathway (K120561). As of August 2020, the system is indicated for use “to ablate, necrotize, or coagulate intracranial soft tissue, including brain structures (eg, brain tumor and epileptic foci as identified by non-invasive and invasive neurodiagnostic testing, including imaging), through interstitial irradiation or thermal therapy in medicine and surgery in the discipline of neurosurgery with 1064 nm lasers” (K201056). The device is intended for planning and monitoring of thermal therapy under MRI guidance, providing real-time thermographic analysis of selected MRI images. The NeuroBlate® system utilizes a laser probe with a sapphire capsule to promote prolonged, pulsed laser firing and a controlled cooling applicator employing pressurized CO2.

On April 25, 2018, the FDA issued a safety alert on MR-guided LITT (MRgLITT) devices with a letter to healthcare providers stating that the FDA is currently evaluating data suggesting that potentially inaccurate MR thermometry information can be displayed during treatment which may contribute to a risk of tissue overheating and potentially associated adverse events, including neurological deficits, increased intracerebral edema or pressure, intracranial bleeding, and/or visual changes.[4] Several risk mitigation strategies were recommended. In an updated letter released on November 8, 2018, risk mitigation recommendations specific to the Visualase™ and NeuroBlate® systems were issued.[5]

EVIDENCE SUMMARY

Evidence reviews assess the clinical evidence to determine whether the use of a technology improves the net health outcome. Broadly defined, health outcomes are length of life, quality of life, and ability to function, including benefits and harms. Every clinical condition has specific outcomes that are important to patients and to managing the course of that condition. Validated outcome measures are necessary to ascertain whether a condition improves or worsens; and whether the magnitude of that change is clinically significant. The net health outcome is a balance of benefits and harms.

To assess whether the evidence is sufficient to draw conclusions about the net health outcome of a technology, two domains are examined: the relevance and the quality and credibility. To be relevant, studies must represent one or more intended clinical use of the technology in the intended population and compare an effective and appropriate alternative at a comparable
intensity. For some conditions, the alternative will be supportive care or surveillance. The quality and credibility of the evidence depend on study design and conduct, minimizing bias and confounding that can generate incorrect findings. The randomized controlled trial is preferred to assess efficacy; however, in some circumstances, nonrandomized studies may be adequate. Randomized controlled trials are rarely large enough or long enough to capture less common adverse events and long-term effects. Other types of studies can be used for these purposes and to assess generalizability to broader clinical populations and settings of clinical practice.

**Study Selection Criteria**

Methodologically credible studies were selected using the following principles:

To assess efficacy outcomes, comparative controlled prospective trials were sought, with a preference for RCTs. In the absence of such trials, comparative observational studies were sought, with a preference for prospective studies. To assess long-term outcomes and adverse events, single-arm studies that capture longer periods of follow-up and/or larger populations were sought.

**PRIMARY OR METASTATIC BRAIN TUMORS**

**Clinical Context and Therapy Purpose**

The purpose of MR-guided LITT is to use a focused thermal therapy technique to ablate primary or metastatic brain tumors and to avoid potential complications associated with alternative surgical interventions.

**Review of Evidence**

**Systematic Reviews**

Viozzi (2021) published a systematic review (SR) of data from 11 studies (N=111) of patients treated with laser interstitial thermal therapy (LITT) for newly diagnosed glioblastoma (nGBM) reported in 11 studies.\(^6\) All included studies were conducted in the US predominantly (81%) using the Neuroblate system. Median overall survival (OS) ranged from 4.1 to 32 months and progression free survival (PFS) from 2 to 31 months. No randomized studies were identified for inclusion. All studies had serious or critical risk of bias, and the quality of evidence was graded as very low according to the GRADE criteria. The mean complication rate was 33.7%. No quality-of-life outcomes were reported. The low quality of available evidence regarding LITT for nGBM precluded the author's ability to draw conclusions regarding the net impact of the technology on health outcomes.

Alattar (2019) published a SR of stereotactic laser ablation (SLA, also known as LITT) for the treatment of brain metastases recurring after radiosurgery (BMRS).\(^7\) Thirteen publications were included. Median survival ranged from 5.8 to 19.8 months. About two-thirds of treated lesions showed postablation expansion of contrast-enhancing volume and fluid-attenuated inversion recovery volume, which reached up to three times the pre-operative lesion volume, typically resolved within six months. Median hospital stay was 1-2 days (range, 1-5 days), and most treated patients were discharged home (range, 59.5%-100%). The incidence of SLA-related permanent neurologic injuries was <10%. The most common complications were hemorrhage, thermal injury causing neurologic deficit, and malignant cerebral edema.
Chen (2021) published a systematic review and meta-analysis of retrospective studies and case series investigating the efficacy of LITT for brain metastases with in-field recurrence or radiation necrosis following treatment with SRS. A meta-analysis of 14 studies (470 patients with 542 lesions) was performed. The overall 12-month local control rate ranged between 56.0% and 84.7% with a pooled rate of 69.0% (95% CI, 60.0% to 76.7%; I² = 50.584%; p = 0.048) and pooled overall survival of 17.15 months (95% CI, 13.27 to 24.8). Among 153 recurrent brain metastasis lesions across 5 studies, the 12-month local control rate was 59.9% (95% CI, 47.9% to 70.9%). Among 75 radiation necrosis lesions across 4 studies, the 12-month local control rate was 76.3% (95% CI, 65.0% to 84.8%). Thus, LITT provided more favorable local control efficacy in patients with radiation necrosis compared to those with brain metastasis recurrence. No significant difference in median overall survival at one year was determined between radiation necrosis and brain metastasis groups (66.5% versus 66.8%; p=.978). Survival outcomes were not stratified by pathology and safety outcomes were not reported. Compared to previously reported estimates for surgical resection with a local control rate ranging from 62% to 93% and a median overall survival of 8.7 months, the authors concluded that LITT demonstrates comparable local control but a more satisfactory survival benefit. The analysis is limited by study heterogeneity, small sample sizes, and the lack of a standardized definition for local disease control.

Montemurro (2020) published a SR of data on LITT in the treatment of recurrent glioblastoma including data from 17 studies (N= 203, 219 LITT sessions). The median age was 57.4 years (65.8 % male). Treatment location was most commonly frontal lobe (29 %), followed by temporal (23.9 %), parietal (21.4 %) and occipital lobes (2.6 %). Thalamus, corpus callosum and cerebellum also were treated (23.1 %). Morbidity was 6.4 % with a median hospital stay of 3.5 days. The most common complications were seizures (2%), motor deficits (1.5 %), wound infection (1.5 %), transient hemiparesis (1%) and hemorrhage (0.5 %). All patients underwent adjuvant chemotherapy after treatment. The median PFS and the median OS after laser interstitial thermal therapy was 5.6 months and 10.2 months, respectively. The median OS from diagnosis was 14.7 months.

de Franca (2020) published a SR and meta-analysis of LITT as a therapy for brain tumors compared to stereotactic radiosurgery (SRS) based on 25 studies. Patient populations included patients with brain metastasis and recurrent glioblastoma multiforme (rGBM). A significant improvement in median overall survival was observed in patients treated with LITT compared to SRS among patients with brain metastasis (12.8 versus 9.8 months; p<0.02) and was associated with a 15% reduction in risk of adverse events overall. The authors concluded that "there is no evidence that LITT can be used as a treatment of choice when compared to SRS," and note specifically there is a “lack of systematic data that were reported in our pooled studies.” The authors do indicate the use of LITT may have a role in lowering the risk of adverse events. The analysis was limited by inclusion of heterogeneous populations, small number of patients treated with LITT (n=39), and a lack of reporting on prior treatments. In particular, patients treated with SRS varied in their degree of radiosensitivity and prior radiation exposure, which may have influenced the higher rate of adverse events observed in this group.

