Over half of clinical practice guidelines use non-systematic methods to inform recommendations: A methods study

Carole Lunny (✉ carole.lunny@ti.ubc.ca)
University of British Columbia   https://orcid.org/0000-0002-7825-6765

Cynthia Ramasubbu
University of British Columbia

Lorri Puil
University of British Columbia

Savannah Gerrish
University of British Columbia

Tracy Liu
University of British Columbia

Douglas M. Salzwedel
University of British Columbia

Barbara Mintzes
Charles Perkins Centre and School of Pharmacy, The University of Sydney

James M. Wright
University of British Columbia

Method Article

Abstract

Introduction Assessing the process used to synthesise the evidence in clinical practice guidelines (CPGs) enables users to determine the trustworthiness of the recommendations. We aimed to assess whether systematic methods were used when synthesizing the evidence for CPGs; and whether reviews or 'overviews of reviews' were cited in support of recommendations.

Methods and analysis We followed a study protocol. CPGs published in 2017 and 2018 were retrieved from TRIP and Epistemonikos. We randomly sorted and sequentially screened the CPGs to select the first 50 that met our inclusion criteria. Our primary outcomes were the numbers and proportions of recommendations that were based on reviews and 'overviews', and CPGs using either a systematic or non-systematic process to gather, assess, and synthesise evidence. We also looked for evidence that critical appraisal was conducted. We also performed a chi-square test of independence to examine the relationship between variables.

Results Of the 50 guidelines, 34% did an exceptional job in systematically synthesising the evidence to inform recommendations. These guidelines clearly reported their objectives and eligibility criteria, conducted comprehensive search strategies, and assessed the quality of the studies. 66% of CPGs reported non-systematic methods to develop their recommendations. This percentage is likely an underestimation because we excluded some CPGs when selecting studies. Overall, 90% of CPGs cited reviews to inform recommendations, and one fifth cited a Cochrane systematic review. Of the 29 CPGs that included reviews, 21% critically appraised the review. 60% of CPGs assessed the quality of primary studies.

Conclusions We used novel methodology to evaluate recommendations in a random sample of CPGs, and found that 62% did not use a systematic process to gather, appraise, and synthesise the evidence. Significant improvement is needed in the conduct and reporting of CPG methods. Guideline developers should use systematic methods endorsed by reputable evidence synthesis organisations.

1. Background

Clinical practice guidelines (CPGs) help healthcare practitioners navigate the complexities in patient care, and facilitate informed, shared decision-making. There is, however, ongoing debate about the extent to which guidelines provide a systematic, evidence-based approach to healthcare [1, 2]. The definitions of, and the processes by which, guidelines are developed vary substantially [6]. Although standards for CPG development are published [3-5], consensus about the optimal methods for their development is lacking [3]. Consequently, the guideline development process can suffer from bias, conflicts of interest, a lack of rigour, and limited applicability in practice.

Generally, the development of CPG recommendations follows a series of steps [7], starting with the convening of a working group, conflict of interest management, and specification of the clinical questions and relevant outcomes. Research questions help define literature searches, inform the planning and process of the evidence synthesis, and act as a guide for the development of recommendations. Evidence synthesis is an integral part of guideline development and typically should involve the following steps: specification of the purpose, objectives and scope of the review; specification of eligibility criteria and literature search methods; data extraction; assessment of risk of bias of included studies, synthesis of findings, and assessment of the quality or certainty of the evidence for each outcome. After the synthesis steps are completed, the working group translates the evidence into recommendations. The higher the certainty of a body of evidence, the *more* likely a strong *recommendation* can be made. However, recommendations incorporate additional considerations such as the net balance of benefits and harms, values and preferences, resource use and acceptability [7].

Guideline developers may use some or all of these steps, and various methods in the conduct of the evidence synthesis. They may conduct a non-systematic literature review (i.e. no systematic methods used), or non-systematic 'overviews of reviews' (also known as umbrella reviews, reviews of reviews or meta-reviews; that retrieve and synthesize the results of systematic reviews), or one of several types of systematic evidence syntheses. The latter include systematic reviews that synthesize the results of original primary studies (e.g. randomized controlled trials, cohort studies), and systematic 'overviews of reviews'. Moreover, guideline developers may search for, and include, a variety, and combination, of different study designs.

Many international standards (i.e. instruction manuals) exist for guideline developers when conducting CPGs, including guidance from the Institute of Medicine (IOM) [8], Guidelines International Network (GIN) [3], the Scottish Intercollegiate Guidelines Network (SIGN) [9], the National Institute for Health and Care Excellence (NICE) [4], the Australian National Health and Medical Research Council (NHMRC) [5], and the World Health Organization (WHO) [10]. The GRADE Working Group provides one of the most rigorous approaches, a framework for assessing the certainty of a body of evidence in an evidence synthesis, then interpreting the evidence into recommendations, and judging the strength of the recommendations [11, 12]. **Box 1** summarizes the international standards on how to conduct the evidence synthesis process for CPGs by the Institute of Medicine (IOM) [8], AGREE II and Guidelines International Network (GIN) [3].

Despite existing international standards, systematic surveys of guidelines [13-16] indicate many are of moderate to low quality, as assessed by the Appraisal of Guidelines for Research & Evaluation Instrument (*AGREE II*) tool [18]. AGREE II is the most commonly applied methodological quality CPG tool worldwide [15]. The tool's third domain deals with the methodological and/or reporting quality of the evidence

synthesis process in a CPG. The most recent systematic survey of 421 CPGs found that 33% scored low on this domain for "rigor of development" [16]. Although popular, the AGREE II tool fails to provide a comprehensive and thorough evaluation of the methodological rigor of evidence syntheses within CPGs.

