11-2013

Oncology Section Task Force on Breast Cancer Outcomes: Scapular Assessment

Mary Insana Fisher  
*University of Dayton, mary.fisher@udayton.edu*

Pamela K. Levangie  
*MGH Institute of Health Professions*

Follow this and additional works at: [https://ecommons.udayton.edu/dpt_fac_pub](https://ecommons.udayton.edu/dpt_fac_pub)

Part of the [Biomechanics Commons](https://ecommons.udayton.edu/dpt_fac_pub), [Musculoskeletal System Commons](https://ecommons.udayton.edu/dpt_fac_pub), [Oncology Commons](https://ecommons.udayton.edu/dpt_fac_pub), [Orthopedics Commons](https://ecommons.udayton.edu/dpt_fac_pub), [Sports Sciences Commons](https://ecommons.udayton.edu/dpt_fac_pub), and the [Therapeutics Commons](https://ecommons.udayton.edu/dpt_fac_pub)

---

**eCommons Citation**

[https://ecommons.udayton.edu/dpt_fac_pub/19](https://ecommons.udayton.edu/dpt_fac_pub/19)

---

This Article is brought to you for free and open access by the Department of Physical Therapy at eCommons. It has been accepted for inclusion in Physical Therapy Faculty Publications by an authorized administrator of eCommons. For more information, please contact frice1@udayton.edu, mslacken1@udayton.edu.
ABSTRACT

Background: Functional deficits and changes in scapular mechanics following breast cancer (BC) treatments have been documented. Scapular assessment is important when examining the shoulder in survivors of breast cancer to document the need for or effectiveness of physical therapy intervention. The Oncology Section Task Force on Breast Cancer Outcomes sought to identify scapular examination tools that can be recommended for routine use in individuals treated for BC. Methods: A systematic review of the literature on scapular measures was conducted. Relevant studies were examined for psychometric properties and clinical usefulness. Each method was given a recommendation score based on the Breast Cancer EDGE (Evidence Database to Guide Effectiveness) criteria. Results: Only Dynamic Motion Assessment was recommended for clinical use. The remaining tools lacked either good psychometric properties or clinical usefulness. Conclusions: Measurement of scapular motion remains a challenge and reliable and valid measures must precede further research into scapular problems among survivors of breast cancer.

INTRODUCTION

Breast cancer is the most commonly diagnosed cancer among women and has a reported survival rate of 90%.1 As a consequence, more women are living after diagnosis. With survivors estimated at 2.5 million women in the United States alone, a focus on quality of life (QOL) has emerged within the research and clinical communities.2 As a component of examining QOL, upper extremity (UE) function of breast cancer survivors (BCS) has been investigated for its impact on levels of activity and participation.

Gross shoulder motion and strength deficits have been reported both immediately after breast cancer treatment as well as several years following completion of treatment.3,5 These limitations in UE function may be attributed in part to impaired scapular function, given that using the arm in functional tasks requires a balance between scapular control and scapular mobility. This controlled mobility can be impaired following breast cancer treatments. Preliminary research investigating scapular motion and control among BCS has documented changes in the pectoral musculature following mastectomy, as well as abnormal scapular mechanics.6-8

Dysfunction in scapular mechanics, termed scapular dyskinesis, has been implicated in shoulder pathology. Scapular dyskinesis was first described by Kibler, and is defined as abnormal motions of the scapula during upper extremity movement.9,10 This abnormal motion has been reported in individuals with shoulder impingement and rotator cuff pathology.11-14 Given that researchers have noted abnormal scapular motion among BCS5 and that range of motion and strength deficits in the UE in this population may be attributable to scapular dyskinesis, it is imperative to assess scapular function as a component of examination of the shoulder.

