Symptom clusters in patients receiving chemotherapy: A systematic review

Carolyn S. Harris,1 Kord M. Kober,1 Yvette P. Conley,2 Anand A. Dhruva,3 Marilyn J. Hammer,4 Christine A Miaskowski1,3

ABSTRACT

Background and purpose Since 2001, symptom cluster research has grown considerably. However, because multiple methodological considerations remain, ongoing synthesis of the literature is needed to identify gaps in this area of symptom science. This systematic review evaluated the progress in symptom clusters research in adults receiving primary or adjuvant chemotherapy since 2016.

Methods Eligible studies were published in English between 1 January 2017 and 17 May 2021; evaluated for and identified symptom clusters ‘de novo;’ and included only adults being treated with primary or adjuvant chemotherapy. Studies were excluded if patients had advanced cancer or were receiving palliative chemotherapy; symptoms were measured after treatment; symptom clusters were pre-specified or a patient-centred analytic approach was used. For each study, symptom instrument(s); statistical methods and symptom dimension(s) used to create the clusters; whether symptoms were allowed to load on more than one factor; method used to assess for stability of symptom clusters and associations with secondary outcomes and biomarkers were extracted.

Results Twenty-three studies were included. Memorial Symptom Assessment Scale was the most common instrument and exploratory factor analysis was the most common statistical method used to identify symptom clusters. Psychological, gastrointestinal, and nutritional clusters were the most commonly identified clusters. Only the psychological cluster remained relatively stable over time. Only five studies evaluated for secondary outcomes.

Discussion While symptom cluster research has evolved, clear criteria to evaluate the stability of symptom clusters and standardised nomenclature for naming clusters are needed. Additional research is needed to evaluate the biological mechanism(s) for symptom clusters.

Key messages

What was already known?
- Gastrointestinal and psychological clusters are common.

What are the new findings?
- Nutritional cluster was commonly identified.

What is their significance?
- Psychological clusters are relatively stable over time.
- Gastrointestinal clusters appear to fluctuate over time.

INTRODUCTION

As the incidence of new cancer cases and mortality rates increase globally,1 the symptom burden of oncology patients remains high. For example, in one study,2 50% of patients receiving chemotherapy experienced an average of 13 symptoms. Equally important, co-occurring symptoms and/or symptom clusters result in increased distress,3 decreased functional status,4 poorer quality of life (QOL)5 and increased mortality.6,7 Given that 50% of oncology patients may experience these negative effects, research on how and why symptoms co-occur is vital to the development of effective interventions.

In 2001, Dodd and colleagues8 were the first to introduce the concept of a symptom cluster into oncology symptom science. Since then, symptom cluster research has increased dramatically.9–12 While the definition of a symptom cluster has evolved,9,13 most recently, it was defined as the co-occurrence of two
or more symptoms that are stable and independent of other clusters, and may share underlying mechanisms and/or outcomes. This research has grown to include the identification of symptom clusters in children and adolescents; in patients with advanced cancer; and in patients receiving active treatment. An emerging area of research is the evaluation of biomarkers and molecular mechanisms associated with symptom clusters.

While this research provides important foundations in our understanding of cancer-related symptom clusters, two key methodological issues remain unresolved; namely: which statistical approach provides the most consistent identification of symptom clusters (eg, factor analysis, exploratory factor analysis (EFA)) and how the dimension(s) of the symptom experience that are used to create the clusters (ie, occurrence, severity, frequency, distress) influence the number and types of symptom clusters identified. Resolution of these issues is key to the development of effective interventions for symptom clusters. In addition, consistent identification of symptom clusters will facilitate the investigation of their underlying mechanisms.

While Skerman and colleagues suggested that factor analysis methods were the optimal approach to create symptom clusters, and more recently network analysis (NA) have been used. Factor analysis methods, like EFA, are used to identify latent constructs or factors (ie, symptom clusters) that account for the strength of the relationships between variables (ie, symptoms). This type of factor analysis is exploratory in nature as it does not test hypotheses on the nature of the relationships among the variables. Cluster analysis methods (eg, hierarchical cluster analysis (HCA)), use measures of correlation or distance to group related variables (ie, symptoms). An emerging analytical approach for identifying symptom clusters is NA. With this approach, relationships between multiple variables or nodes (ie, symptoms) are quantified and illustrated graphically. Unique strengths of NA are its potential to identify ‘core’ symptoms (ie, symptoms that have a high impact on the network or cluster) and relationships among symptom clusters.

Consensus is lacking on which dimension(s) (ie, occurrence, severity, frequency, distress) of the symptom experience should be used to identify symptom clusters. For example, in one review, a significant amount of variability was found in the dimensions used to identify symptom clusters. This type of evaluation is important because the specific dimension used may influence the number, types and composition of the symptom clusters that are identified, making comparisons across studies difficult. While each symptom dimension provides unique information, little is known about how the symptom clusters identified using different dimensions may affect various patient outcomes or the mechanisms that underlie various symptom clusters.