Barnett (2016) conducted a SR and meta-analysis comparing LITT (8 studies; 77 patients) to open craniotomy (12 studies; 1036 patients) for the treatment of high-grade gliomas in or near areas of eloquence, with a focus on adverse events. Proportions of major complications occurred in 5.7% (95% CI: 1.8–11.6) and 13.8% (95% CI: 10.3–17.9) of patients treated via LITT and craniotomy, respectively. Studies were rated at high risk of bias due to lack of randomization and blinding. The analysis was also limited by heterogeneous patient...
populations (eg, age Karnofsky score, recurrent vs primary disease) and lack of reporting on health outcomes.

**Comparative Observational Studies**

Mohammadi (2019) conducted a multicenter retrospective review of survival outcomes in patients with deep seated newly diagnosed glioblastoma treated with upfront MR-guided LITT prior to chemo/radiotherapy (n=24; median age, 54 y; 50% male; 71% <70 yr) compared to a matched cohort of biopsy-only patients (n=24; median age, 64 yr; 58% male; 75% <70 yr).[12] Patients were matched based on age, gender, tumor location (deep versus lobar), and tumor volume. Median follow-up was 9.3 mo (range, 2 to 43 mo) and 14.7 mo (range, 2 to 41 mo) in LITT and biopsy-only cohorts, respectively. Overall median estimates of overall survival and progression-free survival in the LITT cohort was 14.4 and 4.3 mo compared to 15.8 and 5.9 mo for the biopsy-only cohort. Age <70 y and tumor volume <11 cm3 were identified as favorable prognostic factors for overall survival. The study was limited by its retrospective design, lack of randomization, small sample size, and short follow-up durations. Additionally, concurrent chemotherapy and radiotherapy regimens were not specified.

**Single-Arm Studies**

The Laser Ablation of Abnormal Neurological Tissue Using Robotic NeuroBlate System (LAANTERN) registry is an ongoing industry-sponsored, multicenter, multinational prospective registry of the NeuroBlate device enrolling patients with primary and metastatic brain tumors, epileptic foci, and movement disorders (NCT02392078). Rennert (2019) reported procedural safety outcomes for the first 100 patients enrolled in the LAANTERN registry (42% male, 86% white), including 48 and 34 patients with primary or metastatic intracranial tumors, respectively.[13] The majority of patients (81.2%) had undergone prior surgical or radiation treatment and received LITT for a single lesion (79%). The average length of intensive care and overall hospital stays were 38.1 and 61.1 hours, respectively. A total of 11 adverse events among 9 patients were observed. Five adverse events were attributed to energy deposition from laser ablation, including neurological deficits (n=2), postoperative seizures (n=2), and delayed intraparenchymal hemorrhage (n=1). One mortality occurring within 30 days of laser ablation was reported and was not attributed to LITT.

Kim (2020) reported 12-month survival and quality of life outcomes among 223 patients enrolled in the LAANTERN registry with primary (n=131) or metastatic (n=92) brain tumors who received treatment with the NeuroBlate device.[14] The majority of patients with primary tumors had high-grade glioma (n=90) and patients with metastatic disease had recurrent tumors (n=43) or radionecrosis (n=34). The one year estimated overall survival rate was 73% (95% CI, 65.3% to 79.2%), which was not found to be significantly different between primary or metastatic tumors (74.6% versus 70.7%, respectively). Quality of life assessments with the Functional Assessment of Cancer Therapy - Brain (FACT-Br) questionnaire did not meet the criteria for a clinically meaningful change (>10%) and EQ-5D questionnaires indicated an overall decline of 0.1 points from baseline.

Ahlulwalia (2018) reported results from the multicenter, prospective Laser Ablation After Stereotactic Radiosurgery (LAASR) study, which assessed the efficacy and safety of LITT as salvage treatment in patients with radiographic progression after SRS for brain metastasis.[15] Forty-two patients were enrolled, including 20 patients with recurrent brain tumors, 19 patients with biopsy-proven radiation necrosis, and three patients with no diagnosis. PFS rates for patients with recurrent tumors was 54% at 12 weeks and 62% at 26 weeks. Corresponding OS
rates were 71% at 12 weeks and 64.5% at 26 weeks. Of four tumor lesions that received total ablation, 3/4 achieved a complete response, compared to 0/8 that received subtotal ablation. Patient Karnofsky performance, quality of life, and neurocognitive scores did not change significantly over the duration of survival. Overall, 35/42 (83%) patients developed adverse events, including five cases of immediate LITT-related neurological complications and 14 surgery-related adverse events.

Patel (2016) conducted a retrospective analysis of patients who underwent MR-guided LITT with the Visualase system at a single center in the United States between 2010 and 2014.[16] The majority of patients (87/102) were treated for intracranial tumors. Fourteen (13.7%) developed new neurological deficits following treatment, of which nine achieved complete resolution within one month, one achieved partial resolution within one month, two had no resolution at most recent follow-up, and two died without resolution of symptoms. The authors concluded that LITT, albeit minimally invasive, must be used with caution as unintended thermal damage to critical and eloquent structures may occur despite MRI guidance.

Section Summary: Primary or Metastatic Brain Tumors

Evidence for the use of LITT in primary or metastatic brain tumors includes systematic reviews and meta-analyses, one retrospective matched-cohort study (in newly diagnosed glioblastoma comparing LITT to biopsy only), and several single-arm studies. Overall survival estimates ranged from 12.8 to 14.8 months. Among patients with metastatic tumors receiving LITT following prior SRS, overall survival rates have ranged between 72-76% at six months and 63-65% at 12 months. Systematic reviews comparing LITT to open craniotomy with resection or SRS suggest a reduced incidence of adverse events with LITT; however neurological deficits attributable to LITT-induced thermal damage have been observed despite concurrent MRI guidance. Studies are limited by high risk of bias, predominantly retrospective designs, small sample sizes, and population heterogeneity, with study subjects varying by performance status, lesion volume and location, extent of prior therapies, and extent of ablation. Prospective comparative studies in well-defined and -controlled patient populations are required to assess net health outcomes.