The Research Section of the American Physical Therapy Association (APTA) formed a task force charged with identifying for physical therapists valid, reliable, and clinically useful outcome measures specific to particular patient populations. The EDGE (Evidence Database to Guide Effectiveness) Task Force disseminated its work to APTA’s clinical sections. The Oncology Section Task Force on Breast Cancer Outcomes (Breast Cancer EDGE) was convened in 2010 with the goal of identifying outcome measures with properties that supported routine use by physical therapists in the breast cancer population. The first target was shoulder measures. Three subgroups were formed; this review focuses on scapular assessment methods. The purpose of this paper is to report on a systematic review of the literature that examined measures of scapular position and motion with consideration of the psychometric properties and clinical utility of the measures.

METHODS

Search Strategy

The primary search was conducted by both authors using multiple electronic databases, including: Academic Search Premier, Medline, CINAHL, PubMed, Sport Discus, and Pedro. Search terms included the term scapula along with assessment,
kinematics, dyskinesis, position, and measurement. Bibliographic review of relevant articles was conducted as well as review of journals focusing on orthopedics or shoulders. Studies of scapular measurement methods had to meet the following inclusion criteria: clinically feasible tests of scapular position or function, psychometric properties reported, and published in the English language. Exclusion criteria included use of 3 dimensional (3D) motion analysis or imaging studies (radiographs, magnetic resonance imaging, ultrasound imaging) because these measures cannot be used by physical therapists in most clinics for day-to-day patient care. No limit was placed on the publication dates as long as the inclusion criteria were met. All papers were selected prior to March 2011. Each study was reviewed by both authors separately for inclusion and exclusion criteria. Any disagreements were resolved by discussion, and a final list of measures was compiled.

Data Extraction and Synthesis
Relevant data were extracted and recorded on the CancerEDGE Task Force Outcome Measure Rating Form for each study (see the CancerEDGE form in the Appendix of the Introductory paper in this series). Studies were then grouped together into common categories and a single CancerEDGE form completed for each category of scapular assessment. Upon completion of the CancerEDGE form, a recommendation was made using the Breast Cancer EDGE 4-point scale (Table 1). Determination of good psychometric properties was determined by either intraclass coefficient (ICC) or Kappa values. The ICC values greater than 0.75 were considered good to excellent, 0.5-0.74 moderate, and below 0.5 considered poor. Kappa values greater than 80% demonstrated excellent agreement, 61%-80% substantial agreement, 41%-60% adequate agreement, and less than 40% showed poor agreement.

RESULTS
The initial search resulted in 694 possible studies. Of these, abstract review and elimination of duplicates reduced the number to 59 potential studies. Further review of the studies with application of the inclusion and exclusion criteria resulted in 18 studies that were included in this systematic review. See Figure 1 for the flow of the search process and relevant articles for each category of scapular assessment.

After data extraction, 6 outcome measures were classified into two categories: Positional Assessments (Anterior/Posterior Tilt, Upward/Downward Rotation, Protraction/Retraction) and Movement Assessments (Serial Positional Assessment, Lateral Scapular Slide Test, Dynamic Motion Assessment). Table 2 summarizes the psychometric properties of each scapular assessment technique.

Positional Assessment
Anterior/posterior tilt
Two studies examined the distance from the posterior acromion to a surface parallel to the trunk, either supine or standing. One study examined the degree to which the scapula was tilted anteriorly by use of a specialized tool, the Perry tool, which measured the degree of tilt. This measure was rated 2B (unable to recommend at this time).

Protraction/retraction
Four studies examined methods to describe scapular position in terms of scapular protraction or retraction. Three studies examined the psychometric properties of the Lennie test, which involves measuring distances from scapular landmarks to midline of the spine, and one study examined the use of a tool (PALM) to measure similar distances. This measure was rated 2B (unable to recommend at this time).