In the most recent review of symptom clusters research in oncology patients receiving chemotherapy, findings from studies published between 2000 and 2016 were synthesised. However, the impact of symptom clusters on outcomes (eg, QOL, functional status) and associations with underlying mechanisms were not evaluated. As noted in an expert panel report, ongoing synthesis of symptom clusters research is warranted to identify gaps in this area of scientific inquiry. Therefore, the purpose of this systematic review was to evaluate the progress in symptom clusters research in adult patients receiving primary or adjuvant chemotherapy since 2016. Specifically, this paper will: (1) describe the most common instrument(s), statistical approaches and symptom dimensions used to evaluate symptom clusters; (2) describe the number and types of symptom clusters identified using different dimensions of the symptom experience; (3) determine whether symptom clusters change over time; and (4) describe associations between symptom clusters and patient-reported outcomes (PROs) and biological mechanisms.

**METHODS**

**Search strategy**

This review was performed using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Studies that were published between 1 January 2017 and 17 May 2021 were retrieved from the Cochrane Library, Cumulative Index to Nursing and Allied Health Literature, Embase, PubMed and Web of Science databases. The search strategy for each database is listed in table 1.

**Study selection**

Identified studies were downloaded into a prespecified Endnote Library for review and duplicates were removed. Studies were retained for review if they met the following eligibility criteria: (1) evaluated for and identified at least one symptom cluster; (2) included only adults (aged ≥18 years); (3) included only oncology patients who were being treated with primary or adjuvant chemotherapy; (4) were published in English; (5) had a cohort, case-control, cross-sectional or longitudinal design; and (6) identified symptom clusters ‘de novo’ (ie, used a statistical method to identify clusters). Studies were excluded if they: (1) were published prior to 1 January 2017; (2) included patients with advanced cancer (ie, stage IV) or those receiving palliative chemotherapy; (3) measured symptoms after the completion of treatment; (4) used pre-specified symptom clusters (ie, did not use a statistical method to identify clusters); (5) used a patient-centred analytic approach (eg, latent class analysis); or (6) were a systematic review, meta-analysis, conference abstract, dissertation work, case-report or qualitative study. The title and abstract of each study were reviewed by a single author (CSH) for eligibility based on our prespecified inclusion and exclusion criteria.
The first (CSH) and senior (CAM) authors reviewed the full text of the remaining articles against the inclusion and exclusion criteria.

**Data extraction**
The prespecified study characteristics that were extracted are detailed in box 1. Separate evaluations were done for cross-sectional (online supplemental table 1) and longitudinal (online supplemental table 2) studies. Two reviewers (CSH and CAM) independently reviewed each study and consensus was reached on the data included in the tables.

**Assessment of methodological quality**
Each study’s methodological quality was assessed using the National Heart, Lung, and Blood Institute’s (NHLBI) National Institute of Health Quality Assessment Tool for Observational and Cross-Sectional Studies. Questions on this tool were designed to enable researchers to critically appraise the internal validity of research studies. Each question is answered with ‘yes’, ‘no’ or ‘cannot determine, not reported or not applicable’. Items that receive a ‘no’ or indeterminable response are considered a study weakness that may introduce bias. As recommended by the NHLBI tool guidelines, this potential risk of bias must be further evaluated by a reviewer and is factored into the final rating of ‘good’, ‘fair’ or ‘poor’. Two reviewers (CSH and CAM) independently assessed the quality of each study and combined their results in a shared Excel spreadsheet. All studies that met the inclusion and exclusion criteria were included in this review regardless of the methodological quality assessment rating.

**RESULTS**

**Study selection**
The initial search resulted in 574 articles. Following the removal of duplicates, 319 articles remained. Next, the title and abstract of each study were reviewed against our inclusion and exclusion criteria and 283 studies were excluded. The first (CSH) and senior (CAM) authors reviewed the full text of the remaining 36 articles against the inclusion and exclusion criteria. Following these steps, 23 articles were retained for data extraction and are included in this systematic review (figure 1).

**Methodological quality of studies**
Nine of the 13 cross-sectional studies received a ‘good’ quality rating, four received a ‘fair’ rating and none received a poor rating (table 2). Across the four studies that received a ‘fair’ rating, two sources of bias were: lack of reporting of whether the participation rate of eligible persons was at least 50% (item 3) and lack of clarity on whether the timing of the symptom assessment around the receipt of chemotherapy was sufficient in order to see an effect (item 7). All of the longitudinal studies received a ‘good’ rating. Of note, seven of the

