# Thermography and Temperature Gradient Studies

**Policy Number:** 6.01.12  
**Origination:** 11/2002  
**Last Review:** 11/2020  
**Next Review:** 11/2021

Blue KC has developed medical policies that serve as one of the sets of guidelines for coverage decisions. Benefit plans vary in coverage and some plans may not provide coverage for certain services discussed in the medical policies. Coverage decisions are subject to all terms and conditions of the applicable benefit plan, including specific exclusions and limitations, and to applicable state and/or federal law. Medical policy does not constitute plan authorization, nor is it an explanation of benefits.

When reviewing for a Medicare beneficiary, guidance from National Coverage Determinations (NCD) and Local Coverage Determinations (LCD) supersede the Medical Policies of Blue KC. Blue KC Medical Policies are used in the absence of guidance from an NCD or LCD.

## Policy

Blue Cross and Blue Shield of Kansas City (Blue KC) will not provide coverage for thermography or temperature gradient studies. These are considered investigational.

### When Policy Topic is covered

Not Applicable

### When Policy Topic is not covered

The use of all forms of thermography is considered investigational. The use of temperature gradient studies is considered investigational.

## Description of Procedure or Service

<table>
<thead>
<tr>
<th>Populations</th>
<th>Interventions</th>
<th>Comparators</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individuals:</strong></td>
<td><strong>Interventions of interest are:</strong></td>
<td><strong>Comparators of interest are:</strong></td>
<td><strong>Relevant outcomes include:</strong></td>
</tr>
</tbody>
</table>
| • With an indication for breast cancer screening or diagnosis | Thermography | Mammography | • Overall survival  
| | | | • Disease-specific survival |
Thermography is a noninvasive imaging technique that measures temperature distribution in organs and tissues. The visual display of this temperature information is known as a thermogram. Thermography has been proposed as a diagnostic tool for treatment planning and for evaluation of treatment effects for a variety of conditions.

**Summary of Evidence**
For individuals who have an indication for breast cancer screening or diagnosis who receive thermography, the evidence includes diagnostic accuracy studies and systematic reviews. Relevant outcomes are overall survival, disease-specific survival, and test validity. Using histopathologic findings as to the reference standard, a series of systematic reviews of studies have evaluated the accuracy of thermography to screen and/or diagnose breast cancer and reported wide ranges of sensitivities and specificities. To date, no study has demonstrated whether thermography is sufficiently accurate to replace or supplement mammography for breast cancer diagnosis. Moreover, there are no studies on the impact of thermography on patient management or health outcomes for patients with breast cancer. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have musculoskeletal injuries who receive thermography, the evidence includes diagnostic accuracy studies, a longitudinal prospective study,
and a systematic review. Relevant outcomes are test validity, symptoms, and functional outcomes. A systematic review of studies on thermography for diagnosing musculoskeletal injuries found moderate levels of accuracy compared with other diagnostic imaging tests. There is a lack of a consistent reference standard. This evidence does not permit conclusions as to whether thermography is sufficiently accurate to replace or supplement standard testing. Moreover, there are no high-quality or randomized studies on the impact of thermography on patient management or health outcomes for patients with musculoskeletal injuries. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have temporomandibular joint disorder who receive thermography, the evidence includes a systematic review. Relevant outcomes are test validity, symptoms, and functional outcomes. A systematic review of studies on thermography for diagnosing temporomandibular joint disorder found a wide variation in accuracy compared to other diagnostics. There is a lack of a consistent reference standard. The evidence does not permit conclusions as to whether thermography is sufficiently accurate to replace or supplement standard testing. Moreover, there are no studies on the impact of thermography on patient management or health outcomes for patients with the temporomandibular joint disorder. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have miscellaneous conditions (eg, herpes zoster, pressure ulcers, diabetic foot) who receive thermography, the evidence primarily includes diagnostic accuracy studies. Outcomes in these studies are test validity, symptoms, and functional outcomes. Most studies assessed temperature gradients or the association between temperature differences and the clinical condition. Due to the small number of studies for each indication, diagnostic accuracy could not adequately be evaluated. The clinical utility of thermography has only been considered in 1 study of diabetic foot ulcers. For these other miscellaneous conditions, the clinical utility of thermography has not been investigated. The evidence is insufficient to determine the effects of the technology on health outcomes.

**Background**

Infrared radiation from the skin or organ tissue reveals temperature variations by producing brightly colored patterns on a liquid crystal display. Thermography involves the use of an infrared scanning device and can include various types of telethermographic infrared detector images and heat-sensitive cholesteric liquid crystal systems.

Interpretation of the color patterns is thought to assist in the diagnosis of many disorders such as complex regional pain syndrome (previously known as reflex sympathetic dystrophy), breast cancer, Raynaud phenomenon, digital artery vasospasm in hand-arm vibration syndrome, peripheral nerve damage following trauma, impaired spermatogenesis in infertile men, degree of burns, deep vein
thrombosis, gastric cancer, tear-film layer stability in dry-eye syndrome, Frey syndrome, headaches, lower back pain, and vertebral subluxation.

Thermography may also assist in treatment planning and procedure guidance by accomplishing the following tasks: identifying restricted areas of perfusion in coronary artery bypass grafting, identifying unstable atherosclerotic plaques, assessing response to methylprednisolone in rheumatoid arthritis, and locating high undescended testicles.

**Regulatory Status**
A number of thermographic devices have been cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process. FDA product codes: LHQ, FXN. Devices with product code LHQ may only be marketed for adjunct use. Devices with product code FXN do not provide a diagnosis or therapy. Examples of these devices are shown in Table 1.