RADIATION NECROSIS

Clinical Context and Therapy Purpose

The purpose of LITT is to use a focused thermal therapy technique to ablate regions of cerebral radiation necrosis in symptomatic patients with an insufficient or intolerable response to medications, and to potentially avoid complications associated with alternative surgical interventions.

Populations

The population of interest is patients with symptomatic cranial radiation necrosis with insufficient response or intolerance to medication management. LITT is typically used when open surgery is contraindicated due to high risk of procedural morbidity and/or presence of comorbidities that precludes candidacy for open surgery.

Treatment-induced brain tissue necrosis (also referred to as cranial radiation necrosis or radionecrosis) is a serious delayed complication of cranial irradiation that typically develops after one to three years. Radiation necrosis is more likely to occur with high-dose fractionation and potentially with concurrent chemotherapy or use of radiosensitizers. The risk of radiation
necrosis following stereotactic radiosurgery (SRS) has been reported to be higher, with a steep dose-response relationship. Differentiating radiation necrosis from recurrent brain tumors via imaging can be difficult, as conventional structural MRI may reveal features that overlap with the typical radiographic appearance of high-grade primary or metastatic brain tumors. Biopsy may be required for a definitive diagnosis of radiation necrosis, particularly among patients who are symptomatic or with worsening radiographic findings over time.

Symptoms of radiation necrosis are dependent on the location of the lesion and may include focal neurologic deficits or more generalized signs and symptoms of increased intracranial pressure. Seizures are observed in approximately 20% of patients.

**Interventions**

The therapy being considered is LITT as an alternative to open craniotomy with resection or medication management. LITT is performed under real-time MRI guidance.

**Comparators**

The following therapies are currently being used to treat primary and metastatic brain tumors: surgical resection and medication management. Medications used in the management of radiation necrosis include corticosteroids and bevacizumab, a vascular endothelial growth factor (VEGF) inhibitor.

**Outcomes**

Outcomes of interest are symptom improvement, medication use, quality of life, treatment-related morbidity, overall survival (OS), and progression-free survival (PFS). Follow-up duration of at least 2-3 years is of interest for survival outcomes.

**Review of Evidence**

**Systematic Reviews**

The meta-analysis published by Chen (2021), described previously, included 168 (35.7%) patients with radiation necrosis (RN) who received LITT following prior treatment with SRS.[8] The local control rate for patients with RN at 6 and 12 months was 83.1% (95% CI, 68.4% to 91.8%) and 66.8% (95% CI, 49.1% to 80.8%), respectively, and was more satisfactory compared to patients with recurrent brain metastasis. OS was 83.1% versus 69.2% at six months and 66.8% versus 66.5% at 12 months for RN and recurrent brain metastasis groups, respectively. Pre-ablation biopsy, which can accurately diagnose RN, was not routinely performed in all analyzed studies, highlighting a major limitation of this meta-analysis given that it can be quite challenging to accurately distinguish RN from brain metastases based on radiographic evidence alone.

**Comparative Observational Studies**

Sujijantarat (2020) conducted a retrospective chart review comparing outcomes for patients with biopsy-confirmed radiation necrosis treated with LITT (n=25) or bevacizumab (n=13) at a single center between 2011 and 2018.[17] The LITT group had a significantly longer OS compared to bevacizumab (median 24.8 versus 15.2 months; p=0.003). Time to local recurrence was not statistically significant between groups (p=0.091), but trended longer in the LITT cohort. Among 13 patients with pre-treatment symptoms in the LITT group, nine (69%)
achieved symptom relief. Among 11 patients with pre-treatment symptoms in the bevacizumab group, 4 (36%) achieved symptom relief. No significant difference was noted between groups for the ability to wean off concurrent steroids. Given that only 50% of lesions treated with LITT were symptomatic compared to 80% of lesions treated with bevacizumab, the authors suggest that LITT treatment may be more successful before radiation necrosis lesions become symptomatic. The study is limited by its retrospective design, small samples size, and population heterogeneity.

Hong (2019) conducted a single-center retrospective chart review of patients treated with LITT or craniotomy for previously irradiated brain metastasis, including 42 patients with recurrent brain tumors and 33 patients with radiation necrosis (RN). Among the 33 RN patients, 15 received craniotomy and 18 received LITT, of which 20% and 38.9% received adjuvant post-operative bevacizumab, respectively. No significant differences for mean length of hospital stay, symptom improvement, ability to wean off steroids, or rate of perioperative complications were observed between LITT and craniotomy groups. Overall PFS for patients with RN was 73.2% and 86.7% at 24 months or patients treated with LITT and craniotomy, respectively. OS for patients with RN at 24 months was 64.6% for those receiving craniotomy and 63.2% for those receiving LITT. Study interpretation is limited by its retrospective nature and heterogeneity of prior and adjuvant treatments.

Single-Arm Studies

The LAASR study, described previously [Ahluwalia (2018)], included 19 patients with biopsy-confirmed radiation necrosis who received LITT following prior treatment with SRS for brain tumors. PFS and OS survival was 100% and 91%, respectively, at 12 weeks, and 100% and 82.1%, respectively, at 26 weeks. PFS was significantly higher at 12 weeks for patients with radiation necrosis compared to patients with recurrent tumors (p=0.016) but was not significantly different at 12 weeks (p=0.166). Similar trends were seen for OS in patients with radiation necrosis at 12 weeks (p=0.02) and 26 weeks (p=0.09). Thirty percent of subjects were able to stop or reduce steroid usage by 12 weeks after surgery. For patients with RN, regardless of whether a lesion was totally or subtotally ablated, LITT resulted in close to 100% lesion control and > 80% survival at 6 months. No significant differences in Karnofsky performance status, quality of life, or neurocognitive scores were detected between subgroups.

Section Summary: Radiation Necrosis

Evidence on the use of LITT in patients with radiation necrosis includes one meta-analysis, two nonrandomized comparative studies, and one single-arm study. Studies have reported improved local control and survival outcomes in patients with radiation necrosis compared to those with brain metastases. One study comparing LITT to bevacizumab suggested that LITT treatment may be more successful among patients before radiation necrosis lesions become symptomatic. One study comparing LITT to craniotomy did not report significant survival differences between groups. Studies are limited by retrospective designs, small sample sizes, population heterogeneity, and unclear relevance, as symptomatic status was not consistently reported. Prospective comparative studies in well-defined and -controlled patient populations are required to assess a net health outcome.

DRUG-RESISTANT EPILEPSY

Clinical Context and Therapy Purpose
The purpose of LITT is to use a focused thermal therapy technique to ablate epileptogenic foci when seizures have become drug-resistant or medication-related adverse events are intolerable, and to potentially avoid complications associated with alternative surgical interventions.

