

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Scanxiety: A scoping review about scan-associated anxiety
<b>AUTHORS</b>	Bui, Kim Tam; Liang, Roger; Kiely, Belinda; Brown, Chris; Dhillon, Haryana; Blinman, Prunella

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Leys, Christophe Université Libre de Bruxelles , Psychology
<b>REVIEW RETURNED</b>	30-Aug-2020

<b>GENERAL COMMENTS</b>	<p>Although one could have done other choice related to the inclusion criteria, all of them are well explained and sound. The review is well conducted, I do not have much to say. Just some minor comments:</p> <ul style="list-style-type: none"><li>- I am not sure the impact of the conclusions justify a scientific contribution (it seems a bit weak to me, but I leave that assessment to editors). Scanxiety seems just a term defining nothing more than anxiety related to a specific situation. I do not see any specific psychological process, nor any specific treatment. Therefore I do not see the stakes to distinguish between anxiety and scanxiety. So maybe authors could address a bit more thoroughly the relevance of this distinction.</li><li>- Authors included measure of IES, which does not assess anxiety but PTSD (intrusion, avoidance and hyper-arousal). Please remind that in the latest version of DSM PTSD has been removed from general anxiety syndrome (for what seems to be accurate reasons).</li></ul>
-------------------------	---

<b>REVIEWER</b>	Derry, Heather Weill Cornell Medicine
<b>REVIEW RETURNED</b>	23-Nov-2020

<b>GENERAL COMMENTS</b>	<p>In this scoping review, the authors examine the current quantitative research on scanxiety (levels, correlates, and interventions), an important and under-addressed experience in cancer care. The scoping review approach is needed and novel in the literature on scan-related anxiety, and reflects a large undertaking and detailed data obtained by the authors. The authors utilize methods guided by PRISMA-ScR to increase rigor and transparency, such as coding by multiple reviewers and standardized data abstraction forms, which are strengths. Yet, there are several concerns that limit the clarity and conclusions in the paper.</p>
-------------------------	--

	<p>Major points:</p> <ol style="list-style-type: none"><li>1. The title of the paper are somewhat inconsistent with the approach. I suggest omitting "systematic" from the title, as this is a scoping review. I also suggest replacing "What is scanxiety" – this is not specifically addressed, as authors do not develop a definition or conceptual clarity through this review (instead, they utilize a priori criteria to identify quantitative articles).</li><li>2. The authors acknowledge that scanxiety has not been consistently defined in current literature. Given this, the methods would benefit from further rationale for several eligibility criteria.<ol style="list-style-type: none"><li>a. For example, how was "around the time of the scan" defined?</li><li>b. The methods state that studies of follow-up scans were excluded. I suggest clarifying this point as it appears that some studies were included that would be assumed to be follow-up scans (e.g., staging, surveillance). If follow-up scans were indeed removed, the text would benefit from a more specific definition of these and rationale – the current explanation states that this would be "confounding" but it is unclear what would be confounded in this descriptive review.</li></ol></li><li>3. The selection of thresholds used for prevalence and severity is confusing as currently presented and not well-justified. For example, prevalence in Supp Table 3 appears to include both study-specific or validated cut-offs and the presence of "any degree" of anxiety. These are very different approaches and make it difficult to interpret the findings, since "above threshold" and "any" symptoms are grouped yet do not reflect the same construct. Providing more information about each study's approach and specific threshold used may be useful, and/or the difficulty ascertaining thresholds could be added as part of discussion and future direction.</li><li>4. In the conclusion, the authors interpret "low levels" and "numerically low" in several areas. In my opinion, this interpretation is overstated, since specific comparison to norms or psychometrics of the specific tools seem outside the scope of the review (and quantitative analyses were appropriately not conducted). Also, variability statistics are not presented or considered for severity.</li></ol> <p>Minor points:</p> <ol style="list-style-type: none"><li>1. Several PRISMA-ScR items could use clarification: Was the protocol registered? How was the data charting process developed and was software used (#10)?</li><li>2. The placement of "carcinoma" in Figure 1 seems inconsistent with the description of the search strategy in the protocol.</li><li>3. In the protocol, "experiences" of scanxiety are described</li></ol>
--	---