### Table 1 Summary of search strategy

<table>
<thead>
<tr>
<th>Database</th>
<th>Search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Library</td>
<td>‘symptom cluster’ OR ‘symptom clusters’ OR ‘symptom AND cluster’ OR ‘symptom AND clusters’ OR ‘symptoms AND clusters’ in All Text AND cancer OR neoplasm in All Text AND chemotherapy OR Ctx in All Text NOT reviews NOT protocols. Restricted to 1 January 2017 to 17 May 2017</td>
</tr>
<tr>
<td>Cumulative Index to Nursing and Allied Health Literature</td>
<td>‘symptom cluster’ OR ‘symptom clusters’ OR ‘symptom AND cluster’ OR ‘symptom AND clusters’ OR ‘symptoms AND clusters’ AND (cancer OR neoplasm) AND (chemotherapy OR Ctx). Limiters: Published date: 1 January 2017 to 31 May 2021; Language: English</td>
</tr>
<tr>
<td>Embase</td>
<td>‘symptom cluster’ OR ‘symptom clusters’ OR ‘symptom AND clusters’ OR ‘symptom AND cluster’ AND (cancer OR neoplasm) AND (chemotherapy OR Ctx). Search limited to 1 January 2017 to 17 May 2021; Language: English</td>
</tr>
<tr>
<td>PubMed</td>
<td>(((‘symptom cluster’[All Fields]) OR (‘symptom clusters’[All Fields]))) OR (((‘symptom’[All Fields]) AND (‘clusters’[All Fields]))) OR (‘symptoms’[All Fields]) AND (‘clusters’[All Fields])) AND ((cancer[All Fields]) OR (neoplasm[All Fields])) AND (chemotherapy[All Fields]) OR (ctx[All Fields])). Filter applied: 1 January 2017 to 17 May 2017; Language: English</td>
</tr>
<tr>
<td>Web of Science</td>
<td>Topic=(symptom cluster* OR <em>symptom clusters</em>) OR Topic=(symptom* AND cluster*) OR Topic=(symptom* AND clusters*) OR Topic=(symptoms* AND cluster*) OR Topic=(symptoms* AND clusters*) AND Topic=(cancer* OR neoplasm*) AND (chemotherapy* OR CTX*) AND Topic=(chemotherapy* OR Ctx*). Restricted to: 1 January 2017 to 17 May 2021; Language: English</td>
</tr>
</tbody>
</table>
10 longitudinal studies either lost >20% of patients to follow-up or did not report this information.

**Cross-sectional study results**

**Study characteristics**

Of the 23 studies included in this review, 13 used a cross-sectional design to identify symptom clusters in oncology patients receiving chemotherapy (online supplemental table 1). Seven studies were conducted in the USA, two in China, two in Thailand, one in Austria, and one in Turkey. Sample sizes ranged from 96 to 1328. Across these studies, the majority of patients were female (weighted grand mean 76.8%), outpatients, not working, had a weighted grand mean age of 55.0 years, and were relatively homogeneous in terms of ethnicity and race.

Five studies evaluated for symptom clusters in patients with heterogeneous types of cancer. Of the eight studies that evaluated for clusters in patients with homogeneous types of cancer, four evaluated patients with breast cancer, one with bladder cancer, one with leukaemia, one with lymphoma, and one with lung cancer.

**Symptom instrument(s)**

In terms of the instruments, nine of the 13 studies used the Memorial Symptom Assessment Scale (MSAS). Of these nine studies, six used a modified version of the MSAS and one used a condensed version. One study used multiple symptom assessment tools to assess for clusters; specifically, the Breast Cancer Prevention Trial Symptom Checklist, the Beck Depression Inventory-II, the Brief Pain Inventory, the Patient’s Assessment of Own Functioning, and the Profile of Mood States. One study each used the Edmonton Symptom Assessment Scale, the MD Anderson Symptom Inventory (MDASI), and the Rotterdam Symptom Checklist.

**Statistical approach**

Nine of the 13 studies used EFA to identify symptom clusters. Of the remaining studies, two used principal component analysis (PCA), one used HCA and one used NA.

**Symptom dimension(s)**

In terms of the symptom dimension(s), three of the 13 studies used only severity and three used only distress. Of the seven remaining studies, two used both occurrence and severity, one used severity and distress, one used frequency, severity and distress, and three used occurrence, severity, and distress.

**Occurrence**

Across the five studies that used occurrence, a psychological cluster was identified. The number of symptoms ranged from five to 12. Worrying, feeling nervous, feeling sad, and feeling irritable were common across the five studies. A respiratory or lung cancer-specific cluster was identified across three of the five studies. The number of symptoms ranged from four to nine. Shortness of breath, difficulty breathing and cough were common across the three studies.
Table 2  Quality assessment by the National Heart, Lung, and Blood Institute (NHLBI) of the National Institute of Health Quality Assessment Tool for Observational and Cross-sectional Studies