**Populations**

The population of interest is patients with drug-resistant or medication-intolerant epilepsy, defined as failure to achieve sustained seizure freedom despite adequate trials of two or more appropriately chosen and tolerated antiseizure medications, as specified by the International League Against Epilepsy (ILAE) Commission on Therapeutic Strategies consensus definition for drug resistant epilepsy.[19]

Epilepsy is diagnosed when an individual has unprovoked seizures. Primary seizure disorders include multiple subtypes that are recognizable by the degree and type of impairment of consciousness and motor capacity. Seizure disorders may be secondary to brain tumors or other space-occupying intracranial lesions such as congenital malformations, stroke, genetic syndromes, brain trauma, and cerebral infections. Mesial temporal lobe epilepsy, also known as complex partial seizures, is a focal epilepsy syndrome. The epileptogenic foci may present in the hippocampus, amygdala, or parahippocampal gyrus. The most common non-traumatic or non-infectious etiology of mesial temporal lobe epilepsy is hippocampal sclerosis. The associated neuronal loss is a partial explanation for the difficulties in achieving satisfactory seizure control with antiepileptic medication. Approximately one-third of patients with epilepsy do not achieve adequate seizure control with antiepileptic drugs.

Patients with an identifiable seizure focus that can be targeted to achieve seizure freedom are primary candidates for epilepsy surgery, but patients with multifocal or generalized epilepsy may also be considered.

**Interventions**

The therapy being considered is LITT as an alternative to open craniotomy with resection, stereotactic radiosurgery, or neurostimulation. LITT is performed under real-time MRI guidance.

**Comparators**

The following therapies are currently being used to treat medication-refractory epilepsy: open craniotomy with resection, stereotactic radiosurgery (SRS), vagus nerve stimulation, and responsive cortical neurostimulation. Surgical treatment may be considered in instances where seizures have proven refractory to medical management and when the frequency and severity of the seizures significantly diminish quality of life.

**Outcomes**

Outcomes of interest are symptom improvement, change in disease status, quality of life, hospitalizations, medication use, treatment-related morbidity, and disease-specific survival. Key outcome measures are summarized in Table 1.
Table 1. Epilepsy Outcome Measures

<table>
<thead>
<tr>
<th>Outcome Domain</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptom Improvement</td>
<td>Change in seizure frequency (&gt;50% reduction considered clinically meaningful)</td>
</tr>
<tr>
<td>Change in Disease Status</td>
<td>Time to cessation of seizures; Postoperative outcome status, as measured by the Engel classification:[20]</td>
</tr>
<tr>
<td></td>
<td>• Class I: Free of disabling seizures</td>
</tr>
<tr>
<td></td>
<td>• Class IA: Completely seizure free since surgery</td>
</tr>
<tr>
<td></td>
<td>• Class II: Rare disabling seizures</td>
</tr>
<tr>
<td></td>
<td>• Class III: Worthwhile improvement</td>
</tr>
<tr>
<td></td>
<td>• Class IV: No worthwhile improvement</td>
</tr>
<tr>
<td>Quality of Life</td>
<td>QOLIE-89 or QOLIE-31 multi-scale questionnaires (higher scores indicate improved health outcomes); eligibility to drive</td>
</tr>
<tr>
<td>Treatment-related Morbidity</td>
<td>Neuropsychological and neurocognitive testing</td>
</tr>
<tr>
<td>Disease-specific Survival</td>
<td>Incidence of SUDEP</td>
</tr>
</tbody>
</table>

SUDEP: sudden unexpected death in epilepsy; QOLIE: Quality of Life in Epilepsy questionnaire.

Follow-up duration of at least two years is of interest to evaluate the effect of the procedure when compared to resection or neurostimulation. Follow-up durations of 2-3 years are appropriate when compared to SRS, due its known latency for seizure reduction or remission. Rarely, a transient increase in seizure frequency and severity may be observed following surgical interventions. Therefore, time to cessation of seizures and proportion of patients with increased seizure frequency represent additional outcomes of interest.

Review of Evidence

Systematic Reviews

Barot (2021) published a systematic review (SR) with meta-analysis of outcomes following LITT for the treatment of drug-refractory epilepsy (DRE), comparing outcomes between temporal, extratemporal epilepsies and hypothalamic hamartoma.[21] Twenty-eight studies (N=559) were included. The overall prevalence of Engel class I outcome was 56% (95% CI 0.52% to 0.60%). Hypothalamic hamartomas (HH) patients had the highest seizure freedom rate of 67% (95% CI 0.57% to 0.76%) and outcome was overall comparable between mesial temporal lobe epilepsy (mTLE) (56%, 95% CI 0.50% to 0.61%) and extratemporal epilepsy (50% 95% CI 0.40% to 0.59%). The postoperative adverse event rate was 19% (95% CI 0.14% to 0.25%) and the most common adverse event was visual field deficits. The reoperation rate was 9% (95% CI 0.05% to 0.14%), which included repeat ablation and open resection.

Kohlhase (2021) published a SR with meta-analysis evaluating outcomes and complications following temporal lobe MRgLITT, RFA, and conventional surgical approaches (i.e., anterior temporal lobe resection [ATL] or selective amygdalohippocampectomy [sAHE]) for the treatment of drug-refractory mesial temporal lobe epilepsy (mTLE).[22] Forty-three studies (13 MRgLITT, 6 RFA, and 24 surgery studies) of 554, 123, 1504, and 1326 patients treated by MRgLITT, RFA, ATL, or sAHE, respectively, were included in the review. Engel Class I (Engel-I) outcomes were achieved after MRgLITT in 57% (315/554, range = 33.3%-67.4%), RFA in 44% (54/123, range = 0%-67.2%), ATL in 69% (1032/1504, range = 40%-92.9%), and sAHE in 66% (887/1326, range = 21.4%-93.3%). No significant difference in seizure outcome between MRgLITT and RFA (Q = 2.74, p=0.098) was found, however, ATL and sAHE were both superior to MRgLITT (ATL: Q = 8.92, p=0.002; sAHE: Q = 4.33, p=0.037) with better outcomes.
in patients at follow-up of 60 months or more. The rate of major complications was 3.8% for MRgLITT, 3.7% for RFA, 10.9% for ATL, and 7.4% for sAHE; none of these frequencies were statistically significantly different. While the severity of cognitive impairment was not evaluated across treatment groups directly, the authors note that cognitive impairment following intervention appears to increase with the invasiveness of the respective intervention. The authors conclude “patients undergoing MRgLITT may experience fewer major complications compared to ATL or sAHE and might have a more beneficial neuropsychological outcome.”

Kerezoudis (2021) published a SR with meta-analysis aimed at quantifying the relationship of LITT ablation volume with postoperative outcomes in temporal lobe epilepsy (TLE). A total of 13 studies (551 patients) were analyzed. Meta-regression of seizure freedom rate for the overall cohort and mesial temporal sclerosis (MTS) subset (n = 384) was performed adjusting for overall ablation volume as well as percentage of hippocampal and amygdala ablation. Overall seizure freedom rate was 58% (95% confidence interval [CI], 54%-62%) and was not significantly associated with total ablation volume (p = 0.42), hippocampal ablation (p = 0.67), or amygdala ablation (p = 0.33). Seizure freedom rate for patients with MTS was 66% (95% CI, 58%-74%) and was also not found to be significantly associated with total ablation volume (p = 0.15), hippocampal ablation (p = 0.73), or amygdala ablation (p = 0.43). Overall complication rate was 17% (95% CI, 13%-22%).