Table 1. Breast Cancer EDGE Rating Scale

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Highly recommended; the outcome has good psychometric properties and good clinical utility; the measure has been used in research on individuals with or post breast cancer.</td>
</tr>
<tr>
<td>3</td>
<td>Recommended; the outcome measure has good psychometric properties and good clinical utility; No published evidence that the measure has been applied to research on individuals with or post breast cancer.</td>
</tr>
<tr>
<td>2A</td>
<td>Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; the measure has been used in research on individuals with or post breast cancer.</td>
</tr>
<tr>
<td>2B</td>
<td>Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; no published evidence that the measure has been applied to research on individuals with or post breast cancer.</td>
</tr>
<tr>
<td>1</td>
<td>Not recommended; the outcome measure has poor psychometric properties and/or poor clinical utility.</td>
</tr>
</tbody>
</table>
### Table 2. Psychometric Properties of Scapular Measures

<table>
<thead>
<tr>
<th>Measurement Technique</th>
<th>Rating</th>
<th>Relevant Psychometric Properties</th>
<th>Clinical Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positional Assessment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Anterior/Posterior Tilt | 2B | **Inter-rater reliability:** Supine:\(^{17}\) relaxed ICC = 0.88 – 0.94; retraction ICC = 0.91 – 0.92  
Standing, back to wall:\(^{18}\) relaxed ICC = 0.72; scapular retraction ICC = 0.75  
Perry Tool:\(^{19}\) unweighted ICC = 0.92 – 0.97  
weighted ICC = 0.92 – 0.95  
**Intra-rater reliability:**  
Perry Tool:\(^{19}\) unweighted ICC = 0.98 – 0.99  
weighted ICC = 0.97 – 0.99  
**Concurrent Validity** (pain and disability scales):  
r = 0.02 – 0.20 | Minimal in terms of describing scapular position. Measurement does not indicate pathology. Variability within and between participants is normal. |
| Acromion to posterior parallel surface and Perry Tool | | | |
| Protraction/Retraction | 2B | **Intra-rater reliability:**  
ICC = 0.84 – 0.96;\(^{20}\) 0.97;\(^{21}\) 0.91;\(^{22}\) 0.77 – 0.89\(^{20}\)  
**Inter rater reliability:**  
ICC = 0.76 – 0.94;\(^{23}\) 0.96;\(^{21}\) 0.62 – 0.94\(^{20}\)  
**Concurrent Validity:**  
r = 0.69 – 0.82;\(^{21}\) 0.73-0.79\(^{21}\) | Measurement does not indicate pathology. Variability within and between participants is normal. |
| Lennie Test  
Modified Lennie  
P iLM | | | |
| Upward/Downward Rotation | 2B | **Intra-rater reliability** (across varying elevation angles):  
ICC = 0.84 - 0.96;\(^{23}\) 0.97;\(^{21}\) 0.89- 0.96;\(^{23}\) 0.56 – 0.94;\(^{24}\) 0.81 – 0.94\(^{20}\)  
**Inter-rater reliability:**  
ICC = 0.76 - 0.92;\(^{23}\) 0.97\(^{21}\) | Measurement does not indicate pathology. Variability within and between participants is normal. |
| Lennie Test  
Modified Lennie  
Inclinometer | | | |
| **Movement Assessment** | | | |
| Serial Positional Observation | 2B | **Inter-rater reliability:**  
κ = 0.63 (unloaded),\(^{18}\)  
κ = 0.36 (loaded)\(^{18}\) | This assessment has some level of descriptive value of the scapula in different positions, but lacks norms, and validation. |
| Lateral Scapular Slide Test | 2B | **Intrarater reliability without dysfunction:**  
ICC = 0.75 – 0.80\(^{30}\)  
ICC = 0.94 – 0.97\(^{31}\)  
**Intrarater reliability with dysfunction:**  
ICC = 0.52 – 0.66\(^{30}\)  
ICC = 0.87 – 0.96\(^{31}\)  
**Interrater reliability without dysfunction:**  
ICC = 0.43 – 0.74\(^{30}\)  
ICC = 0.20 – 0.82\(^{29}\)  
ICC = 0.92 – 0.95\(^{31}\)  
ICC = 0.58 – 0.63\(^{15}\)  
(Superior Kibler) ICC = 0.55 – 0.87\(^{29}\)  
**Interrater reliability with dysfunction:**  
ICC = 0.45-0.79\(^{30}\)  
ICC = 0.70 – 0.95\(^{17}\)  
ICC = 0.63 – 0.86\(^{31}\)  
**Concurrent validity (with x-ray):**  
r = 0.91\(^{27}\)  
**Sensitivity:** 28 – 50%\(^{10}\) 80 – 96%\(^{11}\)  
**Specificity:** 52 – 58%\(^{30}\), 26.7%\(^{21}\), 4 – 26%  
**Positive LR:** 0.94 – 1.22  
**Negative LR:** 0.21 – 2.5 | This assessment has some level of descriptive value of the scapula in different positions, but research as indicated that asymmetry is not pathological, and no real norms of scapular position exist. |