**Table 1. Thermography Devices Cleared by the U.S. Food and Drug Administration**

<table>
<thead>
<tr>
<th>Device Name</th>
<th>Manufacturer</th>
<th>Clearance Date</th>
<th>510(K) No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrared Sciences Breastscan IR System</td>
<td>Infrared Sciences</td>
<td>Feb 2004</td>
<td>K032350</td>
</tr>
<tr>
<td>Telethermographic Camera, Series A, E, S, and P</td>
<td>FLIR Systems</td>
<td>Mar 2004</td>
<td>K033967</td>
</tr>
<tr>
<td>Notouch Breastscan</td>
<td>UE Lifesciences</td>
<td>Feb 2012</td>
<td>K113259</td>
</tr>
<tr>
<td>WoundVision Scout</td>
<td>WoundVision</td>
<td>Dec 2013</td>
<td>K131596</td>
</tr>
<tr>
<td>AlfaSight 9000 Thermographic System</td>
<td>Alfa Thermodiagnostics</td>
<td>Apr 2015</td>
<td>K150457</td>
</tr>
<tr>
<td>FirstSense Breast Exam®</td>
<td>First Sense Medical</td>
<td>Jun 2016</td>
<td>K160573</td>
</tr>
<tr>
<td>Sentinel BreastScan II System</td>
<td>First Sense Medical</td>
<td>Jan 2017</td>
<td>K162767</td>
</tr>
<tr>
<td>InTouchThermal Camera</td>
<td>InTouch Technologies</td>
<td>Feb 2019</td>
<td>K181716</td>
</tr>
</tbody>
</table>

**Rationale**

**Literature Review**
This evidence review was created in March 1996 and has been updated regularly with searches of the PubMed database. The most recent literature review was performed through August 4, 2020.

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 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.

**Breast Cancer**

**Clinical Context and Test Purpose**
The purpose of using thermography in patients who are suspected of having breast cancer is to inform a decision whether to proceed to appropriate treatment or not.

The question addressed in this portion of the evidence review is: Does thermography when used to screen or diagnose breast cancer improve the net health outcome compared with standard mammographic techniques? Specifically, does the use of thermography improve diagnostic accuracy compared with standard screening mammography methods and is this increase in accuracy likely to improve health outcomes by leading to earlier diagnosis and treatment?

The following PICO was used to select literature to inform this review.

**Patients**
The relevant populations of interest are asymptomatic individuals being screened for breast cancer or individuals undergoing testing to diagnose breast cancer.

**Interventions**
The intervention of interest is thermography. The test would be performed in an outpatient setting.

**Comparators**
The following test is currently being used to make decisions about breast cancer diagnosis: mammography.

**Outcomes**
The outcome of interest for diagnostic accuracy is test validity (ie, sensitivity, specificity). The primary outcomes of interest for clinical utility are overall survival and breast cancer-specific survival rates.

The potential beneficial outcomes of primary interest in the case of a true-negative would be the avoidance of unnecessary surgery and its associated consequences (eg, morbidity, mortality, resource utilization, patient anxiety). The potential harms from a false-positive could be inappropriate assessment and improper management of patients with breast malignancies, which could result in the following: inappropriate surgical decisions, high frequency of unnecessary further testing, and unnecessary patient anxiety. The potential harms from a false-negative could be a determination that the patient does not have malignancy, which would lead to a delay in surgery and tumor diagnosis.
The timing for routine screening can be guided by national guidelines on breast cancer screening. The timing for diagnosis would be after an initial screening test or clinical examination.

**Technically Reliable**
Assessment of technical reliability focuses on specific tests and operators and requires a review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

**Clinically Valid**
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

**Systematic Reviews**
Several systematic reviews of the published literature on the diagnostic accuracy of thermography were identified. A systematic review by Vreugdenburg et al (2013) identified 8 studies on thermography for diagnosis of breast cancer that included a valid reference standard (eg, biopsy with histopathologic confirmation). A previous systematic review by Fitzgerald and Berentson-Shaw (2012) identified 6 studies, 1 using thermography for breast cancer screening and the others using thermography to diagnose breast cancer among symptomatic women or those with a positive mammogram. A summary of the characteristics of clinical validity for these systematic reviews is provided in Table 2. A summary of the clinical validity results is provided in Table 3. Study findings were not pooled due to heterogeneity in data reporting and assessment methodology utilized.

**Table 2. Systematic Reviews: Characteristics of Clinical Validity of Thermography in Breast Cancer Screening or Diagnosis**

<table>
<thead>
<tr>
<th>Study Population</th>
<th>Designa</th>
<th>Reference Standard</th>
<th>Threshold for Positive Index Test</th>
<th>Timing of Reference and Index Tests</th>
<th>Blinding of Assessors</th>
<th>Commentb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vreugdenburg et al (2013)1</td>
<td>For screening studies: • asymptomatic women with unknown disease status For diagnostic studies: • women with suspicious symptoms, suspicious findings on clinical examination,</td>
<td>Diagnostic, cross-sectional studies: • Retrospective case-control; sample selection consecutive Prospective cohort; sample selection NR NR cohort; sample selection NR</td>
<td>Biopsy with histopathologic confirmation</td>
<td>Various</td>
<td>Various</td>
<td>Studies blind to reference: • Blind: 4/8 • Not blind: 2/8 • Unclear: 2/8 Studies blind to comparator: • Blind: 2/8 • Not</td>
</tr>
</tbody>
</table>
or women with an abnormal mammogram

Fitzgerald et al (2012)²

For screening studies:
- asymptomatic women aged 40-65
For diagnostic studies:
- symptomatic women

Screening studies:
- Prospective cohort; NR sample selection
Diagnostic studies:
- NR case-control; sample selection NR
- NR cohort; sample selection NR

Screening studies:
- Mammo graphy
Diagnostic studies:
- biopsy with histopathologic confirmation

Various

In screening studies, only patients with positive index test received reference test. In diagnostic studies, timing of index and reference tests poorly reported.

In all studies, blinding was poorly reported.

Studies utilized various measurement scales and cut-off scores. Thermograms were scored by software, manually, or through a combination of methods. Screening study utilized more than one thermography device. Poor reporting of index and reference test timing.

NR: not reported.

a Note 2 aspects of design: prospective, retrospective or nonconcurrent prospective; sample selection random or consecutive.

b Note other characteristics that could bias or limit relevance such as use of historical data, evolution of technology, or practice setting.