Wang (2021) published a SR of data on LITT, stereotactic radiosurgery (SRS), radiofrequency thermocoagulation (RF-TC), and focused ultrasound for the treatment of mesial (medial) temporal lobe epilepsy (mTLE). Data from 19 publications were included with 1094 patients (LITT: 434, SRS: 81, RF-TC: 402, Cortico-amygdalohippocampectomy (CAH): 153, and selective amygdalohippocampectomy (SelAH): 24). At six months postoperatively, LITT (9/19) Engel I outcomes ranged from 52% to 80%. Seizure freedom was similar between LITT studies and to rates achieved by CAH and SelAH, however, no direct comparisons were available. Common complications included transient postprocedure headaches (LITT: 0.4%-27%, SRS: 15%-70%, and RF-TC: 23%) and visual field deficits (VFDs) (LITT: 3%-40%, SRS: 34%-50%, and RF-TC: 2%-5%)

Brotis (2021) conducted a meta-analysis to estimate the efficacy of LITT for mTLE. Sixteen retrospective case series published between 2012 and 2019 representing 575 patients (range, 1-231) were identified. Overall, seizure freedom was achieved in 54.7% (95% CI 50.6% to 58.8%; I²=18.7%) of patients undergoing LITT with a median follow-up duration of 18 months (IQR, 12-26 months). Sensitivity analyses yielded similar results. Four studies representing 150 patients indicated that the prevalence of Engel Class IA outcomes decreased with time, estimated at 64.2%, 46.9%, and 42.4% at 12-, 24-, and 36-month follow-up, respectively. The overall quality of evidence was regarded as ‘very low’ according to GRADE recommendations, with only 4 studies included more than 20 patients. The authors concluded that while mTLE resective surgeries are invasive and irreversible, they offer better seizure control rates, with previously reported seizure-free rates ranging from ranging from 60% to 90% for mTLE.

Grewal (2019) published a SR and meta-analysis comparing MR-guided LITT versus SRS for medically intractable temporal lobe epilepsy (TLE). A total of 19 studies published between 2008 and 2018 representing 404 patients (range, 5-58) were identified, including 9 retrospective studies on LITT (n=239). The overall seizure freedom rate was not found to be significantly different between LITT (50%; 95% CI, 44% to 56%) and SRS (42%; 95% CI, 27% to 59%; p=0.39), nor was it significantly different for patients with lesional conditions (62% [95% CI, 48% to 74%] versus 50% [95% CI, 37% to 64%]; p=.23). While LITT was associated

MED177 | 12

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
with a significantly lower procedural complication rate (20% versus 26%; p=0.06), reoperation rates were not significantly different (15% versus 27%; p=0.31). The authors noted that the quality of evidence was low and that large-scale comparative studies directly comparing LITT and SRS are required to validate findings.

Xue (2018) reported postoperative outcomes for MR-guided LITT in the treatment of drug-resistant epilepsy.\[27\] Sixteen nonrandomized studies published between 2014 and 2018 representing 269 patients (range, 5-30) were included in the meta-analysis. The prevalence of Engel Class I, II, III, and IV outcomes was 61%, 12%, 16%, and 15%, respectively. The prevalence of postoperative complications was 24% (95% CI, 16% to 32%). Interpretation of outcomes is limited by small study size and short follow-up durations (range, 7 days - 51 months).

Comparative Observational Studies

Hale (2019) reported postsurgical outcomes in 26 pediatric patients with insular epilepsy treated with LITT (n=14) or open resection (n=12).\[28\] Mean follow-up was 2.43 years. Engel Class I outcomes were achieved in 43% of patients treated with LITT compared to 50% who underwent open insular resection at one year post-surgery. Postoperative complications occurred in six patients treated with LITT and seven patients treated with resection, all of which resolved within 3-4 months. The authors conclude that further studies are needed to determine the noninferiority of LITT with respect to resection in terms of complication rates and seizure freedom, especially in cases of cortical dysplasia that may involve extensive regions of the brain.

Petito (2018) published a retrospective, single center analysis of 100 consecutive neurosurgeries performed between 2013 and 2015 in patients with drug-resistant epilepsy, representing 33 LITT procedures and 21 open resections with mean follow-up durations of 21.7 and 21.3 months, respectively.\[29\] A discrete lesion was radiographically identified in 85% of patients treated with LITT and 65% of patients treated with resection. The mean postoperative hospital length of stay was significantly shorter for LITT compared to resection (1.18 versus 3.43 days; p=.0002). Patients treated with resection were significantly younger, with a mean age of 35.4 years (p=.001). At 12 months, seizure freedom was achieved in 56.3% (95% CI, 39.3% to 71.8%) and 60% (95% CI, 38.7% to 78.12%) of patients treated with LITT and resection, respectively (p=0.79). Among patients with focal lesions, the seizure freedom outcomes were not significantly different between groups (p=0.21). For nonlesional patients, LITT treatment trended towards a better outcome, but did not achieve statistical significance (p=0.05). Study interpretation is limited by small sample size, retrospective analysis, and population heterogeneity.

Single-Arm Studies

Landazuri (2020) reported one-year outcomes following LITT of epileptogenic foci with the NeuroBlate system in patients with drug resistant epilepsy enrolled in the previously described LAANTERN registry (see Rennert [2019]).\[13, 30\] Engel Class I outcomes were achieved in 27/42 (64.3%; 95% CI, 48.0% to 78.5%) patients at one year. No significant difference was observed in patients with mesial TLE (70.8%) versus other etiologies. Five adverse events were reported, with one categorized as serious. Median baseline QOLIE-31 was 51.7 (range, 8.7 to 77.3). Median scores increased by 14.1 points reflecting a 72.4% improvement (95% CI, 52.8% to 87.3%) in quality-of-life measures. However, the total score change was not statistically significant (p=.2173). Seizure worry and social functioning sub-scores were
considered statistically significant (p=0.0219 and p=0.0175, respectively). The authors note that the primary success of LITT remains in well localized lesions/localizations, such as those seen in mesial TLE/mesial temporal sclerosis (MTS), cortical dysplasia, and hypothalamic hamartoma.

Wu (2019) published the results of a multicenter, retrospective cohort study of 234 patients with drug-resistant mTLE who underwent LITT between 2011 and 2017. At both one and two years after LITT, 58% of patients achieved Engel I outcomes. Engel I outcomes were associated with ablations involving more anterior, medial, and inferior temporal lobe structures, which tended to involve greater amygdalar volume. Presence or absence of hippocampal sclerosis did not have a significant effect on seizure outcomes. Overall, Engel I or II outcomes were achieved by 76.9% patients at the time of last follow-up. A total of 42 complications were observed in 35 patients, of which 34 persisted at last follow-up.