Box 1: Guidance on the evidence synthesis process in guideline development from IOM, AGREE II, and GIN

IOM	AGREE II	GIN
"Statements that include **recommendations**, intended to optimize patient care, that are informed by **a systematic review of evidence** and an assessment of the benefits and harms of alternative care options."	Domain 3: Rigor of Development 7. **Systematic methods were used** to search for evidence. 8. The criteria for selecting the evidence are clearly described. 9. The strengths and limitations of the body of evidence are clearly described. 11. The health benefits, side effects and risks have been considered in formulating the recommendations. 12. There is an **explicit link between the recommendations and the supporting evidence**. 13. The guideline has been externally reviewed by experts prior to its publication. 14. A procedure for updating the guideline is provided.	*6. Evidence Reviews* Guideline developers **should use systematic evidence review methods** to identify and evaluate evidence related to the guideline topic.

If we are to improve clinical practice, the evidence underpinning guideline recommendations must be rigorously developed and evaluated [17, 18]. Non-systematic methodology to gather, appraise, and synthesise evidence may lead to biased results and over- or under-estimation of treatment effect estimates, which are especially harmful when used to support CPG recommendations [19]. For example, in 2016, the Canadian Association of Radiologists (CAR) issued a guideline calling for women with average breast cancer risk to begin screening mammography at age 40 [20]. The results from three RCTs from 2010, 2014, and 2015 shows that the risk of cancer is lower for women ages 40 to 44 and the risk of harm from screening (biopsies for false-positive findings, overdiagnosis) is higher compared to women over 50 [21]. These three RCTs could have been included in the CAR guideline but were not. No methods for how the guideline was developed were found in the guideline itself, nor on the association's webpages. In this case, use of non-systematic methodology led to a potentially harmful CPG recommendation.

Assessing the process used to synthesise the evidence underpinning recommendations in CPGs enables knowledge users to determine the trustworthiness of the recommendations [22]. We therefore aimed to (a) assess whether systematic methods were used when synthesizing the evidence for CPGs; and (b) evaluate the type of systematic review (with or without pairwise and network meta-analysis) or overview cited in support of recommendations.

2. Methods

2.1   Design and protocol

Our methods protocol was previously published in BMJ Open [23], and in the Open Science Framework (https://osf.io/rju4f/). The study follows systematic review methods guidance for searching, study selection, data extraction, and critical appraisal. As this is a methods study, no relevant research reporting checklists exist. Formal ethical approval was not required as primary data were not collected. Our raw data files have been uploaded to the repository Open Science Framework (https://osf.io/8rxnp/ with DOI 10.17605/OSF.IO/8RXNP).

2.2   Search

CPGs were retrieved from the Turning Research Into Practice (TRIP) and Epistemonikos databases over a two-year period (January 1, 2017 to December 31, 2018). This time period was selected for current relevance and to limit the number of CPGs screened. In Epistemonikos, we selected the "Broad syntheses" filter for retrieval of CPGs (**Appendix 1**). Epistemonikos includes citations retrieved from the following databases: Cochrane Database of Systematic Reviews; PubMed; Embase; CINAHL (The Cumulative Index to Nursing and Allied Health Literature); PsycINFO; LILACS (Literatura Latinoamericana y del Caribe en

Ciencias de la Salud); DARE (Database of Abstracts of Reviews of Effects); the Campbell Library; the JBI Database of Systematic Reviews and Implementation Reports; and the EPPI-Centre Evidence Library. As the TRIP database only contains CPGs, we downloaded all records without restricting by study type. TRIP retrieves guidelines from over 289 journal publications and has recently migrated all content from AHRQ's Clinical Guidelines Clearinghouse (www.guidelines.gov), which lost funding on July 16, 2018 (Jon Brassey, personal communication, April 10, 2018).

2.3. Study selection

References retrieved from TRIP and Epistemonikos were imported into a single EndNote file and de-duplicated. Subsequently, the retrieved citations were randomly sorted using Microsoft Excel's RAND function and, screened at the full text level, using a form designed in Microsoft Excel (2013).

Screening to identify citations meeting our inclusion criteria was conducted independently by two authors, starting with the lowest random number, until 50 guidelines were included. This sample size was chosen to be large enough to include a variety of clinical conditions. All authors involved in study selection screened ten studies as a calibration exercise to establish agreement in definitions of eligibility criteria. Discrepant decisions were resolved by discussion with a senior author.

2.4   Eligibility criteria

In addition to a requirement for publication between January 1, 2017 and December 31, 2018, we defined clinical practice guidelines according to the following inclusion criteria:

- Pertain to the management or treatment of any clinical condition. CPG recommendations for management may include, for example, recommendations for lifestyle modifications, when to implement or adjust therapy, choice of therapy including treatment combinations, and ways to prevent harms associated with therapy.
- Produced by a group or organisation (i.e. not authored by one person).
- Contain at least two explicit recommendations for treatment or management of a condition.
- Contain a description of their methodology within the guideline or in supporting documents (e.g. inclusion/exclusion criteria, key terms used to search, number of databases searched, number of authors used to select studies, methods used to create recommendations, or quality/risk of bias assessment).
- Contain a reference list (i.e. a bibliography).

If more than one publication from the same organisation or author group was identified, we included the most recent version of the guideline.

Exclusion criteria:

CPGs without recommendations or that focus solely on screening or diagnosis were excluded. CPGs were also excluded for the following reasons:

- The full text is unavailable.
- It is designed for local use only (e.g. in a single health facility or single regional health service).
- It is designed for use restricted to hospitalized patients or patients in long-term care facilities.
- It aims to provide recommendations for patterns of use of medications (e.g. guidance about adherence to medications) but not treatment choice.

The eligibility criteria were piloted by all data extractors (CL, DS, BM, CR, TL, SG) independently on a sample of ten guidelines retrieved from the search to ensure consistent application. Once the guidelines were screened and included, we attempted to retrieve any supplementary files, methods documents, published systematic reviews, or any other documentation supplementary to the guideline.