Table 3. Systematic Reviews: Clinical Validity of Thermography in Breast Cancer Screening or Diagnosis

<table>
<thead>
<tr>
<th>Study; Subgroup</th>
<th>Initial N (Range)</th>
<th>Final N (Range)</th>
<th>Excluded N</th>
<th>Prevalence of Condition</th>
<th>Clinical Validity (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Vreugdenberg et al (2013)¹-Diagnostic studies</td>
<td>NR</td>
<td>1,709 (29-769)</td>
<td>565 (13-524)*</td>
<td>NR</td>
<td>25-97% (NR)</td>
</tr>
<tr>
<td>Fitzgerald et al (2012)²-Diagnostic studies</td>
<td>1,224 (63-769)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>25-97% (NR)</td>
</tr>
<tr>
<td>Fitzgerald et al (2012)²-Screening studies, at initial screening</td>
<td>10,229</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>61% (NR)</td>
</tr>
<tr>
<td>Fitzgerald et al (2012)²-Screening</td>
<td>10,229</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>28% (NR)</td>
</tr>
</tbody>
</table>
studies, at 5-yr follow-up

NPV: negative predictive value; NR: not reported; PPV: positive predictive value.

*Only 3/8 studies reported the number of excluded patients in indicated subgroup.

## Diagnostic Studies

Several studies have been published since the systematic reviews. Morales-Cervantes et al (2018) compared the accuracy of automated or manual thermography screening in 206 women scheduled for mammography in Mexico.\(^3\) A retrospective study conducted in the U.S. by Neal et al (2018) assessed outcomes in 38 women referred for further breast imaging following abnormal thermography testing.\(^4\) Omranipour et al (2016) compared the accuracy of thermography and mammography in 132 patients in Iran who had breast lesions and were candidates for breast biopsy.\(^5\) Rassiwala et al (2014) in India reported on 1008 women being screened for breast cancer.\(^6\) Summaries of characteristics and results of clinical validity for these diagnostic studies are provided in Tables 4 and 5.

### Table 4. Diagnostic Study Characteristics of Clinical Validity of Thermography in Breast Cancer Screening or Diagnosis

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Population</th>
<th>Design(^a)</th>
<th>Reference Standard</th>
<th>Threshold for Positive Index Test</th>
<th>Timing of Reference and Index Tests</th>
<th>Blinding of Assessors</th>
<th>Comment(^b)</th>
</tr>
</thead>
</table>
| Morales-Cervantes et al (2018)\(^3\) | For screening study:  
  - women scheduled for consultation with clinical evidence of tumor suspicious for breast cancer and breast cancer risk factors | Prospective cohort, NR sample allocation | Biopsy with histopathologic confirmation | Automated Thermography (Thermal Score)\(^c\)  
  - + (Thermal Score ≥2.5)  
  - - (Thermal Score < 2.5) | Reference testing performed for women with mammography BI-RADS score indicating suspicion for cancer. Mammography performed after thermography. | Blinding of mammography assessor with respect to thermography not described. Double-blinding indicated for manual assessment of thermograms by oncologist. Blinding of biopsy assessor not described. | Blinding and allocation poorly described. No data reported for mammography despite inclusion as comparator. Reported results may be biased and inaccurate due to selective use of reference tests. |
| Neal et al (2018)\(^4\)    | For diagnostic study:  
  - women referred for conventional breast imaging (mammogram and/or ultrasound) for evaluation of abnormal thermograph findings | Retrospective cohort, NR sample allocation | Biopsy with histopathologic confirmation or at least one year of clinical and/or imaging follow-up | Abnormal Thermography:  
  - Any report of abnormal findings  
  - Mammography (BI-RADS Rating):  
    - + (B4-5)  
    - - (B1-3)  
  Ultrasound (Mammography declined by Patient) of Mammography:  
    - NR | Thermography testing performed prior to mammography and/or ultrasound. Reference testing performed after index tests. Histopathological reference testing offered for women with BI-RADS score 4-5. | Blinding of assessors not described. | Blinding and allocation not described. Limited data reporting. Reference testing not uniform for all patients. Small study size with retrospective design. Long-term health outcomes not described. |
| Omranipour et al (2016)\(^5\) | For diagnostic study:  
  - women with breast lesions and candidates for breast biopsy | Prospective cohort, NR sample | Core needle or surgical biopsy with histopathology | Mammography (BI-RADS Rating): | Reference testing performed after imaging index | Mammography assessors blinded to | Blinding and allocation poorly |
breast lesions based on clinical, mammographic, or ultrasonographic finding in need of breast biopsy

• + (B4-5)
• - (B1-3)
Thermography (Rating):
• + (TH3-5)
• - (TH1-2)

tests.
thermography test results. Blinding of thermography and histopathology assessors not described.
described. Concordance of risk classification cannot be assessed due to limited data reporting.

Thermography (Rating):
• + (TH3-5)
• - (TH1-2)

Reference test provided only to women with abnormal or elevated thermography index test results.

NR

Blinding and allocation not described. Reported results may be biased and inaccurate due to selective use of reference tests.

BI-RADS: breast imaging reporting and data system; NR: not reported; ΔT: temperature gradient.
a Note 2 aspects of design: prospective, retrospective or nonconcurrent prospective; sample selection random or consecutive.
b Note other characteristics that could bias or limit relevance such as use of historical data, evolution of technology, or practice setting.
c Thermal score is defined as the sum of the surface temperature difference at the site of the lesion compared to that of the contralateral breast and the vascularity score, based on the following scale: 1) absence of vascular patterns; 2) symmetrical or moderate vascular patterns; 3) significant vascular asymmetry; 4) vascular asymmetry extended in at least one-third of breast area.

Table 5. Clinical Validity of Thermography in Breast Cancer Screening or Diagnosis