	<p>in the objectives, but this is not discussed in the results paper, and qualitative inquiries were excluded. How did the authors arrive at this change in approach?</p> <p>4. The broad strategy and heterogeneity are listed as limitations. I suggest reconsidering this point as the broad search is key to the goal of the review and provides important information. Other limitations (ex: generalizability or type of articles included or conclusions drawn from thresholds) would be useful to include.</p>
--	--

### VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

- 1. I am not sure the impact of the conclusions justify a scientific contribution (it seems a bit weak to me, but I leave that assessment to editors). Scanxiety seems just a term defining nothing more than anxiety related to a specific situation. I do not see any specific psychological process, nor any specific treatment. Therefore I do not see the stakes to distinguish between anxiety and scanxiety. So maybe authors could address a bit more thoroughly the relevance of this distinction.**

We agree that scanxiety may be the same or similar as generalised anxiety. However, as noted by Reviewer 2, scanxiety is ‘an important and under-addressed experience in cancer care’. It is a specific situation experienced by people with cancer, and warrants further investigation (even if this ultimately shows it is similar to generalized anxiety). We have personally observed scanxiety in our patients, have observed patients using social media or blogs to describe their experiences with scanxiety, and have noted online resources published by cancer organisations that specifically address scanxiety (as referenced in our Introduction). Due to these examples, we believe scanxiety is a clinically important problem and were thus encouraged to proceed with this scoping review.

- 2. Authors included measure of IES, which does not assess anxiety but PTSD (intrusion, avoidance and hyper-arousal). Please remind that in the latest version of DSM PTSD has been removed from general anxiety syndrome (for what seems to be accurate reasons).**

PTSD was included in the DSM under generalised anxiety until the latest version (DSM-5) was published in 2013. Of the six studies which used the IES, four were conducted and published prior to 2013. All studies used the IES as a measure of distress around the time of a scan, and the two studies published after 2013 (Bauml, 2016; Bancroft, 2020) specifically used the IES as a measure of scan-associated distress. In our review, we defined the measurement of scanxiety as any measure of anxiety, distress or worry occurring around the time of a scan (stated in the Methods). We believe this broad definition is necessary because there is no uniform terminology that describes scanxiety in the literature.

We have elaborated on this issue in our Discussion:

*Firm conclusions about prevalence and severity could not be drawn due to considerable methodological heterogeneity of the included studies, especially in relation to scanxiety measurement tools. None were designed and validated for scanxiety, and some tools and their thresholds were not designed and/or validated for anxiety*

Reviewer 2:

Major points:

1. **The title of the paper are somewhat inconsistent with the approach. I suggest omitting “systematic” from the title, as this is a scoping review. I also suggest replacing “What is scanxiety” – this is not specifically addressed, as authors do not develop a definition or conceptual clarity through this review (instead, they utilize a priori criteria to identify quantitative articles).**

The title has been changed to: ‘Scanxiety: A scoping review about scan-associated anxiety’.

**2. The authors acknowledge that scanxiety has not been consistently defined in current literature. Given this, the methods would benefit from further rationale for several eligibility criteria.**

- a. **For example, how was “around the time of the scan” defined?**

We have updated our Methods with:

*The measurement of scanxiety was defined as any measure of anxiety, distress or worry occurring around the time of a scan. This included any period before, during or after a scan where the scan was used as a reference point for the measurement of scanxiety.*

**b. The methods state that studies of follow-up scans were excluded. I suggest clarifying this point as it appears that some studies were included that would be assumed to be follow-up scans (e.g., staging, surveillance). If follow-up scans were indeed removed, the text would benefit from a more specific definition of these and rationale – the current explanation states that this would be “confounding” but it is unclear what would be confounded in this descriptive review.**

We have included the following sentence in our Methods to define the included cancer-related reasons:

*Cancer-related reasons included screening (detection of cancer in asymptomatic person), diagnosis (detection of cancer in symptomatic person), staging (determining extent of cancer in person with confirmed or suspected cancer), surveillance (detection of recurrence in person with cancer treated with curative intent) or monitoring (detection of progression in person with cancer treated with non-curative intent).*

We excluded studies where scans were performed to further investigate an initial positive screening result. The term ‘follow-up’ was confusing and has been removed.