<table>
<thead>
<tr>
<th>Author(s), year</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
<th>Item 8</th>
<th>Item 9</th>
<th>Item 10</th>
<th>Item 11</th>
<th>Item 12</th>
<th>Item 13</th>
<th>Item 14</th>
<th>Final quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-sectional studies (n=13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen et al, 2021</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>CD</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Fair</td>
</tr>
<tr>
<td>Cherwin and Perkhounkova, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Chongkham-ang et al, 2018</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Han et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Li et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Matzka et al, 2018</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>CD</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Fair</td>
</tr>
<tr>
<td>Papachristou et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Pozzar et al, 2021</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Ren et al, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>CD</td>
<td>CD</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Fair</td>
</tr>
<tr>
<td>Sezgin and Bektas, 2020</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>CD</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Fair</td>
</tr>
<tr>
<td>Sullivan et al, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Vuttanon et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Wong et al, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Longitudinal studies (n=10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albusoul et al, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Berger et al, 2020</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Browall et al, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>NR</td>
<td>Good</td>
</tr>
<tr>
<td>Han et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>NR</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Kim, 2018</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Li et al, 2020</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>N</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Lin et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>N</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Russell et al, 2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>NR</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Sullivan et al, 2018</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>NR</td>
<td>NA</td>
<td>Good</td>
</tr>
<tr>
<td>Wiggenraad et al, 2020</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>N</td>
<td>NA</td>
<td>Good</td>
</tr>
</tbody>
</table>

Study methodological quality ratings: good, fair and poor.

NHLBI of the National Institute of Health Quality Assessment Tool for Observational and Cross-Sectional Studies Criteria: item 1 (clear research question); item 2 (define study population); item 3 (participation rate at least 50%); item 4 (uniform eligibility criteria); item 5 (sample size justification); item 6 (exposure assessed prior to outcome measurement); item 7 (sufficient timeframe to see an effect); item 8 (examine different levels of exposure); item 9 (clearly defined exposure measures); item 10 (exposure assessed more than once over time); item 11 (clearly defined outcome measures); item 12 (outcome assessors were blinded to exposure status of participants); item 13 (loss to follow-up less than 20%); item 14 (key confounding variables measured and adjusted statistically).

CD, cannot determine; NA, not applicable; n, no; NR, not reported; Y, yes.
A nutritional or weight change cluster was identified across all five studies. A type of nutritional cluster (ie, nutritional, image and nutrition, discomfort and nutrition, weight change) was identified across all six studies and dimensions.

A gastrointestinal cluster was identified in three studies. However, no common symptoms were identified across the three studies.

Severity
Ten studies used severity to evaluate for clusters. Of the eight studies that named the clusters, all identified a psychological cluster (ie, emotion-related, psychological, psychological). The number of symptoms ranged from two to nine. Feeling sad, sadness, or depression was the only symptom that was identified across all of the studies.

Six studies identified a cluster related to nutritional status or weight (ie, nutritional, weight, weight change). The number of symptoms ranged from two to six. Weight loss was the common symptom across all studies.

A gastrointestinal or gastrointestinal and energy related cluster was identified in five of the eight studies. The number of symptoms ranged from two to eight. While no symptoms were common across all studies, weight loss and weight gain were identified in five of them.

Distress
Eight studies evaluated for clusters using the distress dimension. Similar to occurrence and severity, a type of psychological cluster (ie, anxiety and depression, emotion, energy and pain related, emotions, psychological, psychological/gastrointestinal) was identified in seven of the studies that named the clusters.

The symptoms within this cluster ranged from three to 12. Feeling nervous or anxious and feeling sad or depressed mood were common symptoms across all seven studies.

Five studies identified a type of nutritional cluster (ie, appetite, nutritional, nutrition impaired, weight change). The symptoms ranged from two to seven. Lack of appetite was common across four of the five studies.

Multiple dimensions
Seven studies evaluated for differences in clusters across two or more symptom dimensions. 

Of the six studies that named the clusters, a type of psychological cluster (ie, emotion related, emotion, energy and pain related, psychological/gastrointestinal, psychological) was common across all six studies and dimensions. Feeling irritable, feeling nervous, feeling sad, and worrying were the common symptoms across the six studies and dimensions.

Evaluation of the stability of symptom clusters across symptom dimensions
Of the six studies that named the clusters and evaluated for clusters using two or more dimensions, all of them evaluated the stability of the clusters across dimensions. Five studies used the method described by Kirka and Walsh. The sixth study evaluated for stability through visualisation of differences in the network’s structures.

Analysis of secondary outcomes
In the four studies that evaluated for associations between clusters and other PROs, all of them used QOL. In addition, one evaluated for associations with patients’ functional status. None of the cross-sectional studies evaluated for associations between symptom clusters and biological mechanisms.

Longitudinal study results
Study characteristics
Of the 23 studies included in this review, 10 used a longitudinal design to evaluate for symptom clusters in oncology patients receiving chemotherapy (online supplemental table 2). Six studies were conducted in the USA, two in Sweden, one in China, and one in South Korea. Sample sizes ranged from 51 to 540. Across these studies, the majority of the patients were female (weighted grand mean 84.4%), currently employed, had a weighted grand mean age of 55.1 years, and were relatively homogeneous in terms of ethnicity and race.