<table>
<thead>
<tr>
<th>Study; Subgroup</th>
<th>Initial N</th>
<th>Final N</th>
<th>Excluded Samples</th>
<th>Prevalence of Condition</th>
<th>Clinical Validity (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity</td>
</tr>
<tr>
<td><strong>Morales-Cervantes et al (2018)</strong></td>
<td>NR</td>
<td>206</td>
<td>NR</td>
<td>198 benign; 8 malignant</td>
<td>100% (NR)</td>
</tr>
<tr>
<td>Automated Thermography*</td>
<td>NR</td>
<td>206</td>
<td>NR</td>
<td></td>
<td>87.50% (NR)</td>
</tr>
<tr>
<td>Manual Thermography*</td>
<td>NR</td>
<td>206</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammography</td>
<td>NR</td>
<td>206</td>
<td>NR</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td><strong>Neal et al (2018)</strong></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abnormal Thermography</td>
<td>45</td>
<td>38</td>
<td>7</td>
<td>36 benign; 2 malignant</td>
<td>NA</td>
</tr>
<tr>
<td>Mammography following Abnormal Thermography</td>
<td>45</td>
<td>38</td>
<td>7</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Initial N</td>
<td>Final N</td>
<td>Excluded Samples</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------</td>
<td>---------</td>
<td>------------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Omranipour et al (2016)</strong>&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermography</td>
<td>NR</td>
<td>132</td>
<td>NR</td>
<td>81.6% (NR)</td>
<td>57.8% (NR)</td>
</tr>
<tr>
<td>Mammography</td>
<td>NR</td>
<td>132</td>
<td>NR</td>
<td>80.5% (NR)</td>
<td>73.3% (NR)</td>
</tr>
<tr>
<td><strong>Rassiwala et al (2014)</strong>&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermography**</td>
<td>NR</td>
<td>1,008</td>
<td>NR</td>
<td>97.6% (NR)</td>
<td>99.17% (NR)</td>
</tr>
</tbody>
</table>

NA: not applicable; NPV: negative predictive value; NR: not reported; PPV: positive predictive value.

* Clinical validity results for this subgroup must be interpreted with caution as subjects with normal mammograms did not undergo histopathologic reference testing for diagnostic confirmation.

** Clinical validity results for this subgroup must be interpreted with caution as subjects with normal thermograms did not undergo radiologic and histopathologic reference testing for diagnostic confirmation, only clinical assessment.

The diagnostic accuracy of automated thermography in the study by Morales-Cervantes et al (2018) was 69.9%<sup>3</sup>. The authors did not report on the diagnostic accuracy of manual thermography. While automated thermographic screening improved the sensitivity and specificity of the test compared to a manual, qualitative approach, reported values must be interpreted with caution as only patients with positive mammograms were subjected to diagnostic reference testing. Neal et al (2018) indicated that 95% of patients referred for follow-up imaging evaluation following abnormal thermography testing did not have breast cancer, concluding that conventional breast imaging appears sufficient to manage patients.<sup>4</sup> According to Omranipour et al (2016),<sup>5</sup> the diagnostic accuracy of thermography (67.7%) was lower than for mammography (76.9%; p-values not reported). The reported false-negative rate was not accurately calculated in Rassiwala et al (2014) because women who had normal thermograms only had a clinical examination and did not undergo radiologic and histopathologic reference tests for confirmation, highlighting a major limitation of this study.<sup>6</sup> For patients with positive or abnormal thermograms, 8 results were considered false-positive. One false-negative was reported but it is unclear which subgroup this patient belonged to or how this was determined, given that patients with normal thermograms were only assessed with a clinical examination. Limitations tables (see Tables 6 and 7) display further 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 the evidence supporting the position statement.
Table 6. Study Relevance Limitations: Breast Cancer Screening or Diagnosis

<table>
<thead>
<tr>
<th>Study</th>
<th>Population&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Intervention&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Comparator&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Outcomes&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Duration of Follow-Up&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morales-Cervantes et al (2018)&lt;sup&gt;3&lt;/sup&gt;</td>
<td>1, 4. Intended use population unclear; study population not representative of intended use (screening study enriched with patients with clinical symptoms).</td>
<td>1-2. Classification thresholds for manual thermographic assessment not described; BI-RADS version used unclear with no description of classification thresholds.</td>
<td>1-2. BI-RADS classification thresholds for mammography not defined; normal mammograms not compared to credible reference standard.</td>
<td>1, 3, 5. Study does not directly assess a key health outcome; key clinical validity outcomes not reported; adverse events of the test not described.</td>
<td>1. Follow-up duration not sufficient for patients not evaluated by biopsy.</td>
</tr>
<tr>
<td>Neal et al (2018)&lt;sup&gt;4&lt;/sup&gt;</td>
<td>1. Classification thresholds for patients receiving ultrasounds after declining mammography not described; classification thresholds for thermography not evaluated.</td>
<td>1. Not compared to consistent reference standard.</td>
<td>1. Study does not report on key long-term health outcomes; key clinical validity outcomes not reported.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omranipour et al (2016)&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td>1, 5. Study does not directly assess a key health outcome; adverse events of the test not described.</td>
<td></td>
</tr>
<tr>
<td>Rassiwala et al (2014)&lt;sup&gt;6&lt;/sup&gt;</td>
<td>4. Study population not representative of intended use (age for screening).</td>
<td>1-2. Classification thresholds not defined; normal index tests not compared to credible reference standard.</td>
<td>1-2. Classification thresholds not defined; normal index tests not compared to credible reference standard.</td>
<td>1, 4-5. Study does not directly assess a key health outcome; reclassification of diagnostic or risk categories not reported; adverse events of the test not described.</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.

BI-RADS: breast imaging reporting and data system.

<sup>a</sup> 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.

<sup>b</sup> Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.

<sup>c</sup> Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).

Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

Table 7. Study Design and Conduct Limitations: Breast Cancer Screening or Diagnosis

<table>
<thead>
<tr>
<th>Study</th>
<th>Selectiona</th>
<th>Blindingb</th>
<th>Delivery of Testc</th>
<th>Selective Reportingd</th>
<th>Data Completenesse</th>
<th>Statisticalf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morales-Cervantes et al (2018)³</td>
<td>1. Selection not described.</td>
<td>1. Blinding to index and reference tests not fully described.</td>
<td>3, 4. Procedure for manual interpretation of thermograms and mammograms not described; expertise of all evaluators not described.</td>
<td>1-2. Not registered; evidence of selective reporting (mammography data not reported).</td>
<td>1. No description of indeterminate or missing samples.</td>
<td>1-2. Confidence intervals and/or p values not reported; comparison to mammography not reported.</td>
</tr>
<tr>
<td>Neal et al (2018)⁴</td>
<td>1. Selection not described.</td>
<td>1. Blinding not described.</td>
<td>2-3. Timing of index and comparator tests not same; procedures for interpreting all tests not described</td>
<td>1. Not registered.</td>
<td>3. High loss to follow-up or missing data.</td>
<td>1-2. Confidence intervals and/or p values not reported; comparison to other tests not reported.</td>
</tr>
<tr>
<td>Rassiwala et al (2014)⁶</td>
<td>1. Selection not described.</td>
<td>1. Blinding not described.</td>
<td>1,3-4. Timing of delivery of index and reference tests not fully described; procedure for interpreting reference tests not described; expertise of evaluators not described.</td>
<td>1. Not registered.</td>
<td>1. Inadequate description of indeterminate or missing samples.</td>
<td>1. Confidence intervals and/or p values not reported.</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 Selection key: 1. Selection not described; 2. Selection not random or consecutive (ie, convenience).

b Blinding key: 1. Not blinded to results of reference or other comparator tests.

c Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.


e Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.

f Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

**Clinically Useful**
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

**Direct Evidence**
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

No studies have demonstrated how the results of thermography could be used to enhance the management of breast cancer patients in a manner that would improve their health outcomes.