2.5 Definitions of evidence syntheses

We classified approaches to evidence synthesis according to the following definitions.

Literature reviews or non-systematic reviews are summaries of the literature on a particular topic that are not developed systematically. A literature review may identify a search strategy but does not fulfill the minimum requirements of a systematic process as listed below.

Systematic review. A review of evidence is considered systematic if it reports, at a minimum:

- A research question or questions formatted using PICOS (participants, interventions, comparisons, outcomes, and study design);
- Eligibility criteria for all included study types;
- Full search strategy for at least one database (i.e. keywords reported and a full search strategy reported in an appendix);
- Search strategy in the main body of the manuscript (i.e. not only in the abstract) using 2 or more electronic databases; and
- Process for selecting/screening studies (e.g. number of authors; independent process).

Systematic or non-systematic reviews may contain a *pairwise meta-analysis* or *network meta-analysis*. A pairwise meta-analysis is a traditional meta-analysis in which the effect estimates of two interventions are compared directly, following a judgment that the included studies are sufficiently similar to warrant pooling. *A network meta-analysis* compares multiple interventions using both direct comparisons of interventions within randomised trials and indirect comparisons across trials based on a common comparator [9].

An overview of reviews aims to primarily identify, include, and synthesise the results of secondary analyses (reviews, systematic reviews, guidelines, or health technology assessments) and may or may not have used systematic methods as outlined above [10-12].

2.6 Outcomes

The study's primary outcome consisted of the numbers and proportions of recommendations within CPGs that were based on the following types of evidence syntheses:

1. Systematic reviews without meta-analysis

2. Systematic reviews with pairwise meta-analysis

3. Systematic reviews with network meta-analyses

4. Overviews of systematic reviews

We also evaluated the number of CPGs using either a systematic or non-systematic process to gather, assess, and synthesise evidence (**Figure 1**; and section 2.10).

The secondary outcomes, calculated as numbers or proportions, are:

5. CPGs that cited a Cochrane review or overview

6. CPGs that report using GRADE methodology

7. CPGs that report using other systems evaluating the strength of the recommendation and type of tool used (e.g. American Heart Association [24])

8. CPGs that report using a level of evidence system and type of system used

9. Currency of the CPG (calculated by the time from last search to full publication)

10. CPGs that report conflicts of interest disclosures by authors

2.7   Data extraction

Data from 50 guidelines were extracted. Each included CPG was first examined to determine whether reviews, or overviews of reviews, were used and cited in support of one or more of the guideline's recommendations (yes or no for each review type). If yes, we evaluated all treatment or management recommendations citing each review type.

A data extraction form was developed in Microsoft Excel (2013). Ten CPGs were independently extracted by six authors and then discussed to come to consensus about definitions and to calibrate the coding (**Appendix 2**). Full data extraction was completed independently by two authors and compared. Any discrepancies were discussed, and a senior author arbitrated conflicts. A senior author checked that the data was consistently coded across similar or related items.

Data extracted at the guideline level included: our primary and secondary outcomes, name of the guideline, year of publication, country, the organisations or commissioning agency (publisher), type of publisher (government, medical society, university, other [specify]), aim of the guideline, publishing journal (if applicable), open source/paywall, the date of the last search for evidence to be included in the guideline, funding, declaration of conflicts of interest by developers, stakeholder affiliation with/honoraria from pharmaceutical companies, target population (general population, or specific subpopulations such as those identified by age (e.g. children and adolescents, adults of any age, older adults), sex/gender, or co-morbidities), and scope (pharmacological or non-pharmacological treatment (e.g. surgical, medical device).

If a review or overview was cited within a recommendation, we also looked for evidence that critical appraisal of the review or overview was conducted, and recorded which tool was used (e.g. Assessing the Methodological Quality of Systematic Reviews (AMSTAR) 1 or 2 [25], Risk of Bias Assessment Tool for Systematic Reviews (ROBIS) [26]).

2.8   Gaps in evidence supporting a recommendation

If a guideline did not cite a Cochrane systematic review, we searched the Cochrane Database of Systematic Reviews using the keywords used in the main search strategy of the guideline. We noted whether a Cochrane systematic review or overview could have been identified and used to inform the recommendations by checking the search dates of the CPG to ensure it could have been included.

2.10 Risk of bias assessment of the review process for informing the guideline recommendations

To determine if a systematic process was used to gather, assess and synthesise evidence to inform recommendations, we used the following four criteria:

1. Explicit statement of the questions or objectives reported in terms of PICOS (Populations, Interventions, Comparisons, Outcomes, and Study design) elements.
2. Eligibility criteria reported for all included study designs.
3. A systematic search conducted (i.e. two or more databases searched).
4. Process reported for selecting/screening studies (e.g. number of authors, independent process).

We considered these criteria to be the minimum that can be used by a CPG to reduce bias and limitations when gathering evidence to inform recommendations. We also assessed whether the CPG working group reported the following methods:

- A full search strategy for at least one database (i.e. keywords reported and a full search strategy reported in an appendix).
- Assessment of the quality/risk of bias of the review or overview supporting/refuting the recommendation.
- Assessment of primary studies for quality/risk of bias.

We adapted these risk of bias items from the ROBIS tool, which comprehensively assesses the risk of bias of a systematic review [26]. The tool includes items relating to internal validity and classifies them in the following domains: study eligibility criteria; identification and selection of studies; data collection and study appraisal; and synthesis and findings.

2.10 Open access

All data management and study processes were conducted and recorded in the Open Science Framework.

2.11 Data analysis

We calculated the number and frequencies of citations of reviews and overviews, and their characteristics, found in the 50 included guidelines. We described and tabulated all primary and secondary outcomes. Additional information was put into appendices. To estimate the time taken to

conduct each guideline, we calculated the difference between the initial literature search date and publication date using the month and day function in Excel 2013.