Only one study evaluated for symptom clusters in a sample of patients with heterogeneous cancer diagnoses. Of the nine studies that evaluated for clusters in patients with homogeneous diagnoses, six evaluated patients with breast cancer, one with acute myelogenous leukaemia, one with brain cancer, and one with lung cancer.

Symptom instrument(s)
In terms of the instruments, seven of the 10 studies used the MSAS. Of these seven studies, three used a modified version of the MSAS. Two studies used the Hospital Anxiety and Depression Scale, the Symptom Experience Scale, and the Medical Outcomes Study Short-Form Survey v2.46 One study used the Breast Cancer Prevention Trial Symptom Checklist, the Beck Depression Inventory-II, the Brief Pain Inventory, the Patient’s Assessment of Own Functioning, and the Profile of Mood States.
Symptom dimension(s)
In terms of the symptom dimension(s), four studies used only the severity dimension. While two studies evaluated for clusters using both occurrence and severity, two used occurrence, severity, and distress. The remaining two studies created a symptom burden score (ie, the average of the frequency, severity, and distress scores for each symptom on the MSAS).

Occurrence
Four studies used occurrence to identify clusters across three timepoints. For three of these studies, these timepoints were: approximately one week before the second or third cycle of chemotherapy (T1), approximately one week after chemotherapy administration (T2), and approximately two weeks after chemotherapy administration (T3). For the fourth study, these timepoints were: within six days of the start of induction chemotherapy (T1a), one to seven days during induction chemotherapy (T2a), and one to seven days after induction chemotherapy (T3a).

A psychological cluster was identified across all four studies and all three timepoints, except for one study where the cluster was not identified until T2a. Feeling nervous and feeling sad were common across each study and timepoint. In addition, difficulty concentrating, feeling irritable, and worrying were common to the three studies that identified a psychological cluster at T1. Across these four studies, the symptoms within this cluster remained relatively consistent across time. While a nutritional or weight change cluster was identified across all four studies, it was not identified at each timepoint. For three of the studies, lack of appetite was present at T2 and lack of appetite and weight gain were present at T3. Except for one study, the symptoms identified within this cluster were relatively consistent across timepoints within each study. While an epithelial or epithelial/gastrointestinal, or body image cluster was identified across all four studies, it was not identified at each timepoint and the symptoms within this cluster changed over time. Hair loss was identified at T2 in three studies. Itching was identified at T3 and T3a in three studies. Changes in skin was identified across all four studies at T3 and T3a.

A gastrointestinal cluster was identified across three studies at one or more timepoints. However, this cluster was not identified at each timepoint and no common symptoms were consistent across each of the three studies. Abdominal cramps appeared across two of the studies that identified this cluster at T1.

Severity
Eight studies used severity to identify clusters across three or four timepoints. Of the two studies that evaluated for clusters over four timepoints, one evaluated for clusters throughout all cycles of chemotherapy (ie, prior to the first cycle to post-chemotherapy) and the other evaluated for clusters from prior to and at 18 months post-chemotherapy. Five of the remaining six studies evaluated for clusters over three timepoints around the receipt of active treatment (eg, prior to and post-chemotherapy). The sixth study evaluated for clusters after the completion of chemotherapy (ie, prior to chemotherapy to one year after initial chemotherapy treatment).

While no single cluster was common across the eight studies, a gastrointestinal cluster was identified across seven of them. This cluster was not identified across all timepoints and no common symptoms were identified. In addition, a type of psychological cluster (ie, negative emotion, negative emotion and decreased vitality, psychological, psychoneurocognitive) was identified in six of the eight studies. This cluster was not identified across all of the timepoints. However, when the cluster was identified, feeling sad or depression was consistent across all of the studies.

Distress
Only two studies evaluated for clusters using distress across three timepoints. A psychological cluster was identified across both studies and at two of the three timepoints. Across these timepoints, feeling nervous and feeling sad were consistent. While an epithelial or body image cluster was identified across both studies, it was not present across all three timepoints. When the cluster did occur, itching was identified across both studies and timepoints.

Burden score
In the two studies that used a symptom burden score to identify clusters, one evaluated for clusters over four timepoints across multiple cycles of chemotherapy, and the other evaluated for clusters over three timepoints prior to the start of the second cycle of chemotherapy to 12 months post cycle two. An emotional cluster was identified across both studies and timepoints. Feeling sad was common across both studies and all timepoints. While a physical cluster was identified across both studies and timepoints, no common symptoms were identified.

Multiple dimensions
Four studies evaluated for clusters using two or more dimensions over three timepoints. In three of these studies, a psychological cluster was...
identified across all of these studies, dimensions and timepoints. In the fourth study, this cluster occurred with some variability across timepoints and dimensions. Feeling nervous and feeling sad occurred consistently across studies, dimensions, and timepoints.