(Continued On Page 14)
Upward/downward rotation

Five studies examined methods to assess upward and downward rotation of the scapula.\textsuperscript{21,23-26} Upward and downward rotation can be measured using the Lennie test and its modification, calculating distances from scapular landmarks to the inferior angle of the spine, as well as through use of an inclinometer along the spine of the scapula. This measure was rated 2B (unable to recommend at this time).

Movement Assessment

Lateral Scapular Slide Test

By far the most studied test of scapular dyskinesis, the Lateral Scapular Slide Test (LSST), was examined by 7 researchers.\textsuperscript{17,18,27-31} The LSST measures scapular motion with the arm in 3 positions: (1) glenohumeral neutral with arms at side, (2) hands on hips with arms in humeral internal rotation with 45° abduction, and (3) arms in maximal humeral internal rotation with arms elevated to 90° abduction. The distance from the inferior angle of the scapula to the adjacent spinous process is measured and recorded for each side. A bilateral difference of 1.5 cm or more is indicative of scapular dyskinesis. The recommendation rating was 2B (unable to recommend at this time).

Serial positional observation

One researcher examined the movement of the scapula in serial, static positions as the upper extremity is moved into elevation, with and without loads.\textsuperscript{18} These positions were the same as those in the LSST. The scapular motion was observed and if any one of the following 5 types was noted, the scapula was then determined to be impaired: (1) prominent inferior angle, (2) prominent medial border, (3) protraction, (4) elevation/depression, (5) medial border parallel to spine only at rest. Serial positional observation was rated 2B (unable to recommend at this time).

Dynamic movement assessment

This method of scapular motion assessment requires visual observation of the scapula while the upper extremities are actively moving through an elevation motion. Initially described by Kibler and Uhl,\textsuperscript{32} this test requires the examiner to categorize scapular motion while observing the scapula during active elevation. One of 4 descriptive categories is assigned to each scapula. Inferior angle (type I) is characterized by a prominent inferior angle of the scapula with humeral elevation. Medial border (type II) is characterized by a prominent medial border with humeral elevation. Superior border (type III) is characterized by scapular elevation and anterior displacement without winging during humeral elevation. Symmetric scapulohumeral (type IV) is considered normal. This visual categorization test was modified by two different researchers (Uhl et al\textsuperscript{33} and McClure et al\textsuperscript{34}) into a yes/no categorization (scapular dyskinesis is either present or absent). This test in the modified form was given a recommendation rating of 3.

DISCUSSION

The scapula is difficult to measure statically or dynamically, in part because of its multiplanar movements. However, abnormal scapular motion has been implicated in shoulder pathology in multiple populations. Researchers have examined scapular kinematics in relationship to shoulder impingement and documented that abnormal mechanics exist in the group with pathology when compared to controls with healthy shoulders.\textsuperscript{11,13,35,36} Furthermore, researchers have established that changes exist in scapular mechanics in individuals treated for breast cancer using 3D kinematic analysis. It is important, therefore, for clinicians to have measures that accurately assess the scapula. Valid and reliable clinical measures of the scapula will facilitate appropriate surveillance in individuals treated for breast cancer. Surveillance over time should be conducted in this population to identify changes that may lead to later pathology. Such measures are also necessary to document changes with intervention, a necessity in determining appropriate treatment strategies to be used when shoulder problems arise.