As sufficient studies were collected to make meaningful comparisons (≥10), we performed a chi-square test of independence to examine the relation between using the GRADE approach and whether the guideline used a systematic process (Model 3; post hoc analysis). Dependent categorical variables were type of organization (medical association, pharmaceutical, government, no funding, not reported), scope (narrow, broad), and continent (Europe, North America, Intercontinental). We also performed a chi-square test of independence to examine the relationship between GRADE use and type of funder, and CPG having conducted a systematic process and type of funder. We planned to explore whether the characteristics of CPGs differed in terms of pharmacological vs. non-pharmacological scope. However, there were too few CPGs with these characteristics to permit reliable comparisons (≤10 in each group). We formally tested the associations using a chi-square test for one independent variable with 2 levels with categorical dependent variables in R.

3. Results

3.1   Search results

From a total of 713 records retrieved from the TRIP and Epistemonikos databases, 691 remained after duplicate removal (**Figure 1** flowchart). The 691 records were then randomly sorted and screened sequentially. A total of 419 records were screened at full text to obtain our target of 50 eligible CPGs (see included studies in **Appendix 3**). Of the 369 excluded records, 47 CPGs were excluded as they did not have a methods section, and 16 did not include a reference section.
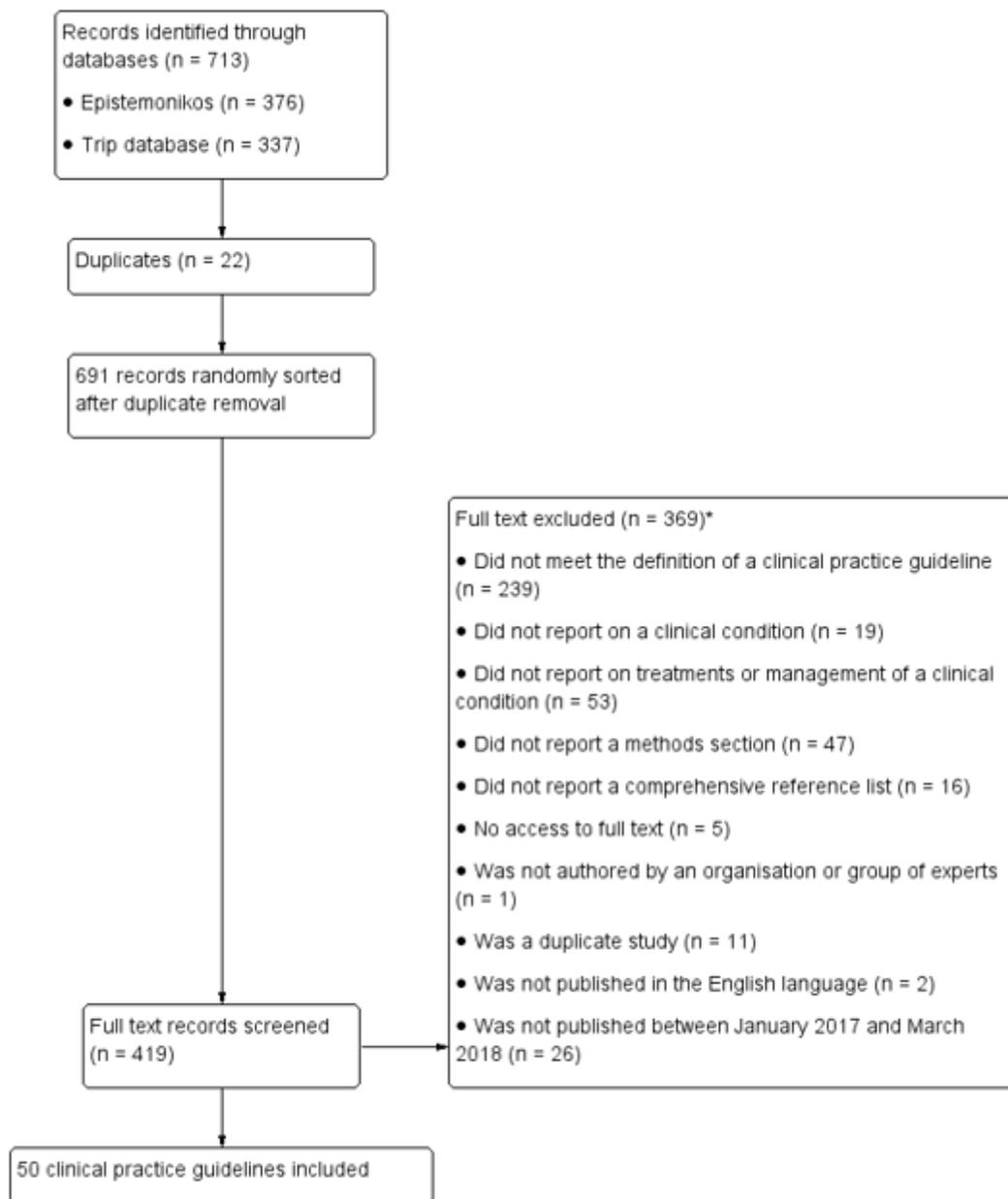
Figure 1. Flowchart of the study selection process

3.2   Characteristics of CPGs

The majority of the randomly selected CPGs were from Canada or the United States (31/50 [62%]) and were published in 2017 (40/50 [80%]); **Table 1**). CPGs conducted in Europe constituted 32% (16/50). The most frequent medical condition addressed was malignant neoplasms (11/50 [22%]); however, the CPGs covered a broad range of clinical topics. Half of the CPGs (25/50 [50%]) had both a pharmacological and non-pharmacological scope. The average time from search to full publication was 24 months (range 2 – 204 months).

The three associations most frequently commissioning or conducting guidelines were the European Society for Medical Oncology (5/50 [10%]), the American Society of Clinical Oncology (4/50 [8%]), and the American Urological Association (3/50 [6%]). The majority of CPGs were funded by a medical society (18/50 [36%]), the pharmaceutical industry (9/50 [18%]), or funding was not reported (9/50 [18%]). A smaller number of CPGs were funded by government (8/50 [16%]), or did not receive any funding (6/50 [12%]).