**Chain of Evidence**
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

It is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence that the diagnostic accuracy of thermography is at least as high as mammographic techniques for breast cancer screening and diagnosis.

**Section Summary: Breast Cancer**
Systematic reviews of studies evaluating the accuracy of thermography for diagnosing breast cancer found wide ranges of sensitivities and specificities and, where data are available, relatively low diagnostic accuracy compared with mammography. To date, no study has demonstrated that thermography is sufficiently accurate to replace or supplement mammography for breast cancer diagnosis. Moreover, there are no studies on the impact of thermography on patient management or health outcomes for patients with breast cancer.
Musculoskeletal Injuries

Clinical Context and Test Purpose
The purpose of using thermography in patients who have a musculoskeletal injury is to inform a decision whether to proceed to appropriate treatment or not.

The question addressed in this portion of the evidence review is: Does thermography when used to diagnose musculoskeletal injuries, improve the net health outcome compared with standard approaches. Specifically, does the use of thermography improve diagnostic accuracy compared with standard approaches (eg, clinical examination, imaging with radiography or magnetic resonance imaging), and is this degree of increased accuracy likely to improve health outcomes by leading to earlier diagnosis and treatment?

The following PICO was used to select literature to inform this review.

Patients
The relevant population of interest is individuals with musculoskeletal pain.

Interventions
The intervention of interest is thermography. The test would be performed in an outpatient setting.

Comparators
The following tests and practices are currently being used to make decisions about musculoskeletal injuries: standard care without imaging and other forms of imaging (eg, with radiography, magnetic resonance imaging).

Outcomes
The outcomes of interest for diagnostic accuracy include test accuracy and test validity (ie, sensitivity, specificity). The primary outcomes of interest for clinical utility are a reduction in pain symptoms and improvement in functional ability. The timing would be following a musculoskeletal injury.

Technically Reliable
Assessment of technical reliability focuses on specific tests and operators and requires a review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Systematic Reviews
A systematic review by Sanchis-Sanchez et al (2014) evaluated the literature on thermography for diagnosing musculoskeletal injuries. Six studies met the
eligibility criteria (n=416); 3 included patients with suspected stress fractures (n=119) and the remainder addressed other musculoskeletal injuries. Characteristics and results of clinical validity for stress fracture diagnostic studies were reported and summaries are provided in Tables 8 and 9. A systematic review by Vardasca et al (2019) evaluated the literature on musculoskeletal applications of thermography specific to the arm and forearm. However, the review mainly focused on correlations between skin surface temperatures and physical condition or health recovery monitoring. As diagnostic accuracy data was not extracted or pooled from included studies, this review was not assessed for evidence of clinical validity.

Table 8. Systematic Review: Characteristics of Clinical Validity of Thermography in Musculoskeletal Injury

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Population</th>
<th>Designa</th>
<th>Reference Standard</th>
<th>Threshold for Positive Index Test</th>
<th>Timing of Reference and Index Tests</th>
<th>Blinding of Assessors</th>
<th>Commentb</th>
</tr>
</thead>
</table>

Thermography and Temperature Gradient Studies 6.01.12
For diagnostic studies:
• studies reporting on the diagnostic accuracy of infrared thermal imaging in the diagnosis of musculoskeletal injuries (e.g., bone fractures, dislocations, sprains, muscle contractures, tendinopathy, contusions, or compartment syndrome) that utilized a recognized reference standard (e.g., radiographs, CT, MRI, or ultrasound scanning)

- Prospective cohort; sample selection consecutive (4/6)
- Prospective cohort; sample selection NR (1/6)
- Prospective cohort; sample selection by convenience (1/6)

High-quality radiographic imaging (various)

Reported (1/6 studies)
Unclear (4/6 studies, including all studies on stress fractures)
NR (1/6 studies)

Reported (2/6 studies)
Unclear (4/6 studies, including all studies on stress fractures)

High heterogeneity in thermography index test methodologies and diagnostic accuracy. QUADAS assessment by authors indicates moderate-to-high risk of bias in studies on stress fractures.

CT: computed tomography; MRI: magnetic resonance imaging; NR: not reported; QUADAS: Quality Assessment of Diagnostic Accuracy Studies.

Note 2 aspects of design: prospective, retrospective or nonconcurrent prospective; sample selection random or consecutive.

Note other characteristics that could bias or limit relevance such as use of historical data, evolution of technology, or practice setting.

Table 9. Systematic Review: Clinical Validity of Thermography in Musculoskeletal Injury

<table>
<thead>
<tr>
<th>Study; Subgroup</th>
<th>Initial N (Range)</th>
<th>Final N (Range)</th>
<th>Excluded N</th>
<th>Prevalence of Condition</th>
<th>Clinical Validity (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Sanchis-Sanchez</td>
<td>NR</td>
<td>119 (17-84)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>
### Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

### Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

### Longitudinal Studies

Côrte et al (2019) published pilot data from a longitudinal prospective study on the screening and prevention of muscle injuries in 28 professional Brazilian soccer players. Players were monitored for musculoskeletal imaging during the 2015-2016 seasons with ultrasound. In the second season, a thermographic monitoring regimen was added twice-weekly 48 hours after matches, and an injury prevention protocol was followed based on the results of thermographic imaging. The number of musculoskeletal injuries was compared for both seasons based on these management protocols. The total number of muscle injuries reported decreased from 11 in 2015 to 4 in 2016 (p=0.04). Seven players were on the team roster across both seasons. There was no statistically significant reduction in muscle injury in this subgroup (p=0.06). Limitations of this study are addressed in Tables 10 and 11.