In assessing the quality of scapular assessment techniques in the literature, no assessment method was assigned a recommendation rating of 4 as none of the scapular assessment tools were specifically designed for the breast cancer population and none have any been validated in this population. Furthermore, the complex nature of scapular mechanics makes it quite challenging to create clinical tools to assess the scapula, with most available tools lacking good psychometric properties, demonstrated clinical usefulness, or both.
Positional Assessment

The static measures of scapular position received a recommendation rating of 2B despite good psychometric properties. For anterior/posterior tipping, the ICCs ranged from good to excellent. Scapular measures of protraction and retraction demonstrated ICCs for interrater reliability that are good to excellent, and concurrent validity with radiographs also good. Upward/downward rotation psychometric assessment resulted in good to excellent interrater reliability measures, and fair to good concurrent validity. Static scapular tests may adequately discriminate side-to-side differences in scapular position or posture. However, no one has been able to establish what constitutes normal values for scapular positions or that the presence of scapular asymmetries is indicative of pathology. Data across studies confirm substantial variability even among individuals without shoulder dysfunction. Most importantly, the authors of the papers included in this study often do not agree on what they are measuring. Nijs et al.17 identified their assessment of posterior acromion to the table simply as a measure “scapular position,” while Struyf et al.18 considered acromion to wall to be a measure of protraction. Plafcan et al.19 considered their measure of “posterior displacement of the scapula” to be inseparable combination of scapular winging and tipping. When the measures by different authors were categorized for this paper, an attempt was made to determine the plane of scapular measurement based on what was done rather than what the authors claimed. However, it is inarguable the planes of the scapula cannot be isolated, and that raises important issues about the validity and utility of these static measurements in spite of, in some cases, good psychometric properties.

Movement Assessment

Serial positional observation would appear to be a step toward identifying whether abnormal scapular motion is present. Like static measures, however, lack of clinical utility resulted in a recommendation rating of 2B. Kappa values for serial positional observation indicate substantial agreement in testing when the UE was unloaded. However, in loaded testing, agreement was only fair.

Unlike serial positional observation, the LSST has been validated against radiographs, and its reliability has been extensively studied. Although there was a high correlation between the LSST and radiograph measurements,27 the sensitivity and specificity of this test remain lower, and the ICCs for interrater reliability vary greatly from poor to good. It is important to recognize that serial static testing has the same limitations as static testing, where investigators are taking linear measurements of a 3D phenomenon. The resulting recommendation for serial static testing was 2B.

The serial positional assessment and the LSST were designed to capture changes in scapular position throughout humeral elevation. However, serial static positions do not provide adequate information about what the scapula does in dynamic activity. Only dynamic movement assessment, among the clinical assessment strategies for the scapula, includes performance of the scapula throughout its range of motion.

Visual assessment of scapular dyskinesis as done in dynamic movement assessment has evolved in the last decade and holds the most promise as a screen to determine if the movement of the scapula is abnormal and, therefore, a potential source of shoulder dysfunction. Kibler et al.32 first identified a clinical method to evaluate scapular dyskinesis in an active manner. The 4-level categorization of scapular motion by Kibler and colleagues demonstrated Kappa values in the poor to adequate range. When the test was modified into a yes/no categorization,33 Kappa values improved. Furthermore, a positive predictive value of 74% and a sensitivity of 76% indicate that this method may have clinical value as a screen for scapular dysfunction.33 Validation of this method by 3D motion analysis makes this a good clinical test that can be immediately put into practice as a screening tool,37 although validation in the breast cancer population still needs to be done.