The majority of CPGs were published in peer reviewed journals (42/50 [96%]), and were open access (45/50 [90%]). In 48 CPGs (96%), guideline authors declared their conflicts of interest, and in 33 (66%), authors declared affiliations with pharmaceutical companies (66%).

Table 1: Characteristics of clinical practice guidelines (n = 50)

Characteristics	clinical practice guidelines n (%)
Year of publication	
2017	40/50 (80)
2018	10 (20)
Region*	
Canada or United States	31 (62)
Europe	16 (32)
All other regions	3 (6)
Medical classification	
Malignant neoplasms	11 (22)
Diseases of the cardiovascular system	5 (10)
Diseases of the genitourinary system	4 (8)
Endocrine, nutritional and metabolic diseases	4 (8)
Diseases of the gastrointestinal system	3 (6)
Mental and behavioural disorders	3 (6)
Diseases of the skin and subcutaneous tissue (n=2);	
Factors influencing health status and contact with health services (n=2);	
HIV/AIDS (n=2);	
Symptoms, signs and abnormal clinical and laboratory findings (n=2).	
Other classifications**	23 (46)
Scope	
Both pharmacological and non-pharmacological scope	25 (50)
Pharmacological scope	13 (26)
Non-pharmacological scope (e.g. surgery, medical device)	7 (14)
Funding of the association writing the guideline	
Medical society	18 (36)

Pharmaceutical industry	9 (18)
Government	8 (16)
No funding received	6 (12)
Not reported	9 (18)
Publishing	
Published in a peer reviewed journal	42 (84)
Open source publishing (open access)	45 (90)
Conflict of interest	
Declaration of conflicts of interest by developers	48 (96)
≥1 stakeholder affiliated with*** pharmaceutical companies	33 (66)
Average time from search to full publication (months [range])	24 (2 − 204)

*Numbers do not total 50 as multiple categories can apply to one CPG.

**Ten additional topics were identified: complications of surgical and medical care (n=1); congenital malformations (n=1); diseases of oral cavity, salivary glands and jaws (n=1); diseases of the blood (n=1); diseases of the ear (n=1); diseases of the eye (n=1); diseases of the nervous system (n=1); injuries to the abdomen, lower back, lumbar spine and pelvis (n=1); osteopathies and chondropathies (n=1); pregnancy and childbirth (n=1).

*** Includes both employees of the pharmaceutical companies and consultant fees.

3.3   Approach to evidence synthesis in CPGs

According to our definition of a systematic process, as outlined in the methods, 17/50 (34%) of CPGs were systematic in their approach to evidence syntheses, and over half (33/50 [66%]) of CPGs were non-systematic (**Figure 2**). Of the 3/50 (6%) CPGs that used an overview method (synthesised systematic reviews, CPGs, or health technology assessments (HTAs)), one CPG was systematic in its approach to evidence synthesis, and two were non-systematic (**Table 2**).

Of the 17 CPGs using a systematic approach to evidence syntheses, 65% (11/17) used a qualitative synthesis approach, 18% used a pairwise meta-analysis (3/17), 12% (2/17) used a network meta-analysis approach, and one overview of systematic reviews used a qualitative synthesis approach (1/17)

(**Table 2**). Of the three CPGs that conducted overviews, two (2/50 [4%])) did an overview of systematic reviews, and one (1/50 [2%])) did an overview of CPGs.

Table 2: Clinical practice guidelines using a systematic or non-systematic approach to evidence syntheses (n=50)

Review types	All CPGs n (%*)
Reviews:	**47 (94%)**
Literature reviews	**31 (62%)**
with qualitative synthesis	22 (44%)
with pairwise meta-analysis	7 (14%)
with network meta-analyses	2 (4%)
Systematic reviews	**16 (32%)**
with qualitative synthesis	11 (22%)
with pairwise meta-analysis	3 (6%)
with network meta-analyses	2 (4%)
Overviews:	**3 (6%)**
Non-systematic overviews of reviews	**2 (4%)**
Overview of reviews with non-statistical summary	1 (2%)
Overview of guidelines with non-statistical summary	1 (2%)
Systematic overview of reviews	**1 (2%)**
with non-statistical summary	1 (2%)

*Percentages do not sum to 100 as some CPGs used more than one approach

3.4   Specific methods used to formulate guideline recommendations

More than half of the guidelines (30/50 [60%]) used an explicit statement to develop guideline questions and/or objectives and structured these using the PICOS format (**Table 3**). Eligibility criteria were specified clearly in over half of the CPGs (29/50 [58%]), and a systematic search reporting keywords used was found in half of the CPGs (25/50 [50%]). A total of 31 CPGs searched two or more databases (31/50 [62%]), and a slightly lower number reported the process for selecting studies (27/50 [54%]).

Approximately one third of CPGs (17/50 [34%]) used the GRADE approach, and two thirds (33/50 [66%]) used another system to assess the strength of the evidence, such as the Guidelines Into Decision Support methodology [27] (**Table 3**).

Over half of CPGs (30/50 [60%]) reported assessing the quality of primary studies <u>or</u> reported using the GRADE approach (and therefore must have assessed the risk of bias of the primary studies although the assessments weren't provided) (**Table 3**). Of these 30 CPGs, eight CPGs (8/30 [27%]) reported using the Cochrane risk of bias tool to assess the quality of randomised trials, 11 CPGs reported using another tool (11/30 [37%]) such as the Drug Effectiveness Review Project instrument [28], QUADAS 2 tool [29], or the Newcastle Ottawa scale [30], and 11 did not report the specific tool. Of the 29 CPGs that had defined eligibility criteria for inclusion of reviews, only 6 (6/29 [21%]) assessed the risk of bias or quality using an appropriate tool like ROBIS [26] or AMSTAR 2 [31].