We have updated our Methods with:

*Exclusion criteria were... studies of scans performed to investigate a positive initial screening result because the psychological experiences of asymptomatic persons facing a potential new cancer diagnosis may lead to higher anxiety than is attributable to scanxiety.*

**3. The selection of thresholds used for prevalence and severity is confusing as currently presented and not well-justified. For example, prevalence in Supp Table 3 appears to include both study-specific or validated cut-offs and the presence of “any degree” of anxiety. These are very different approaches and make it difficult to interpret the findings, since “above threshold” and “any” symptoms are grouped yet do not reflect the same construct. Providing more information about each study’s approach and specific threshold used may be useful, and/or the difficulty ascertaining thresholds could be added as part of discussion and future direction.**

We accept that our thresholds for prevalence and severity are arbitrary. The rationale for these thresholds was our desire to descriptively compare results between studies, where the heterogeneity between measurement tools prevented a meaningful quantitative comparison – particularly as some measurement tools and/or their pre-specified thresholds were also not validated.

To enable transparency of our results and conclusions, we reported results from the purpose-designed definitions separately from studies with pre-specified thresholds, and we provided all raw data in Table 3 which shows each study, the scanxiety tool used and the thresholds where provided.

We have updated our Methods with:

*The definitions of prevalence and severity were purposive-designed to allow descriptive comparisons between the studies as we anticipated heterogeneity in scanxiety measurement would preclude meaningful summary statistics.*

We have updated our Discussion in two sections:

1) *This review did use purpose-designed definitions of prevalence and severity to allow some comparison between studies; however, the lack of a universal definition or specific measurement tool for scanxiety limits confidence in the interpretation of the results and interstudy comparisons.*

and;

2) *Limitations include the use of purpose-designed definitions of prevalence and severity*

**4. In the conclusion, the authors interpret “low levels” and “numerically low” in several areas. In my opinion, this interpretation is overstated, since specific comparison to norms or psychometrics of the specific tools seem outside the scope of the review (and quantitative analyses were appropriately not conducted). Also, variability statistics are not presented or considered for severity.**

We agree. We have adjusted the strength of our wording to reflect the uncertainty of our conclusions, and have removed our numerical interpretation of the pre- to post-scan scanxiety levels. We do not believe variability statistics can be meaningfully reported because of the frequent use of unvalidated tools and/or thresholds.

Minor points:

**1. Several PRISMA-ScR items could use clarification: Was the protocol registered? How was the data charting process developed and was software used (#10)?**

We did not register the protocol – as a scoping review, we were not aware of where we could register the protocol. We included the original protocol for transparency of our methods.

The data charting process was developed based on the objectives of the review, which themselves were developed in consultation with the research team of medical oncologists, a behavioural scientist and a statistician.

We have also added the following sentence to the Methods:

*Relevant data were independently extracted by two authors (KTB, RL) into an electronic data extraction form in Microsoft Excel.*

**2. The placement of “carcinoma” in Figure 1 seems inconsistent with the description of the search strategy in the protocol.**

Correct. This was a typographic error and has been amended.

**3. In the protocol, “experiences” of scanxiety are described in the objectives, but this is not discussed in the results paper, and qualitative inquiries were excluded. How did the authors arrive at this change in approach?**

We did initially intend to include qualitative experiences of scanxiety. However, we realised this was not feasible given the volume of data, differences in analysis approach for quantitative and qualitative data, and our own limited resources. As per scoping review methodology, we developed our exclusion criteria after familiarising ourselves with the literature, and decided then to exclude qualitative data.