Limitations to the Study

As is true for all systematic reviews, this study is limited by the possibility of incomplete ascertainment of relevant papers. Additional limitations include the varied populations that were used, as well as the varied measurement methods and terminology used across papers.

Recommendations for the Future

Dynamic movement assessment of the scapula can be recommended as components of shoulder examination. However, further information is needed relative to its value in the breast cancer population. Most scapular tests have been investigated in those with impingement or in over-head athletes, and it is possible that the breast cancer population may, given the accompanying treatments such as surgery and radiation, present with different mechanisms and sources of pain. The challenge to develop valid, reliable, and clinically feasible measures that objectively measure the multiplanar components of scapular position and motion remains unanswered.

CONCLUSION

The role of the scapula in normal shoulder function has been well established, and the presence of scapular dyskinesis has been documented in relation to shoulder impingement. Many scapular tests have been proposed to describe abnormal scapular mechanics, but visual observation of whether the scapula is moving in a normal or abnormal fashion appears to be the most reliable and useful clinical tool at this time.

REFERENCES

3. Rietman JS, Dijkstra PU, Geertzen JH, et al. Treatment-related upper limb morbidity 1 year after sentinel lymph


---

**PRISMA 2009 Checklist**

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
<td>1</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured summary</td>
<td>2</td>
<td>Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.</td>
<td>2</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
<td>3-4</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
<td>4</td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol and registration</td>
<td>5</td>
<td>Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.</td>
<td>NA</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.</td>
<td>4-5</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.</td>
<td>4-5</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
<td>4-5</td>
</tr>
<tr>
<td>Study selection</td>
<td>9</td>
<td>State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
<td>5</td>
</tr>
<tr>
<td>Data collection process</td>
<td>10</td>
<td>Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
<td>5</td>
</tr>
<tr>
<td>Data items</td>
<td>11</td>
<td>List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.</td>
<td>See EDGE form in Intro Appendix</td>
</tr>
<tr>
<td>Risk of bias in individual studies</td>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.</td>
<td>None specified</td>
</tr>
<tr>
<td>Summary measures</td>
<td>13</td>
<td>State the principal summary measures (e.g., risk ratio, difference in means).</td>
<td>NA</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>14</td>
<td>Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $i^2$) for each meta-analysis.</td>
<td>NA</td>
</tr>
<tr>
<td>Section/topic</td>
<td>#</td>
<td>Checklist item</td>
<td>Reported on page #</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>15</td>
<td>Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).</td>
<td>None specified</td>
</tr>
<tr>
<td>Additional analyses</td>
<td>16</td>
<td>Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.</td>
<td>NA</td>
</tr>
<tr>
<td>RESULTS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study selection</td>
<td>17</td>
<td>Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.</td>
<td>p. 5-6 figure 1</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>18</td>
<td>For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.</td>
<td>6-7</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>19</td>
<td>Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).</td>
<td>None specified</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>20</td>
<td>For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.</td>
<td>Table 2</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>21</td>
<td>Present results of each meta-analysis done, including confidence intervals and measures of consistency.</td>
<td>NA</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>22</td>
<td>Present results of any assessment of risk of bias across studies (see Item 15).</td>
<td>NA</td>
</tr>
<tr>
<td>Additional analysis</td>
<td>23</td>
<td>Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).</td>
<td>NA</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary of evidence</td>
<td>24</td>
<td>Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).</td>
<td>8-11</td>
</tr>
<tr>
<td>Limitations</td>
<td>25</td>
<td>Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).</td>
<td>11</td>
</tr>
<tr>
<td>Conclusions</td>
<td>26</td>
<td>Provide a general interpretation of the results in the context of other evidence, and implications for future research.</td>
<td>11-12</td>
</tr>
<tr>
<td>FUNDING</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding</td>
<td>27</td>
<td>Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.</td>
<td>NA</td>
</tr>
</tbody>
</table>