Table 3: Specific methods used for evidence synthesis in CPGs

Method	All CPGs n/N (%)
Formulation of the guideline question(s) or objective(s) in terms of PICOS elements	30/50 (60)
Inclusion and exclusion criteria of studies reported (e.g. all study types; RCTs; reviews; overviews)	29/50 (58)
Systematic search strategy	25/50 (50)
Two or more databases searched	31/50 (61)
Process reported for selecting/screening studies	27/50 (54)
Assessment of methodological quality (risk of bias) of review/overview	5/50 (10)
Assessment of methodological quality (risk of bias) of primary studies	30/50 (60)
Cochrane risk of bias tool for randomized trials	8/30 (27)
Other tool	11/30 (37)
GRADE used but risk of bias methods not reported	11/30 (37)
GRADE approach reported	17/50 (34)
Other system for assessing the strength of recommendation reported	33/50 (66)
Highest "level of evidence" rating	
High quality SR/MAs	17/50 (34)
SR/MAs	9/50 (18)
High quality RCTs	13/50 (26)
RCTs	4/50 (8)

3.5   *Assessment of whether reviews or overviews were used to inform recommendations*

Of the 50 CPGs, 46/50 (92%) cited reviews to inform recommendations. There were a total of 128 recommendations citing 249 reviews of any type. Of the cited reviews, 160/249 (64%) were systematic

reviews with pairwise meta-analysis, 7/249 (3%) were systematic reviews with network meta-analysis, and 23/249 were systematic reviews without meta-analysis (**Table 4**). Of the 190 systematic reviews to inform recommendations, 46/190 (23%) of these were Cochrane systematic reviews, representing 46/249 (17%) of all reviews.

Of the 45/50 (90%) CPGs that cite either a review or overview to inform recommendations, only 29/50 (58%) guideline developers specified in their eligibility criteria that reviews or overviews were included.

Table 4: Types of reviews used in 128 recommendations

Types of reviews used in the 128 recommendations	Reviews (n = 249) n (%)
Literature reviews:	*59 (24%)*
Literature review without meta-analysis	30 (12%)
Literature review with pairwise meta-analysis	29 (12%)
Systematic reviews:	*190 (76%)*
Systematic review* without meta-analysis	23 (9%)
Systematic review with pairwise meta-analysis	160 (64%)
Systematic review with network meta-analyses	7 (3%)

3.4   *Gaps in evidence supporting a recommendation*

Of the 50 guidelines, 16/50 (32%) cited a Cochrane systematic review or an 'overview of systematic reviews'. Of the 34 remaining CPGs, 27/34 (79%) CPGs could have used and cited Cochrane reviews to inform the recommendations. The median number of Cochrane reviews that could have been cited based on the CPGs search strategy was 1 [range 0-18].

3.5   *Potential associations*

A chi-square test of independence was performed to examine the association between use of the GRADE framework (pre-specified analysis) and having used a systematic process for evidence synthesis (post

hoc analysis). The association between these variables was insignificant, X^2 (1, N = 50) = 0.023, p = 0.9. CPGs that used the GRADE framework were not more likely to have used a systematic process.

We also explored whether a link existed between GRADE use and type of funder (X^2 [4, N = 50] = 9.05, p = 0.05), conflict of interest (X^2 [1, N = 50] = 0.18, p = 0.07), scope (X^2 [1, N = 50] =1.4 , p = 0.2), affiliation with the pharmaceutical industry (X^2 [1, N = 50] =2.4 , p = 0.11), and continent (X^2 [2, N = 50] =6 , p = 0.05). Upon further exploration, CPGs using GRADE were more likely to have been funded by government or the pharmaceutical industry, and conducted internationally (with an organisation like the WHO).

We also tested the relationship between the CPG having conducted a systematic process and type of funder (X^2 [4, N = 50] = 3.60, p = .46), conflict of interest (X^2 [1, N = 50] = 0.9804, p = 0.322), scope (X^2 [1, N = 50] 0 , p = 1.0), affiliation with the pharmaceutical industry (X^2 [1, N = 50] =0.61, p = .43), and continent (X^2 [2, N = 50] =7.55 , p = .02). CPGs that used a systematic process were more likely to have been conducted internationally (with organisation like the WHO). A table of these associations is found in **Appendix 4**.

4. Discussion

In this sample, only a minority of CPGs did an exceptional job in systematically synthesising the evidence to inform recommendations, notably the guidelines by the WHO [32], the UK National Institute for Health and Care Excellence (NICE) [33, 34] and the Thoracic Society of Australia and New Zealand [35]. These guidelines explicitly and clearly reported their objectives and eligibility criteria, conducted comprehensive search strategies, and assessed the methodological quality of the studies included in the review of the evidence. High quality, systematically developed evidence syntheses produced by CPG working groups provide the best available evidence to inform recommendations [36].

Nearly two thirds, or 62%, of CPGs reported non-systematic methods to develop their recommendations. This percentage is likely an underestimation because we excluded some CPGs when selecting studies. A total of 47 CPGs (47/691 [7%]) were excluded because they did not contain a methods section, and 16 were excluded (16/691 [2%]) because they did not cite any references. This is a small improvement from an assessment of guidelines done two decades ago [37], which stated that only 20% of the 269 included CPGs specified search methods, and 25% did not cite any references.

Several possible explanations exist for why guideline developers may use non-systematic methods when reviewing the evidence to inform recommendations. First, guideline working groups may lack the required resources to undertake a full systematic process, which is time consuming and labour intensive. Second,

guideline developers may be unaware of the importance of using systematic methods to minimise bias and error when synthesising evidence, and/or be unaware of the guidance available from evidence synthesis organisations, such as Cochrane [32], the GRADE Working Group [38], and other organisations [8, 33]. Some organisations and societies require that CPG developers adhere to established methodological standards (e.g. NICE [39]) and undergo mandatory training in these methodologies (e.g. **Infectious Diseases Society of America [40])**. However, guidance provided by other medical associations and societies on how to gather, appraise and synthesise evidence to inform recommendations can vary and may not adequately emphasize the steps needed to minimise biases [6].