While an epithelial, epithelial/gastrointestinal, or body image cluster was identified across all four studies, it was not stable across dimensions or timepoints. Only changes in skin appeared across dimensions and studies at the third timepoint (ie, two weeks postcycle two or three, one to seven days after induction). In addition, gastrointestinal and nutritional or weight change clusters were identified across three of the four studies. No common symptoms were identified consistently across studies, dimensions and/or timepoints for either cluster.

Evaluation of the stability of symptom clusters across symptom dimensions and/or timepoints
Six studies used the method described by Kirkova and Walsh to evaluate the stability of symptom clusters across dimensions and timepoints. Two studies relied on an investigator’s appraisal of the stability. The remaining two studies did not report on a method to evaluate stability.

Analysis of secondary outcome(s)
In the only longitudinal study that evaluated for associations between symptom clusters and a PRO, measures of QOL were used. In the only study that evaluated for associations between symptom clusters and biological mechanisms, levels of lipid peroxidation were examined in patients with primary brain tumours.

DISCUSSION
This systematic review evaluated the progress of symptom clusters research in adult patients receiving primary or adjuvant chemotherapy from 2017 through 2021. Given the relative infancy of symptom cluster research, this type of ongoing review and synthesis is needed to advance this area of scientific inquiry. This discussion focuses on how the science has evolved since the previous review.

Symptom assessment instruments
The MSAS was the most common instrument used in 69.6% of the studies. While it was found to be one of the most commonly used instruments in the previous review, its use grew from 26.3% to 69.6%. This growth may be due to the multiple strengths of the MSAS. First, because it evaluates 32 common symptoms, it is cited as one of the most comprehensive instruments to use in research and clinical practice. In addition, the MSAS evaluates multiple dimensions of the symptom experience (ie, occurrence, severity, frequency, distress); has well established validity and reliability; and is available in more than eight languages (eg, Arabic, Chinese, Spanish).

In contrast with the previous review that noted that the MDASI was used in 26.3% of the studies, it was used in only 4.3% of the studies in this review. This change may be due to a shift among researchers to use more comprehensive symptom instruments. Instruments like the MDASI (13 symptoms) and the Edmonton Symptom Assessment Scale (nine symptoms) are limited because they assess a relatively small number of symptoms using only severity ratings. Given that oncology patients receiving active treatment report an average of 13 unrelieved symptoms, and the optimal symptom dimension to evaluate for symptom clusters has yet to be determined, use of a comprehensive, multidimensional instrument is warranted.

Statistical approaches
EFA was the most common method used in 73.9% of the studies, followed by PCA in 17.4%. These findings are consistent with the previous review that reported that 68.4% of the studies used a factor analytic approach. Given that one conceptual basis for the use of EFA is that symptoms cluster together because they share common underlying mechanism(s), EFA is preferred over HCA or PCA.

One of the key strengths of EFA is that it allows symptoms to load on more than one factor. As a result, the authors of the previous review recommended that the most common symptoms that load on more than one cluster be identified. Of the studies that used EFA, 10 allowed for symptoms to load on multiple factors. While the symptoms that loaded on more than one factor were not specified in most studies, in the two studies that evaluated for symptom clusters in patients with lung cancer, difficulty concentrating, feeling nervous, feeling sad, swelling of the arms and legs, and worrying cross-loaded on multiple clusters. For the four studies that evaluated for clusters in patients with breast cancer, change in the way food tastes cross-loaded in three studies and difficulty concentrating cross-loaded in two.

Symptom dimensions
While severity was the most common dimension used to create the clusters (78.3%), use grew from 43.5% to 69.6%. This growth may be due to the multiple strengths of the MSAS. First, because it evaluates 32 common symptoms, it is cited as one of the most comprehensive instruments to use in research and clinical practice. In addition, the MSAS evaluates multiple dimensions of the symptom experience (ie, occurrence, severity, frequency, distress); has well established validity and reliability; and is available in more than eight languages (eg, Arabic, Chinese, Spanish).

Among the 10 studies that evaluated for clusters using two or more dimensions and named the clusters, psychological and nutritional clusters were the two common clusters identified across all of the studies and dimensions. However, none of the
symptoms within these clusters were consistent across studies. This finding may be partially explained by the variability in cancer diagnoses across the studies. In the previous review,\textsuperscript{11} the authors were unable to compare the number and types of clusters identified across dimensions due to the fact that only 15.8% (n=3) of the studies used two or more dimensions. The growth in the number of studies from 15.8% to 47.8% may be a result of multiple reports recommending that research be done on the stability of symptom clusters across the different dimensions.\textsuperscript{9–11}

**Number and types of symptom clusters**

Across the 23 studies included in this review, the number of clusters identified ranged from two to eight. A psychological cluster was the most common cluster identified in 82.6% of the 23 studies in this review.\textsuperscript{29–40 42–54} Similar to the previous review,\textsuperscript{11} feeling sad or depressed was common across 18 of the 19 studies, while feeling anxious or nervous was common across 16.