### Table 10. Study Relevance Limitations: Musculoskeletal Injury

<table>
<thead>
<tr>
<th>Study</th>
<th>Population&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Intervention&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Comparator&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Outcomes&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Duration of Follow-Up&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Côrte et al (2019)&lt;sup&gt;8&lt;/sup&gt;</td>
<td>2. Clinical context is unclear (definition and reporting of)</td>
<td>2. Version used unclear (therapy utilized in prevention protocol was based on)</td>
<td>1-2. Classification thresholds for ultrasound not defined; comparison to</td>
<td>3-5. Key clinical validity outcomes not reported; reclassification of diagnostic or risk</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.

| Selection key: 1. Selection not described; 2. Selection not random or consecutive (ie, convenience). |
| Blinding key: 1. Not blinded to results of reference or other comparator tests. |
| Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described. |
| Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data. |
| Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported. |

Table 11. Study Design and Conduct Limitations: Musculoskeletal Injury

<table>
<thead>
<tr>
<th>Study</th>
<th>Selectiona</th>
<th>Blindingb</th>
<th>Delivery of Testc</th>
<th>Selective Reportingd</th>
<th>Data Completenesse</th>
<th>Statisticalf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Côrte et al (2019)g</td>
<td>1. Selection not random or consecutive.</td>
<td>1. Blinding to index and reference tests not described.</td>
<td>1-4. Timing of delivery of index or reference tests not described; timing of index and comparator tests not described; procedure for interpreting comparator and/or reference tests not described; expertise of evaluators not described.</td>
<td>1. Not registered.</td>
<td>1. No description of indeterminate or missing samples.</td>
<td>1-2. Confidence intervals and/or p values not reported; diagnostic comparison to other tests not reported.</td>
</tr>
</tbody>
</table>
No high-quality or randomized studies have been published that evaluate health outcomes in patients with musculoskeletal injuries who were managed with and without thermography.

**Chain of Evidence**
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

It is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence that the diagnostic accuracy of thermography is at least as high as standard techniques for diagnosing musculoskeletal injuries.

**Section Summary: Musculoskeletal Injuries**
A systematic review of studies on thermography for diagnosing musculoskeletal injuries found moderate levels of accuracy compared with other diagnostic imaging tests. There was a lack of a consistent reference standard. This evidence does not permit conclusions as to whether thermography is sufficiently accurate to replace or supplement standard testing. Moreover, there are insufficient studies on the impact of thermography on patient management or health outcomes for patients with musculoskeletal injuries.

**Temporomandibular Joint Disorder**

**Clinical Context and Test Purpose**
The purpose of using thermography in patients who have TMJ disorder is to inform a decision whether to proceed to appropriate treatment or not.

The question addressed in this portion of the evidence review is: Does thermography when used to diagnose TMJ disorder, improve the net health outcome compared with standard approaches. Specifically, does the use of thermography improve diagnostic accuracy compared with standard approaches (eg, clinical examination, imaging with radiography or magnetic resonance imaging), and is this degree of increased accuracy likely to improve health outcomes by leading to earlier diagnosis and treatment?

The following PICO was used to select literature to inform this review.

**Patients**
The relevant population of interest is individuals with TMJ pain.

**Interventions**
The intervention of interest is thermography. The test would be performed in an outpatient setting.

**Comparators**
The following tests and practices are currently being used to make decisions about TMJ disorder: standard clinical examination without imaging, diagnostic scales (eg,
Research Diagnostic Criteria for Temporomandibular Disorders [RDC/TMD], Fonseca Anamnestic Index, Anamnestic Index), and other forms of imaging (eg, with radiography, arthrotomography, magnetic resonance imaging).

**Outcomes**
The outcomes of interest for diagnostic accuracy include test accuracy and test validity (ie, sensitivity, specificity). The primary outcomes of interest for clinical utility are a reduction in pain symptoms and improvement in functional ability.

**Technically Reliable**
Assessment of technical reliability focuses on specific tests and operators and requires a review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

**Clinically Valid**
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

**Systematic Reviews**
A systematic review by de Melo et al (2019) evaluated the diagnostic accuracy of infrared thermography in TMJ disorder.\(^9\) Nine studies were identified utilizing a variety of comparators. The authors note that while no specific diagnostic tool is currently considered the gold standard for the diagnosis of TMJ disorder, the RDC/TMD diagnostic is commonly used with a reported sensitivity and specificity of 87% and 92%, respectively. Four out of 9 studies utilized RDC/TMD, whereas the remaining studies utilized clinical examination or other methods. Characteristics and results of clinical validity for temporomandibular joint disorder diagnostic accuracy in this systematic review are summarized in Tables 12 and 13.

**Table 12. Systematic Review: Characteristics of Clinical Validity of Thermography in Temporomandibular Joint Disorder**

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Population</th>
<th>Design(^a)</th>
<th>Reference Standard</th>
<th>Threshold for Positive Index Test</th>
<th>Timing of Reference and Index Tests</th>
<th>Blinding of Assessors</th>
<th>Comment(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Melo et al (2019)(^9)</td>
<td>For diagnostic studies: • studies reporting on the diagnostic accuracy of infrared thermography vs other</td>
<td>• NR; sample selection consecutive (1/9 studies) or by convenience (8/9)</td>
<td>RDC/TMD diagnostic, clinical examination, or other imaging methods</td>
<td>NR</td>
<td>NR</td>
<td>High-risk of bias based on flow and timing: 4/9 studies; Unclear risk of bias based on</td>
<td>Thermography index test methodologies unclear. Heterogeneity in use of comparator and/or reference standard. Assessment</td>
</tr>
</tbody>
</table>
diagnostic tests and imaging methods in patients with temporomandibular disorder

• flow and timing: 5/9 studies.

by authors indicates high-risk of bias in all studies.

NR: not reported; RDC/TMD: Research Diagnostic Criteria for Temporomandibular Disorders.

a Note 2 aspects of design: prospective, retrospective or nonconcurrent prospective; sample selection random or consecutive.

b Note other characteristics that could bias or limit relevance such as use of historical data, evolution of technology, or practice setting.