We have updated our Methods with:

*Due to feasibility of conducting quantitative and qualitative analysis with the volume of literature identified, studies reporting only a qualitative assessment of scanxiety were also excluded, and the objective to explore patient experiences was abandoned.*

We have also provided a Supplementary File which summarises the changes between the original protocol and final methodology.

4. The broad strategy and heterogeneity are listed as limitations. I suggest reconsidering this point as the broad search is key to the goal of the review and provides important information. Other limitations (ex: generalizability or type of articles included or conclusions drawn from thresholds) would be useful to include.

Thank you for your insight. We have amended limitations in our Discussion to incorporate your recommendations.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Leys, Christophe Université Libre de Bruxelles , Psychology
<b>REVIEW RETURNED</b>	15-Feb-2021

<b>GENERAL COMMENTS</b>	All comment have been correctly addressed.
-------------------------	--

<b>REVIEWER</b>	Derry, Heather Weill Cornell Medicine
<b>REVIEW RETURNED</b>	02-Mar-2021

<b>GENERAL COMMENTS</b>	<p>The authors have been largely responsive to the points raised in review, which has strengthened the paper. The paper continues to address an important topic and I believe the unique contribution and extensive review will be of interest to readers. There are several notable remaining issues:</p> <p>1. Most notably is the presentation of results. The authors still combine estimates for “any” and “pre-specified cutoffs” to indicate prevalence. These approaches refer to much different symptom levels/interpretations and it is at times difficult to tell what a specific score is referring to in Table 3. It also is not appropriate to interpret these estimates in similar ways (page 18); estimates should be reported separately for those reporting “any” vs. “above threshold” anxiety symptoms.</p> <p>In fact, the distinction in presenting prevalence vs. severity also seems somewhat arbitrary, because some of the studies summarized for severity measures actually seem to fit under prevalence (e.g., “moderate to severe, 25%”) given lack of clarity for each.</p> <p>To address these, one solution would be instead providing one brief summary phrase for each study (without separating into severity or prevalence), such as “study findings” or “estimate obtained.”</p> <p>2. Similarly, the second method for determining severity (ex--“any score that was at least half the total score”) is unusual – many self-report instruments have cut-offs for</p>
-------------------------	--

	<p>clinically-significant scores that are less than half the total score. It would be helpful to list rationale or citations for this approach – or again de-emphasize this by focusing on descriptively relaying each paper’s finding.</p> <p>3. The paper still references “low” scores and “not severe” in some areas, and it is unclear what norms this statement is based on. This broad statement also seems premature since scores could not be compared across studies and the included studies have quite the range of time periods for when anxiety was assessed.</p> <p>4. Revisiting the abstract to increase precision with respect to the above points would be helpful.</p> <p>5. Table 3 – it is unclear if Range refers to observed range of scores or possible range of instruments.</p> <p>6. The first paragraph of the Discussion seems to use conclusive language (“interventions are more likely...”) although quantitative comparisons were not made.</p> <p>6. Table 6 – the p-values are not very useful without other context, and the total N could be interpreted as though the authors made quantitative comparisons across studies/the total N of participants, but I don’t think this was the case. Listing the N with each study might be clearer.</p> <p>7. The authors made several deletions on page 6 about the observational studies’ methods – I felt these were useful to include given the goals of the paper, if space is not an issue.</p>
--	---

### VERSION 2 – AUTHOR RESPONSE

Comments from Reviewer 2:

1. **Most notably is the presentation of results. The authors still combine estimates for “any” and “pre-specified cutoffs” to indicate prevalence. These approaches refer to much different symptom levels/interpretations and it is at times difficult to tell what a specific score is referring to in Table 3. It also is not appropriate to interpret these estimates in similar ways (page 18); estimates should be reported separately for those reporting “any” vs. “above threshold” anxiety symptoms.**

**In fact, the distinction in presenting prevalence vs. severity also seems somewhat arbitrary, because some of the studies summarized for severity measures actually seem to fit under prevalence (e.g., “moderate to severe, 25%”) given lack of clarity for each. To address these, one solution would be instead providing one brief summary phrase for each study (without separating into severity or prevalence), such as “study findings” or “estimate obtained.”**

We have added the following footnote to Table 3:

*NA is listed as the anxiety threshold when the study did not state a pre-specified threshold. In these cases, the definition of scanxiety prevalence was the percentage of people who reported any degree of anxiety, and the definition of scanxiety severity was at least half the total instrument score*

In the manuscript text (page 18), the prevalence estimates for ‘above threshold’ and for ‘any’ anxiety are reported separately.