Consistent with the previous review,\textsuperscript{11} a gastrointestinal cluster was another common cluster identified in 69.6% of the studies.\textsuperscript{29–36 39–40 42–45 48 50} Nausea was the most common symptom in this cluster that occurred in 13 of the 16 studies, followed by diarrhoea in eight. This finding is similar to the previous review\textsuperscript{11} that identified nausea as one of the most common symptoms across 10 of the 13 studies.

In a departure from the previous review that identified a nutrition or nutritional cluster in only 15.8% of the studies,\textsuperscript{11} a nutritional or weight change cluster was identified across 56.5% of the studies in this review.\textsuperscript{29–32 38–50} Lack of appetite was the most common symptom in 12 of the 13 studies, followed by weight loss in 11.\textsuperscript{29–34 38–50}

The emergence of a nutritional or weight cluster may be due to the inclusion of an increased number of symptoms related to these two problems. For example, in nine of the 13 studies that identified a nutritional or weight change cluster, the MSAS was modified to include additional symptoms (eg, abdominal cramps, increased appetite, weight gain).\textsuperscript{29,32–33 37–47 49–55} Weight gain was common across nine studies,\textsuperscript{29–32 35–37 47–50} and increased appetite was common across six.\textsuperscript{29,33 35 37 47 49}

Additional research is needed to determine the optimal number, as well as the most common and disease and treatment-specific symptoms, to assess in order to obtain more specific and mechanistically based symptom clusters.

In factor analytic methods, factor loading scores are standardised partial regression coefficients that provide an estimate of the strength of the association between a variable (ie, symptom) and a factor (ie, symptom cluster) while controlling for the impact of other factors.\textsuperscript{25} This score is used to determine which symptoms load on which factors using a predetermined cut-off that indicates a meaningful relationship. While factor loadings of ≥0.30 or ≥0.40 are commonly accepted,\textsuperscript{61} it is not clear what the optimal minimum factor loading score should be to include a symptom within a cluster.

In this review, ≥0.40 was the most common minimum factor loading score (n=11),\textsuperscript{33–36 38–40 47–50} followed by ≥0.30 (n=3),\textsuperscript{35 45 46} and ≥0.50 (n=1).\textsuperscript{52} Of note, seven studies did not report this score. In the studies that used a minimum factor loading score of 0.40, two to eight symptom clusters were identified. While no clear pattern emerged in terms of sample size, this wide gap may be due to differences in the instruments used (eg, disease specific vs cancer specific); the type of treatment (eg, adjuvant vs induction chemotherapy); or the timing of the symptom assessments (eg, during chemotherapy, post-chemotherapy). Two of the three studies that used a factor loading of 0.30 identified only two clusters (n=219,\textsuperscript{45} n=219\textsuperscript{46}) and the third identified five (n=232).\textsuperscript{35} This difference may be due to the fact that two of these studies\textsuperscript{35 46} used only 10 symptoms to evaluate for clusters.

**Unique symptom clusters**

While it is important to identify which clusters are consistent across cancer types and treatments, it is equally important to identify clusters that are unique to a specific cancer and/or treatment. A hormonal or vasomotor cluster was identified in 26.1% of the studies.\textsuperscript{29–40 48 50} Of note, four of these studies evaluated for clusters in women with breast cancer\textsuperscript{34 36 48 50} and one in women with a gynaecological cancer.\textsuperscript{35} In the sixth study,\textsuperscript{29} the majority of women had either breast (40.2%) or gynaecological cancer (17.3%).

**Changes in symptom clusters over time**

Ten studies evaluated for changes in clusters over three\textsuperscript{46 47 49 50 52–54} or four timepoints.\textsuperscript{45 48 51} While three studies evaluated for clusters beyond the completion of chemotherapy (eg, six months post-chemotherapy),\textsuperscript{46 48 52} the other seven studies evaluated for clusters around and during active treatment.\textsuperscript{45 47 49–51 53 54} Of these studies, six reported a psychological or emotional cluster that remained relatively stable over time.\textsuperscript{47–52} In contrast, six studies identified a gastrointestinal cluster that varied over time.\textsuperscript{45–48 51–53}

**Methods to evaluate the stability of symptom clusters across dimensions and/or over time**

Stability was evaluated using the method proposed by Kirkova and Walsh\textsuperscript{44} in 81.1% of the studies that evaluated for differences in symptom clusters across two or more dimensions,\textsuperscript{33 35–37 40 47 49 50} and in 60% of the longitudinal studies\textsuperscript{37–50} that evaluated for the stability across dimensions and timepoints. The method proposed by Kirkova and Walsh\textsuperscript{44} specifies that 75% of the symptoms in a cluster should be in agreement in order for a symptom cluster to
be stable across timepoints or dimensions. In addition, the most prominent or important symptom(s) needs to be present. The majority of studies that evaluated for stability of symptom clusters across dimensions or time used Kirkova and Walsh’s method, the criteria were applied with relative subjectivity (eg, described clusters as ‘relatively stable’) (p. 47); described symptoms within clusters as ‘relatively stable’ (p. 787).