Section Summary: Drug-Resistant Epilepsy

The evidence for the use of LITT in drug-resistant epilepsy includes several large systematic reviews (N>500 patients treated with LITT) and meta-analyses, two nonrandomized comparative studies, and two single-arm studies. Meta-analyses have reported seizure freedom rates ranging from 50 to 61% and six months postoperatively, Engel I outcomes have been observed between 52% to 80%. Studies comparing outcomes following LITT to open resection or radiofrequency ablation have reported comparable outcomes in patients with drug-refractory temporal lobe epilepsy. While one systematic review found lower rates of Engel-1 outcomes with LITT compared to conventional surgical intervention, this review also reported LITT may be associated with fewer major complications and improved cognitive outcomes compared to open approaches. Total quality of life scores reported in the ongoing LAANTERN registry study increased by 72.4%, however this change did not reach statistical significance (p=0.2173).

PRACTICE GUIDELINE SUMMARY

AMERICAN ASSOCIATION OF NEUROLOGICAL SURGEONS

In September 2021, the American Association of Neurological Surgeons (AANS) and Congress of Neurological Surgeons (CNS) Joint Section on Tumors issued a position statement regarding the use of LITT for brain tumors and radiation necrosis. The statement concludes that "LITT is an appealing option because it offers a method of minimally invasive, targeted thermal ablation of a lesion with minimal damage to healthy tissue. There is a growing body of evidence to demonstrate that LITT is an effective and well tolerated cytoreductive option for treatment of [newly diagnosed glioblastoma multiforme (GBM), recurrent GBM, and primary or recurrent brain metastases.] Intracranial LITT is also an effective option for addressing radiation necrosis with an overall reduction in steroid dependence for these patients. Especially in instances where the therapeutic window is narrowed such that craniotomy is not a viable option, LITT can play an important role in treatment for glioma or metastatic brain cancer."

AMERICAN SOCIETY FOR RADIATION ONCOLOGY

The American Society for Radiation Oncology (ASTRO) clinical practice guideline on radiotherapeutic and surgical management for newly diagnosed brain metastases (2012) does not address the use of LITT.
AMERICAN SOCIETY FOR STEREOTACTIC AND FUNCTIONAL NEUROSURGERY

In September 2021, the American Society for Stereotactic and Functional Neurosurgery (ASSFN) issued a position statement on the use of LITT in drug-resistant epilepsy.[34] The statement recommends consideration of MR-guided LITT (MRgLITT) as a treatment option when all of the following criteria are met:

- "Failure to respond to, or intolerance of, at least 2 appropriately chosen medications at appropriate doses for disabling, localization-related epilepsy AND
- Well-defined epileptogenic foci or critical pathways of seizure propagation accessible by MRgLITT."

CONGRESS OF NEUROLOGICAL SURGEONS

The Congress of Neurological Surgeons (CNS) guidelines for the treatment of adults with metastatic brain tumors (2019) state that "there is insufficient evidence to make a recommendation regarding the routine use of laser interstitial thermal therapy (LITT), aside from use as part of approved clinical trials."[35]

NATIONAL COMPREHENSIVE CANCER NETWORK

The National Comprehensive Cancer Network (NCCN) clinical practice guidelines for central nervous system cancers (v.2.2021) states that MRI-guided laser interstitial thermal therapy "may be considered for patients who are not surgical candidates (craniotomy or resection). Potential indications include relapsed brain metastases and radiation necrosis." (Category 2B)[36]

SUMMARY

Studies comparing laser interstitial thermal therapy (LITT) to open resection or radiofrequency ablation have found comparable outcomes in the treatment of drug-resistant epilepsy. In addition, there is evidence that this treatment approach may be associated with fewer major complications and improved cognitive outcomes than open approaches. Evidence-based clinical practice guidelines recommend LITT for the treatment of drug-resistant epilepsy when criteria are met. Therefore, LITT for the treatment of drug-resistant epilepsy may be considered medically necessary when there is documentation of disabling seizures despite use of two or more antiepileptic drug regimens (i.e., medically refractory epilepsy) and there is a well-defined epileptogenic focus of seizure propagation in the temporal lobe or hypothalamus. The evidence for the use of laser interstitial thermal therapy (LITT) for all other neurological indications is limited by retrospective designs, small sample sizes, and population heterogeneity. In addition, neurological deficits attributable to LITT-induced thermal damage have been observed despite concurrent MRI guidance. The evidence is insufficient to determine that the use of laser interstitial thermal therapy (LITT) results in an improvement in the net health outcome for these patients. Therefore, laser interstitial thermal therapy (LITT) is considered investigational for all other neurological indications, including but not limited to treatment of primary or metastatic brain tumors or radiation necrosis.
REFERENCES


11. GH Barnett, JD Voigt, MS Alhuwalia. A Systematic Review and Meta-Analysis of Studies Examining the Use of Brain Laser Interstitial Thermal Therapy versus Craniotomy for the Treatment of High-Grade Tumors in or near Areas of Elocution: An Examination of the Extent of Resection and Major Complication Rates Associated with Each Type of Surgery. *Stereotact Funct Neurosurg.* 2016;94:164-73. PMID: 27322392


35. JB Elder, BV Nahed, ME Linskey, JJ Olson. Congress of Neurological Surgeons Systematic Review and Evidence-Based Guidelines on the Role of Emerging and Investigational Therapies for the Treatment of Adults With Metastatic Brain Tumors. *Neurosurgery.* 2019;84:E201-E03. PMID: 30629215


**CODES**

**NOTE:** Coding note would be entered here to explain something related to the codes listed in the policy.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>61736</td>
<td>Laser interstitial thermal therapy (LITT) of lesion, intracranial, including burr hole(s), with magnetic resonance imaging guidance, when performed; single trajectory for 1 simple lesion</td>
</tr>
<tr>
<td></td>
<td>61737</td>
<td>Laser interstitial thermal therapy (LITT) of lesion, intracranial, including burr hole(s), with magnetic resonance imaging guidance, when performed; multiple trajectories for multiple or complex lesion(s)</td>
</tr>
</tbody>
</table>

*These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.*
<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>64999</td>
<td>Unlisted procedure, nervous system</td>
</tr>
<tr>
<td>HCPCS</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

*Date of Origin: December 2021*
Ovarian, Internal Iliac, and Gonadal Vein Embolization, Ablation, and Sclerotherapy

Effective: August 1, 2021

Next Review: April 2022
Last Review: June 2021

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Embolization involves occlusion of blood flow through the ovarian, internal iliac, and gonadal veins with coils, foam, or a chemical sclerosant as a treatment of pelvic congestion syndrome or varicoceles.

MEDICAL POLICY CRITERIA

Note: This policy does not address surgical ligation of the spermatic vein(s) or uterine artery embolization.