Second, it is possible that the opinion of guideline developers may outweigh or ignore relevant evidence in guideline development. Indeed, prior evaluations of clinical guidelines in a range of clinical specialty areas have found that many recommendations are based on expert opinion [41-44]. Many groups may vote on recommendations based on scientific evidence, but others may base voting on opinions drawn from consensus or clinical experience. Expert opinion may appropriately be combined with empirical evidence in clinical practice. In the absence of relevant research, expert opinion may be considered the best available evidence. The process used for making recommendations, including the role of expert opinion in decision making should be transparently reported [45]. Caution is advised in situations when a non-systematic process for synthesising evidence to inform recommendations is used, as relevant literature may not be found, and research can be selected to confirm expert opinion. This potentially allows for self-serving biases, such as confirmation bias (selective gathering of or ignoring evidence), consensus bias (believing that one's opinions are relatively common and appropriate), and bias associated with conflicts of interest [46].

Our investigation confirms previous findings that investigators often fail to cite and use earlier research when preparing new research [47-50]. Many guideline developers, such as the WHO and the NHMRC, recommend the use of systematic reviews and overviews to underpin guideline recommendations [5, 51]. The majority of CPGs in our sample used any type of review to underpin recommendations, but only about a third cited Cochrane reviews. About 80% of CPGs could have cited a Cochrane review but did not. Our findings are similar to other studies that identified 40-70% of guideline recommendations did not use all the relevant Cochrane reviews [52-54]. Recommendations not underpinned by review-level evidence may indicate problems with the CPG methodology (e.g. search strategy [missing relevant systematic reviews]) or gaps in the evidence base (i.e. a lack of adequately-designed relevant studies).

Overviews are potentially useful to guideline developers [24,48] who use secondary data (i.e. SRs, guidelines, and HTAs), who need to make decisions about the most effective and safe interventions for patient centred care in the development of CPGs [5]. As 94% of CPGs in our study used reviews to

underpin recommendations, guideline authors can benefit from using overviews of review methods to resolve discrepant, conflicting, and overlapping data in multiple systematic reviews on the same topic [55]. A catalogue of methods for overviews has been systematically developed, and these methods can be used by guideline developers to synthesise the results of multiple systematic reviews across similar topics [56-58]. If CPGs are to use overview of reviews in development of recommendations, the guidance relating to their conduct and reporting (e.g. [56-58]) must be used, and all stakeholders involved in their conduct [59].

Evaluating how well a study has been conducted is essential to determine if the findings are trustworthy and relevant to patient care and outcomes. Two thirds of guideline developers in our sample of CPGs did not assess the risk of bias (quality) of included primary studies, and one fifth did not assess the methodological quality of included reviews. Risk of bias assessment is about identifying systematic flaws, bias or limitations in the design, conduct, or analysis of research. If an overview or NMA is at risk of bias and inappropriate methods are used, the findings can be misleading. Several studies have shown that bias can obscure up to 60% of the real effect of a treatment [60, 61]. Evidence shows biased results from poorly designed and reported studies can mislead decision-making in healthcare at all levels [62-64].

Significant improvement is needed in the reporting of methods in CPGs. Improved reporting will help users of CPGs assess the methodological quality of the included evidence to inform recommendations. While the length of CPGs is often unwieldy, links in text to the full methodology should be provided. As with clinical trials and systematic reviews, all CPGs should conduct a priori protocols, plans, or registered reports [65]. Without these details, researchers will not be able to assess biases such as selective reporting of outcomes, or selective handling of multiple measures or analyses. At a minimum, guideline developers should develop explicit research questions, define all outcomes within the domains of interest, and pre-specify plans for handling many different outcomes, measures, and analyses. Prospective registration of CPGs would promote transparency, help reduce potential for bias, and would serve to avoid unintended duplication of multiple CPGs on the same topic [66-68].

Implications for clinical care

Systematically developed evidence syntheses in CPGs provide the high quality evidence base that is needed to inform recommendations. Using non-systematic methods compromises the validity and reliability of the evidence used to inform guideline recommendations, leading to potentially misleading and untrustworthy results. Patients, healthcare providers and policy makers need in turn the highest quality CPGs to inform decisions about which treatments should be used in healthcare practice. While many therapeutic decisions must be made on the basis of weak guideline recommendations, these

recommendations must still be based on systematic methodology for the CPG to be trusted. The consequences of providing patient care and rolling out population health policy from results of CPGs that are based on non-systematic is compromised patient care and safety [69].

Strengths and Limitations

The strengths of our methods include the adoption of systematic and transparent methods, specific and explicit eligibility criteria, broad search strategies, randomised selection of CPGs, and duplicate and independent processes for CPG selection and data extraction.

A main limitation of our study is the narrow search dates of the test set of CPGs. In addition, when coding guidelines, substantial judgment was required. To mitigate the subjectivity of classifying and coding characteristics and methods used in reporting CPG recommendations, all data extractors piloted the items on ten studies. The piloting results were discussed to refine the wording of the items, come to consensus about definitions, and calibrate the coding. We also used minimum criteria to assess the evidence synthesis process used by the CPGs. A full assessment of the biases in an evidence synthesis would involve using a methodological quality assessment tool for reviews like ROBIS [26] or AMSTAR 2 [31]. Reporting was inadequate across CPGs, thus limiting our assessment of the methodological quality of the evidence gathering, appraising, and synthesising process. Complete reporting of methods will help knowledge users to assess the risk of bias of the evidence underpinning recommendations.