We believe we have presented the results in the most objective manner by stating the raw data from each study in Table 3. Summary phrases for each study may introduce interpretation bias. In the Results, we provide a range for prevalence and descriptively report severity, rather than providing summary statistics such as a mean or median. In the Discussion, we state that firm conclusions about prevalence and severity cannot be drawn and provide reasons for this.

2. **Similarly, the second method for determining severity (ex--“any score that was at least half the total score”) is unusual – many self-report instruments have cut-offs for clinically-significant scores that are less than half the total score. It would be helpful to list rationale or citations for this approach – or again de-emphasize this by focusing on descriptively relaying each paper’s finding.**

As stated in the Methods, the definitions for prevalence and severity were purpose-designed to allow descriptive comparison between studies. The second definition was only used where no threshold was stated, and a level of ‘at least half’ was used to avoid over-reporting severity.

3. **The paper still references “low” scores and “not severe” in some areas, and it is unclear what norms this statement is based on. This broad statement also seems premature since scores could not be compared across studies and the included studies have quite the range of time periods for when anxiety was assessed.**

We used equivocal terminology when describing scanxiety severity (e.g., scanxiety severity appeared low) rather than definitive wording (e.g., scanxiety severity was low). We believe readers will appreciate the uncertainty around the level of severity because we provide a detailed explanation of the ‘high’ severity findings in the Results (i.e., measurement tool used, the mean score and the pre-specified threshold, if applicable, used in each study) and because we provide commentary about why firm conclusions about scanxiety severity cannot be made in the Discussion. In the Conclusions, we deliberately did not include an estimate or subjective statement on scanxiety prevalence and severity.

Timing of scanxiety assessment has been added specifically as a limitation:

*Limitations include... the limited generalisability of the results due to heterogeneity in cancer type, reason for scan, imaging modality and timing of scanxiety measurement between the studies*

4. **Revisiting the abstract to increase precision with respect to the above points would be helpful.**

We have not changed the abstract for the reasons detailed above.

5. **Table 3 – it is unclear if Range refers to observed range of scores or possible range of instruments.**

To improve clarify on this, the column headings for Table 3 have been amended to:

Measurement of scanxiety			Results of scanxiety measurement		
Name of tool	Range of tool (Anxiety threshold <sup>a</sup> )	Timing of assessment	Prevalence (%)	Severity (Mean <sup>b</sup> )	Pre- & post-scan comparison

The column headings for Table 4 have also been amended to:

Measurement of scanxiety			Impact of intervention on scanxiety	
Name of tool	Range of tool (Anxiety threshold)	Timing of assessment	Description of results	P-value

6. The first paragraph of the Discussion seems to use conclusive language (“interventions are more likely...”) although quantitative comparisons were not made.

We have amended this to:

*Interventions **may be** more likely to reduce scanxiety if they...*

7. Table 6 – the p-values are not very useful without other context, and the total N could be interpreted as though the authors made quantitative comparisons across studies/the total N of participants, but I don’t think this was the case. Listing the N with each study might be clearer.

We have added the following footnote to Table 6:

*The P-values listed in this table were reported by individual studies based on their own datasets. This scoping review has not performed additional analysis or attempted quantitative comparisons between studies.*

We believe stating the N for each study may detract from the readability of the study.

8. The authors made several deletions on page 6 about the observational studies’ methods – I felt these were useful to include given the goals of the paper, if space is not an issue.

Unfortunately, the word count is an issue. These details are included in Table 3.