Table 13. Systematic Review: Clinical Validity of Thermography in Temporomandibular Joint Disorder

<table>
<thead>
<tr>
<th>Study</th>
<th>Initial N (Range)</th>
<th>Final N (Range)</th>
<th>Excluded N</th>
<th>Prevalence of Condition</th>
<th>Clinical Validity (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Melo et al (2019)⁵</td>
<td>NR</td>
<td>548 (23-104)</td>
<td>NR</td>
<td>NR</td>
<td>NR; Range: 38.5-90%</td>
</tr>
</tbody>
</table>

NPV: negative predictive value; NR: not reported; PPV: positive predictive value.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

No studies have been published that evaluate health outcomes in patients with TMJ disorder who were managed with and without thermography.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

It is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence that the diagnostic accuracy of thermography is at least as high as standard techniques for diagnosing TMJ disorder.
Section Summary: Temporomandibular Joint Disorder

A systematic review of studies on thermography for diagnosing TMJ disorder found a wide variation in accuracy compared with other diagnostics. There was a lack of a consistent reference standard. This evidence does not permit conclusions as to whether thermography is sufficiently accurate to replace or supplement standard testing. Moreover, there are no studies on the impact of thermography on patient management or health outcomes for patients with TMJ disorder.

Miscellaneous Conditions

A number of studies have assessed a range of potential thermography applications. To date, no randomized study has examined the impact of thermography on patient management decisions or health outcomes. Examples of other studies on thermography, mainly conducted outside of the U. S., include those evaluating the association between thermographic findings and post-herpetic neuralgia in patients with herpes zoster,10,11 surgical site healing in patients who underwent knee replacements,12 predicting pressure ulcers13 and pressure ulcer healing,14 posttreatment pain in patients with coccygodynia,15 evaluation of allergic conjunctivitis,16 evaluation of burn depth17 and association between thermographic findings and burn treatment,18 detecting cervical lymph node metastasis from oral cavity cancer,19 monitoring lesions or inflammation in patients with scleroderma,20,21 detection of vascular obstruction22 or perforator vessels during surgery,23,24 diagnosis of lower extremity cellulitis,25 prediction of infrainguinal bypass surgery,26 detection of melanoma,27 detection of contact dermatitis during allergy patch testing,28 and measuring disease activity in patients with rheumatoid arthritis.29,30

Several studies evaluating the clinical validity of thermography to assess potential complications of the diabetic foot have been conducted. Thermographic images of nondiabetic feet, nonulcerated diabetic feet and ulcerated diabetic feet have been compared.31,32,33,34,35 Another study used thermography to diagnose infections in patients admitted with diabetic foot complications.36 The only study to date to investigate the clinical utility of thermography compared with no thermography assessed diabetic foot ulcer incidence in 110 participants with a history of diabetic neuropathy and foot ulcers.37 After 12 months followup, the study found no significant difference between use of monthly thermography versus no thermography and foot ulcer incidence (62% versus 56%; adjusted OR 0.55, 95% CI 0.21 to 1.40) or time to ulcer recurrence (adjusted HR 0.67, 95% CI 0.34 to1.3).

Section Summary: Miscellaneous Conditions

For most of these potential indications, there are 1 or 2 preliminary studies on each of the indications. Several studies evaluated the clinical validity of thermography in assessing diabetic foot and related complications. For all indications, the studies described temperature gradients or the association between temperature differences and the clinical condition. Due to the small number of studies for each indication, the diagnostic accuracy could not adequately be evaluated. The clinical utility of thermography for these miscellaneous conditions was not investigated in any study.
Summary of Evidence
For individuals who have an indication for breast cancer screening or diagnosis who receive thermography, the evidence includes diagnostic accuracy studies and systematic reviews. Relevant outcomes are overall survival, disease-specific survival, and test validity. Using histopathologic findings as to the reference standard, a series of systematic reviews of studies have evaluated the accuracy of thermography to screen and/or diagnose breast cancer and reported wide ranges of sensitivities and specificities. To date, no study has demonstrated whether thermography is sufficiently accurate to replace or supplement mammography for breast cancer diagnosis. Moreover, there are no studies on the impact of thermography on patient management or health outcomes for patients with breast cancer. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have musculoskeletal injuries who receive thermography, the evidence includes diagnostic accuracy studies, a longitudinal prospective study, and a systematic review. Relevant outcomes are test validity, symptoms, and functional outcomes. A systematic review of studies on thermography for diagnosing musculoskeletal injuries found moderate levels of accuracy compared with other diagnostic imaging tests. There is a lack of a consistent reference standard. This evidence does not permit conclusions as to whether thermography is sufficiently accurate to replace or supplement standard testing. Moreover, there are no high-quality or randomized studies on the impact of thermography on patient management or health outcomes for patients with musculoskeletal injuries. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have temporomandibular joint disorder who receive thermography, the evidence includes a systematic review. Relevant outcomes are test validity, symptoms, and functional outcomes. A systematic review of studies on thermography for diagnosing temporomandibular joint disorder found a wide variation in accuracy compared to other diagnostics. There is a lack of a consistent reference standard. The evidence does not permit conclusions as to whether thermography is sufficiently accurate to replace or supplement standard testing. Moreover, there are no studies on the impact of thermography on patient management or health outcomes for patients with the temporomandibular joint disorder. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have miscellaneous conditions (eg, herpes zoster, pressure ulcers, diabetic foot) who receive thermography, the evidence primarily includes diagnostic accuracy studies. Outcomes in these studies are test validity, symptoms, and functional outcomes. Most studies assessed temperature gradients or the association between temperature differences and the clinical condition. Due to the small number of studies for each indication, diagnostic accuracy could not adequately be evaluated. The clinical utility of thermography has only been considered in 1 study of of diabetic foot ulcers. For a these other miscellaneous
conditions, the clinical utility of thermography has not been investigated. The evidence is insufficient to determine the effects of the technology on health outcomes.

SUPPLEMENTAL INFORMATION

Practice Guidelines and Position Statements

European Society of Breast Imaging
A position paper by the European Society of Breast Imaging (2017) and 30 other national breast radiology bodies on screening for breast cancer stated that "screening with thermography or other optical tools as alternatives to mammography is discouraged."38.