I. Embolization, ablation, and sclerotherapy of ovarian veins, internal iliac veins, or gonadal veins is considered investigational for the treatment of pelvic congestion syndrome and varicoceles.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Varicose Vein Treatment, Surgery, Policy No. 104
BACKGROUND

Enlarged ovarian and internal iliac veins can lead to pelvic congestion syndrome in women, and enlarged gonadal and internal iliac veins can lead to a varicoceles in men. Each are discussed separately below.

PELVIC CONGESTION SYNDROME

Pelvic congestion syndrome (PCS), also called pelvic venous incompetence, is a rare condition characterized by chronic pelvic pain. Although this condition is primarily found in women it can also be found in men. PCS is often aggravated by standing for long periods of time, and often manifests during or after pregnancy. The syndrome is thought to be associated with dilated and refluxing incompetent pelvic veins, similar to what happens in varicose veins of the legs. However, the cause of PCS is unclear. Furthermore, there are no definitive diagnostic criteria for PCS. Instead the diagnosis is generally based on a combination of symptoms, tenderness on physical exam, and documentation of pelvic vein dilation or incompetence after excluding all other causes for the nonspecific findings. Although imaging may show vein dilation or incompetence, these findings are common nonspecific findings and therefore no diagnostic.

There is no standard treatment approach for PCS, and the optimum treatment is unknown. Instead, therapy is individualized and based on symptoms. Medical therapy is generally the first line of treatment, as it is low risk and non-invasive. Other methods, such as embolization has been proposed as an alternative to surgical treatment for patients who fail medical therapy with analgesics. Embolization therapy involves the occlusion of blood flow through the ovarian and internal iliac veins with coils, glue, or chemical sclerosants. The internal iliac veins may be treated at the same time or a later date to prevent recurrence.

VARICOCELES

A varicocele is the dilation of the pampiniform plexus of the gonadal veins. Varicocele’s are present in 15 to 20% of post-pubertal males, and generally get larger over time. Most varicoceles occur in the left hemiscrotum because the left gonadal vein is one of the longest veins in the body and it enters the left renal vein at a perpendicular angle increasing pressure which can dilate the veins and cause incompetence of the valves, similar to what happens in varicose veins of the legs. Although varicoceles on the left are more common, bilateral varicoceles can occur; however, this could be caused by a possible underlying pathology warranting more investigation. Symptoms of a varicocele include dull, aching, left scrotal pain, which is often aggravated by standing for long periods of time, testicular atrophy, and decreased fertility. Although there are no clear guidelines regarding the established treatment for varicoceles, surgical ligation is the preferred first-line treatment.

EVIDENCE SUMMARY

The primary beneficial outcomes of interest for treatments of pelvic pain in both men and woman are symptom reduction and improvement in the ability to function. These are subjective outcomes that are typically associated with a placebo effect. Therefore, data from adequately powered, randomized controlled trials (RCTs) with sufficient long-term follow-up are required to control for the placebo effect, determine its magnitude, and to determine whether any treatment effect from provides a significant advantage over placebo or other treatment options.
TREATMENT FOR PELVIC CONGESTION SYNDROME

Health Technology Assessments

In 2016, Champaneria published a health technology assessment from the National Institute for Health Research that examined the diagnosis and treatment of pelvic vein incompetence and chronic pelvic pain in women.[1] Forty studies were included in the review; six association studies, ten studies involving ultrasound, two studies involving magnetic resonance venography, 21 case series, and one poor-quality randomized trial of embolization. The authors found that there were no consistent diagnostic criteria for pelvic congestion syndrome (PCS). Although the studies have showed associations between chronic pelvic pain (CPP) and pelvic vein incompetence (PVI), the prevalence of PVI ranged widely. The authors identified that transvaginal ultrasound with doppler and magnetic resonance venography are both useful screening methods; however, there is limited data on the accuracy of these methods for PCS. Finally, although the research showed embolization provides symptomatic relief in the majority of women, these studies were small case series. The authors concluded that more research is needed to determine what the diagnostic criteria for PCS are, and the efficacy of embolization as a treatment for PCS.

Systematic Reviews

A 2016 systematic review by Mahmoud identified 20 case series (total N=1081 patients) who underwent vein embolization for pelvic congestion syndrome.[2] The authors did not require any particular diagnostic criteria for pelvic congestion syndrome. The length of follow-up in the studies ranged from one month to six years. Seventeen studies (n=648 patients) reported the proportion of patients who reported symptom relief. Overall, 571 (88.1%) patients reported short-term symptom relief and 77 (11.9%) reported little or no relief. Seventeen studies (n=721 patients) reported symptom relief at 12 months. A total of 88.6% had symptom improvement and 13.4% reported little or no relief. Only one study used a comparison group, but patients in it received conservative treatment because they were ineligible for vein embolization therapy, so outcomes after the two interventions cannot be compared.

A systematic review by Daniels (2016) assessed the effectiveness of sclerotherapy or embolization for the treatment of chronic pelvic pain.[3] The review included 21 case series and one poor-quality randomized trial. Due to the overall low quality and heterogeneity of the studies, a meta-analysis was not performed. However, the authors reported that approximately 75% of women who underwent embolization experienced early pain relief. Adverse events noted included, transient pain following foam embolization and a small (<2%) risk of coil migration.

In 2015 Hansrani published a systematic review that evaluated the effectiveness of transvenous occlusion as a treatment of chronic pelvic pain.[4] Thirteen studies were included comprising 866 women. The authors noted that all 13 studies were of poor methodological quality, and most studies did not use objective outcome measures or have consistent follow-up of outcomes. Studies on embolization for treatment of PCS were rated as poor due to lack of randomization and control groups, unclear patient selection criteria, and heterogeneous outcome measures that did not permit between-study comparison or estimates of overall treatment effects. There was one RCT included in the review, in which embolization resulted in significantly better pain reduction than hysterectomy, but the study also had significant limitations, including but not limited to, the randomization protocol was not described, and the hysterectomy patients (bilateral compared to unilateral salpingo-oophorectomy) were not

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.
Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
blinded to their treatment allocation, small sample size limits the ability to rule out the role of chance as an explanation of study findings, and a discrepancy between reported outcomes in text and data tables. The authors recommended that more high quality studies are needed that compare embolization, with other treatments, including surgical treatments, hormonal therapy, and other noninvasive treatments.