### VERSION 3 – REVIEW

<b>REVIEWER</b>	Derry, Heather Weill Cornell Medicine
<b>REVIEW RETURNED</b>	26-Mar-2021
<b>GENERAL COMMENTS</b>	<p>The authors have responded to the majority of issues raised, and again the extensive review is valuable. My remaining recommendations involve resolving confusion around the definition of severity.</p> <p>1. In the Table 3 footnote: “definition of scanxiety severity</p>

	<p>was at least half the total instrument score” – It is not clear what this means. Symptom severity would typically be reported as the mean and standard deviation (or similar statistics) of the scale. Is that what the authors summarize in Table 3? It does appear that way. If that is the case (if the mean score in each sample is what’s actually represented -- if not this is an issue), this part of the footnote should be removed, and I believe that would resolve the confusion -- the SD’s are not included, which limits an otherwise useful summary and it is unclear why they are not present.</p> <p>2. Regarding separating the estimates of prevalence for any vs. above-threshold symptoms: The section “between 13% and 83% using the 14 measures without pre-specified anxiety thresholds” is still unclear that this is referring to rates of any anxiety. Recommend stating this more clearly, such as “Prevalence of any anxiety symptoms ranged from” and “prevalence of anxiety symptoms above pre-specified thresholds ranged from” and likewise this would improve clarity in the abstract, where prevalence is still reported as one range (0 to 83) .</p> <p>3. Similarly, in the abstract/text “mean severity scores appeared low” – It would be helpful to know what defines low in this review such as “i.e., below pre-specified thresholds and/or purpose-driven criteria selected for this review ”</p>
--	--

### VERSION 3 – AUTHOR RESPONSE

Comments from Reviewer 2:

- 1. In the Table 3 footnote: “definition of scanxiety severity was at least half the total instrument score” – It is not clear what this means. Symptom severity would typically be reported as the mean and standard deviation (or similar statistics) of the scale. Is that what the authors summarize in Table 3? It does appear that way. If that is the case (if the mean score in each sample is what’s actually represented -- if not this is an issue), this part of the footnote should be removed, and I believe that would resolve the confusion -- the SD’s are not included, which limits an otherwise useful summary and it is unclear why they are not present.**

We have removed the statement ‘*definition of scanxiety severity was at least half the total instrument score*’ from Table 3.

We have added the standard deviations to all severity scores in Table 3, where this has been provided in the research article.

- 2. Regarding separating the estimates of prevalence for any vs. above-threshold symptoms: The section “between 13% and 83% using the 14 measures without pre-specified anxiety thresholds” is still unclear that this is referring to rates of any anxiety. Recommend stating this more clearly, such as “Prevalence of any anxiety symptoms ranged from” and “prevalence of anxiety symptoms above pre-specified**

**thresholds ranged from” and likewise this would improve clarity in the abstract, where prevalence is still reported as one range (0 to 83).**

This Results section has been amended to say:

*The prevalence of scanxiety above pre-specified anxiety thresholds ranged between 0% and 64% across the 16 measures, though eight of these measures came from only two studies. In the 14 measures without a pre-specified anxiety threshold, the prevalence of any degree of scanxiety ranged between 13% and 83%.*

The Abstract has been amended to say:

*Scanxiety prevalence ranged from 0% to 64% (above pre-specified thresholds) or 13% to 83% (‘any’ anxiety, if no threshold).*

**3. Similarly, in the abstract/text “mean severity scores appeared low” – It would be helpful to know what defines low in this review such as “i.e., below pre-specified thresholds and/or purpose-driven criteria selected for this review ”**

In the Results section, the statement ‘mean severity scores appeared low...’ is immediately preceded by two paragraphs. The first paragraph states:

*The mean severity scores were below pre-specified anxiety thresholds on 17 of the 19 measures where a threshold was reported.*

The second paragraph states:

*Of the 43 measures without a pre-specified threshold, the majority had mean scores that were less than half the total scores.*

The Abstract has been amended to:

Mean severity scores appeared low in almost all measures which quantitatively measured scanxiety (54/62, 87%), regardless of whether anxiety thresholds were pre-specified.