American College of Physicians
The American College of Physicians (2019) issued a guidance statement for breast cancer screening in average-risk women that reviews existing screening guidelines. While the use of thermography was not mentioned in this statement, the authors conclude that evidence is insufficient to understand the benefits and harms of primary or adjunctive screening strategies in women who are found to have dense breasts on screening mammography.39.

American College of Radiology
The American College of Radiology guidelines for breast cancer screening (revised 2017) do not mention the use of thermography for breast cancer screening.40.

National Comprehensive Cancer Network
National Comprehensive Cancer Network guidelines on breast cancer screening and diagnosis (v.1.2019) states that: "Current evidence does not support the routine use of thermography or ductal lavage as screening procedures."41.

U.S. Preventive Services Task Force Recommendations
The U.S. Preventive Services Task Force (2016) recommendations on breast cancer screening do not mention thermography. Additionally, there is insufficient evidence for the use of adjunctive screening methods for breast cancer (ultrasonography, magnetic resonance imaging, digital breast tomosynthesis, or other methods) in women identified to have dense breasts on a negative screening mammogram.42.

Medicare National Coverage
Medicare does not cover thermography. Current Medicare coverage policy states: "Thermography for any indication (including breast lesions which were excluded from Medicare coverage ...) is excluded from Medicare coverage because the available evidence does not support this test as a useful aid in the diagnosis or treatment of illness or injury. Therefore, it is not considered effective..."43.
**Ongoing and Unpublished Clinical Trials**

Some currently ongoing and unpublished trials that might influence this review are listed in Table 14.

**Table 14. Summary of Key Trials**

<table>
<thead>
<tr>
<th>NCT No.</th>
<th>Trial Name</th>
<th>Planned Enrollment</th>
<th>Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ongoing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT03089463</td>
<td>Foot Assessment in People with Diabetes: A Quantitative Diagnostic Approach</td>
<td>153</td>
<td>Apr 2020 (recruiting)</td>
</tr>
<tr>
<td>NCT03735550</td>
<td>Investigation of the Effectiveness of Liquid Crystal Contact Thermography in Detecting Pathological Changes in Female Breasts Compared to Standard Diagnostic Methods of Breast Cancer</td>
<td>3000</td>
<td>Jan 2019 (ongoing)</td>
</tr>
<tr>
<td>NCT03217214</td>
<td>Investigation of Contact Based Method for Diagnosis of Cardiovascular Disease</td>
<td>100</td>
<td>Sep 2019 (recruiting)</td>
</tr>
<tr>
<td>NCT02776995</td>
<td>Tumor Monitoring Using Thermography During Radiation Therapy</td>
<td>80</td>
<td>Dec 2020 (recruiting)</td>
</tr>
<tr>
<td>NCT04013711</td>
<td>Quantitative Thermal Imaging to Evaluate Skin Toxicity from Radiation Treatment</td>
<td>200</td>
<td>Aug 2021 (recruiting)</td>
</tr>
<tr>
<td><strong>Unpublished</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT03254095</td>
<td>Thermal and Biomechanical Characterization of Diabetic Foot Patients, Predictors of Skin Temperature, Barefoot Plantar Pressure and Ulceration</td>
<td>100</td>
<td>Sep 2018 (completed)</td>
</tr>
</tbody>
</table>

NCT: national clinical trial.

REFERENCES


38. Sardanelli F, Aase HS, Alvarez M, et al. Position paper on screening for breast cancer by the European Society of Breast Imaging (EUSOBI) and 30 national breast radiology bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, The Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey. Eur Radiol. Jul 2017; 27(7): 2737-2743. PMID 27807699


Billing Coding/Physician Documentation Information

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>93740</td>
<td>Temperature gradient studies</td>
</tr>
<tr>
<td>93799</td>
<td>Unlisted cardiovascular service or procedure</td>
</tr>
<tr>
<td>K1004</td>
<td>Low frequency ultrasonic diathermy treatment device for home use, includes all components and accessories</td>
</tr>
</tbody>
</table>
ICD-10 Codes

G56.40- G56.42  Causalgia of upper limb code range
G57.70- G57.72  Causalgia of lower limb code range
G89.0- G89.4  Pain, not elsewhere classified code range
G90.50- G90.59  Complex regional pain syndrome I code range
M25.50- M25.579  Pain in joint code range
M54.00- M54.9  Dorsalgia code range
M79.60- M79.676  Pain in limb, hand, foot, fingers and toes code range
R52  Pain, unspecified

CPT codes 93760 and 93762 for thermography were deleted effective 12/31/2008. These services would now be reported using the unlisted code 93799.

Additional Policy Key Words
N/A

Policy Implementation/Update Information

11/1/02 New policy added to the Medical section.
11/1/03 No policy statement changes. Added to the radiology section.
11/1/04 No policy statement changes.
11/1/05 No policy statement changes.
5/1/06 No policy statement changes.
11/1/06 No policy statement changes.
5/1/07 No policy statement changes.
11/1/07 No policy statement changes.
5/1/08 No policy statement changes.
8/1/08 Policy updated to include discussion regarding Temperature Gradient Studies. Policy statement revised to indicate this is considered investigational. Policy title updated to include Temperature Gradient Studies.
11/1/08 No policy statement changes.
5/1/09 No policy statement changes.
11/1/09 No policy statement changes.
5/1/10 No policy statement changes.
11/1/10 No policy statement changes.
5/1/11 No policy statement changes.
11/1/12 No policy statement changes.
11/1/13 No policy statement changes.
11/1/14 No policy statement changes.
11/1/15 No policy statement changes.
11/1/16  No policy statement changes.
11/1/17  No policy statement changes.
11/1/18  No policy statement changes.
11/1/19  No policy statement changes.
11/1/20  No policy statement changes.

State and Federal mandates and health plan contract language, including specific provisions/exclusions, take precedence over Medical Policy and must be considered first in determining eligibility for coverage. The medical policies contained herein are for informational purposes. The medical policies do not constitute medical advice or medical care. Treating health care providers are independent contractors and are neither employees nor agents Blue KC and are solely responsible for diagnosis, treatment and medical advice. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, photocopying, or otherwise, without permission from Blue KC.