This subjectivity may be due in part to a lack of clarity and consensus on the definition of ‘stability’. Similar to Kirkova and Walsh, in their definition of a symptom cluster, Kim and colleagues used stability as a characteristic to describe the group of symptoms within the cluster. In contrast, other researchers have described stability in terms of the type of cluster that is identified. Skerman and colleagues suggested that for a cluster to be stable, it must be ‘reproducible’ (ie, replicated in a similar sample) or appear reliably over time. Barsevick went further to describe stability as how consistently clusters appeared across statistical methods, within homogeneous populations, or over time. From these descriptions, it is unclear if stability refers to the the stability of a specific cluster itself (eg, gastrointestinal, nutritional) across time and/or symptom dimensions, or the symptoms within the cluster. Adding to this confusion, only one of these reports provided criteria to evaluate stability.

Building on Barsevick’s description, we suggest that the term stability should be used to describe whether or not the same clusters are identified across study samples, dimensions and/or over time. While consistency should be used to describe whether the symptoms within a cluster remain the same across these conditions. The use of separate terms to describe these characteristics of symptom clusters may provide clarity and move the science forward. In addition, consensus on how stability is used in the definition of a symptom cluster research warrants consideration.

Secondary outcomes and biomarker evaluation

Of the five studies that evaluated for associations between symptom clusters and other PROs, all used measures of QOL. In addition, Chen and colleagues examined the relationships between symptom clusters and functional performance. Cherwin and Perkhounkova examined how symptom clusters impact symptom interference with daily life and QOL. Of the 23 studies included in this review, only one evaluated for associations between symptom clusters and a biological mechanism.

Limitations

Despite the strict criteria that were employed to ensure a comprehensive review of the literature, only one author made the initial study selection and only two authors did the data extraction. Therefore, it is possible that some studies and/or information were missed. Because the majority of the studies in this review included patients who were homogeneous in terms of gender, race, ethnicity, and cancer diagnosis, our findings may not generalise to all patients with cancer. In addition, because this review focused on adults with stage I to III cancer, our findings may not generalise to patients with advanced cancer or cancer survivors. Finally, 34.8% of the studies came from a single, large study of patients undergoing chemotherapy and may influence the findings of this review.

CONCLUSIONS

This review highlighted numerous areas of growth within symptom clusters research, and identified multiple areas that warrant consideration. One ongoing issue in symptom cluster research is the lack of consistent methods for naming the clusters. In 2016, Miaskowski stressed that a standardised nomenclature needed to be developed in order to facilitate comparisons of clusters across studies. However, as demonstrated in this review, a large amount of variability exists in how clusters were named. For example, the psychological cluster had 10 different names. In addition, researchers must name their clusters to allow for comparisons. In this review, symptom clusters were unnamed in 8.7% of the studies compared with 26.3% in the previous review.

We identified only one study that evaluated for symptom clusters using NA. An advantage of NA is that it allows for an examination of the strengths of the relationships among the symptoms within a cluster and how symptom clusters relate to each other within the network. Additional research using NA is needed to explore the inter-relationships among symptoms within clusters and whether these relationships differ based on the dimension used.

One of the aims of this review was to describe associations between symptom clusters and biological mechanisms. Of the 23 studies included in this review, only one study evaluated for associations between symptom clusters and a biological mechanism. Investigation of the mechanisms that underlie symptoms and symptom clusters is a key priority set by the National Institute of Nursing Research. Future research needs to incorporate the evaluation of biological mechanisms that may underlie symptom clusters in order to better understand why these symptoms cluster and to develop interventions to target clusters of symptoms rather than single symptoms.

Contributors CSH and CAM both contributed to conception and design; acquisition and analysis of the literature and drafted and critically revised the manuscript. KK contributed to conception and design; acquisition of the literature and critically revised the manuscript. YC, AAD and MJH each contributed to the analysis of the literature and critically revised the manuscript. CSH is the guarantor of this work and accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish. All
Systematic review

authors gave final approval and agree to be accountable for all aspects of work ensuring integrity and accuracy.

Funding CSH is supported by a grant from the American Cancer Society (134336-DCSN-20-073-01-SCN) and the National Institute of Nursing Research of the National Institutes of Health (T32NR016920). CAM is an American Cancer Society Clinical Research Professor.

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no data sets generated and/or analysed for this study. Not applicable.

ORCID iD Carolyn S. Harris http://orcid.org/0000-0002-7080-4990

REFERENCES