**Randomized Controlled Trials**

A randomized, prospective trial by Guirola (2018) compared the safety and efficacy of embolization with vascular plugs (VP) or fibered platinum coils (FPC) in women with pelvic congestion syndrome. Patients were enrolled (N=100) and randomly assigned to each treatment group via block randomization (N=50). Diagnosis of pelvic congestion syndrome was accomplished through a symptom screening questionnaire followed by an ultrasound study. Patients with 3 or more positive symptom responses advanced to the ultrasound screening, and patients with pelvic veins >6 mm in diameter and/or venous reflux or dilated midline communicating veins were advanced to randomization. Follow-up screening occurred at 1, 3, 6, and 12 months. The primary outcome was clinical success assessed subjectively through patient responses regarding relief of symptoms and pain scores assessed with the visual analog scale. Clinical success was achieved in 89.7% of the FPC group and 90.6% of the VP group. Improvement in visual analog scale pain scores at the end of 12 months was 90.2% overall and improvement was seen in 95.9% of the FPC group and 96% of the VP group. A total of 11 (22%) complications were seen in the FPC group and 5 (10%) in the VP group. Minor adverse events included access site hematoma and ovarian vein extravasation. Device migrations were considered major complications. A major limitation in the study is the significant difference in age and pre-treatment visual analog scale pain score between groups, both of which were higher in the VP group despite randomization.

**Nonrandomized Studies**

The remainder of the published literature regarding the clinical outcomes of embolization therapy consists of nonrandomized studies, case series, and retrospective reviews. Collectively, conclusions concerning safety and effectiveness cannot be reached from these studies due to significant limitations in the data, including but not limited to:

- Lack of established diagnostic criteria for pelvic congestion syndrome. Without consistent criteria for patient selection it is unknown which patients are most likely to benefit, or not benefit, from treatment. Furthermore, it is unknown how results from the various case series can be applied to the overall population of patients with this condition.
- Lack of randomization and comparison groups. Failure to randomize patients to different treatment groups may introduce bias on the part of both the study participant and researchers in favor of the new technology. As noted above, for pain treatments, a comparator (preferably sham treatment) is necessary, in order to guard against this bias and to distinguish treatment from placebo effects.
- Retrospective design and failure to control for other treatments. Retrospective study designs do not allow for control of co-treatments or confounding factors that may influence results. This design may also introduce bias to interpretation of results. Control for additional factors, such as other medical therapies, is necessary to isolate treatment response to embolization therapy.
- Failure to define relevant study endpoints. Bias may also be introduced by failure to define study endpoints and treatment success prior to commencement of the study.
Adverse Effects

The following adverse effects associated with embolization of the uterine and internal iliac veins, though uncommon, have been reported in the literature.\cite{6,14}

- Embolization of coils to the pulmonary circulation
- Embolization of coils to the renal circulation
- Accidental embolization of glue fragments
- Perforations of the ovarian vein with extravasation of contrast
- Transient cardiac arrhythmia

Treatment of Varicoceles

Systematic Reviews

Belczak (2021) published a systematic review regarding semen parameter improvement after varicocele coil embolization.\cite{29} There were six retrospective studies and two observational studies included involving 701 patients where semen concentration and motility were the primary outcomes. The authors concluded that semen concentration was improved significantly in all five studies using that outcome and semen motility was significantly improved in seven studies. This review is limited by a small number of studies and no randomized or comparative studies being included.

In 2012 Kroese published results from a systematic review and meta-analysis that examined the effect of treatment, surgery or embolization, for varicoceles in subfertile men.\cite{30} Ten studies were included in the review, which comprised 894 men. The authors concluded that there is evidence to suggest treatment improves a couple’s chance of pregnancy; however, findings are inconclusive. Furthermore, the available evidence is of low quality and limited to men from couples with subfertility problems. Therefore further research is needed to determine the efficacy of treatment, surgery or embolization, for the treatment of varicoceles.

Randomized-Controlled Trials

No randomized controlled trials have been published comparing embolization therapy for the treatment of varicoceles to an alternative or sham/placebo treatment. Randomized controlled trials are especially needed in situations such as this where the primary symptom is pain, a subjective outcome for which a placebo response to treatment is likely.

Nonrandomized studies

The remainder of the published literature regarding the clinical outcomes of embolization therapy consists of case series and retrospective reviews.\cite{31-48} Collectively, conclusions concerning safety and effectiveness cannot be reached from these studies due to significant limitations in the data.

PRACTICE GUIDELINE SUMMARY

PELVIC CONGESTION SYNDROME

American Congress of Obstetricians and Gynecologists
No relevant policy positions on embolization for treating pelvic congestion syndrome were identified on the American Congress of Obstetricians and Gynecologists (ACOG) website.\[49]\n
**Society for Vascular Surgery (SVS) and the American Venous Forum**

The 2011 Society for Vascular Surgery (SVS) and the American Venous Forum (AVF) guidelines for the care of patients with varicose veins and associated chronic venous diseases provided a Grade 2B recommendation in favor of coil embolization, plugs, or transcatheter sclerotherapy for treatment of PCS. A Grade 2B recommendation is defined as a weak recommendation based on medium quality evidence.\[50]\n
**SUMMARY**

There is not enough research to show that embolization, ablation, or sclerotherapy improves long term health outcomes for people with pelvic congestion syndrome or varicoceles, compared to other forms of therapy. Therefore, embolization, ablation, or sclerotherapy of ovarian veins, internal iliac veins, or gonadal veins are considered investigational for the treatment of pelvic congestion syndrome or varicoceles.

**REFERENCES**


51. Bluecross BlueShield Association Medical Policy Reference Manual "Ovarian and Internal Iliac Vein Embolization as a Treatment of Pelvic Congestion Syndrome." Policy No. 4.01.18

### CODES

**NOTE:** There are no specific codes for ovarian and internal iliac vein embolization; however, the following codes may be used:

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>36012</td>
<td>Selective catheter placement, venous system: second order or more selective, branch (eg, left adrenal vein, petrosal sinus)</td>
</tr>
<tr>
<td></td>
<td>37241</td>
<td>Vascular embolization or occlusion, inclusive of all radiological supervision and interpretation, intra procedural roadmapping, and imaging guidance necessary to complete the intervention; venous, other than hemorrhage (eg, congenital or acquired venous malformations, venous and capillary hemangiomas, varices, varicoceles)</td>
</tr>
<tr>
<td>Codes</td>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>75894</td>
<td>Transcatheter therapy, embolization, any method, radiological supervision and interpretation</td>
</tr>
</tbody>
</table>

**HCPCS**: None

*Date of Origin: October 2005*
56625, 56800, 56805, 57106, 57110, 57291, 57292, 57295, 57296, 57335, 57426, 58353, 58356, 58563, C1813, C2622, L8600 Use code 17999 to request laser hair removal. Gender affirming surgical interventions for gender dysphoria require pre-authorization. Codes for specific procedures might also be listed as requiring pre-authorization in other medical policies, including but not limited to:

Abdominoplasty -15830
Breast Reconstruction -19316, 19318, 19325, 19350, L8600
Blepharoplasty and Brow Lift -15820, 15821, 15822, 15823, 67900, 67901, 67902, 67903, 67904, 67906, 67908, 67909, 67950
Chin Implants -21120, 21121, 21122, 21123, 21209
Collagen Injections -11950, 11951, 11952, 11954
Endometrial Ablation -58353, 58356, 58563
Panniculectomy -15830
Rhinoplasty -30400, 30410, 30420, 30430, 30435, 30450

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.