A further limitation is that our study is focused on CPGs for the management or treatment of any clinical condition. We may also have missed high quality reviews by not searching for 'non-Cochrane' reviews. Cochrane reviews are known for using robust methodology [70-72], and by searching for missed Cochrane evidence, we evaluated whether a guideline might be missing high quality evidence. However, Cochrane reviews are prone to biases like non-Cochrane reviews, and should not be considered high quality without assessment of the risks of bias.

Future studies looking into the use of reviews in screening or diagnostic recommendations would also be useful to determine the quality of recommendations.

Conclusions

We used novel methodology to evaluate recommendations for clinical treatment in a random sample of recent CPGs, and found that 62% of the guidelines did not use a systematic process to gather, appraise, and synthesise the evidence to inform recommendations. Significant improvement is needed in the conduct as well as the reporting of methods in CPGs. Guideline developers should use the systematic

methods endorsed by reputable evidence synthesis organisations. It is important for health care practitioners to appreciate this major limitation of CPGs.

This methods study is one of only a few studies to assess the use and citation of systematic reviews with or without pairwise meta-analysis, systematic reviews with network meta-analyses, and 'overviews of systematic reviews' in CPGs. The majority of CPGs in our sample used reviews to underpin recommendations, but only a third cited Cochrane reviews. A systematic process should be followed to ensure the evidence synthesis is accurate, valid, of the highest methodological quality, and based on all available scientific information.

References

1. Marciano, N., et al., *To what extent are current guidelines for cutaneous melanoma follow up based on scientific evidence?* International journal of clinical practice, 2014. **68**(6): p. 761-770.

2. McAlister, F.A., et al., *How evidence-based are the recommendations in evidence-based guidelines?* PLoS medicine, 2007. **4**(8).

3. Qaseem, A., et al., *Guidelines International Network: toward international standards for clinical practice guidelines.* Annals of internal medicine, 2012. **156**(7): p. 525-531.

4. National Institute for Health and Clinical Excellence [Internet]. *National Institute for Health and Clinical Excellence. 2019. Avilable from:* http://www.nice.org.uk.

5. NHMRC, *Guidelines for Guidelines Handbook [Draft].* https://nhmrc.gov.au/guidelinesforguidelines. 2018, National Health and Medical Research Council, Australia Government: NSW, Australia.

6. Ansari, S. and A. Rashidian, *Guidelines for Guidelines: Are They Up to the Task? A Comparative Assessment of Clinical Practice Guideline Development Handbooks.* PLOS ONE, 2012. **7**(11): p. e49864.

7. Woolf, S., et al., *Developing clinical practice guidelines: types of evidence and outcomes; values and economics, synthesis, grading, and presentation and deriving recommendations.* Implementation science : IS, 2012. **7**: p. 61-61.

8. Graham, R., et al., *Clinical Practice Guidelines We Can Trust*, Institute Of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, Editor. 2011, National Academies Press. Available from: https://www.ncbi.nlm.nih.gov/books/NBK209546/: US.

9. Scottish Intercollegiate Guidelines Network, *SIGN 50: a guideline developer's handbook. Available at:* https://www.sign.ac.uk/sign-50.html. 2015: Edinburgh, Scotland.

10. World Health Organization, *WHO handbook for guideline development.* 2014: World Health Organization.

11. Brożek, J., et al., *Grading quality of evidence and strength of recommendations in clinical practice guidelines: part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions.* Allergy, 2009. **64**(5): p. 669-677.

12. Schünemann, H.J., et al., *Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise.* Cmaj, 2014. **186**(3): p. E123-E142.

13. Alonso-Coello, P., et al., *The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies.* Qual Saf Health Care, 2010. **19**(6): p. e58-e58.

14. Armstrong, J.J., et al., *Improvement evident but still necessary in clinical practice guideline quality: a systematic review.* Journal of clinical epidemiology, 2017. **81**: p. 13-21.

15. Kung, J., R.R. Miller, and P.A. Mackowiak, *Failure of clinical practice guidelines to meet institute of medicine standards: two more decades of little, if any, progress.* Archives of internal medicine, 2012. **172**(21): p. 1628-1633.

16. Molino, C., et al., *Factors Associated With High-Quality Guidelines for the Pharmacologic Management of Chronic Diseases in Primary Care: A Systematic Review.* JAMA Intern Med, 2019. **179**(4): p. 553-560.

17. Djulbegovic, B., C.L. Bennett, and G. Guyatt, *Failure to place evidence at the centre of quality improvement remains a major barrier for advances in quality improvement.* Journal of evaluation in clinical practice, 2019. **25**(3): p. 369-372.

18. Djulbegovic, B., C.L. Bennett, and G. Guyatt, *A unifying framework for improving health care.* Journal of evaluation in clinical practice, 2019. **25**(3): p. 358-362.

19. Wayant, C., et al., *Risk of Bias and Quality of Reporting in Colon and Rectal Cancer Systematic Reviews Cited by National Comprehensive Cancer Network Guidelines.* Journal of General Internal Medicine, 2020: p. 1-5.

20. Canadian Association of Radiologists, *CAR Practice Guidelines and Technical Standards for Breast Imaging and Intervention.* 2016: https://car.ca/wp-content/uploads/Breast-Imaging-and-Intervention-2016.pdf.

21. van den Ende, C., et al., *Benefits and harms of breast cancer screening with mammography in women aged 40-49 years: A systematic review.* Int J Cancer, 2017. **141**(7): p. 1295-1306.

22. Eikermann, M., et al., *Tools for assessing the content of guidelines are needed to enable their effective use–a systematic comparison.* BMC research notes, 2014. **7**(1): p. 853.

23. Lunny, C., et al., *Impact and use of reviews and 'overviews of reviews' to inform clinical practice guideline recommendations: protocol for a methods study.* BMJ Open, 2020. **10**(1): p. e031442.

24. American Heart Association, *American Heart Association Methodologies and policies from the ACC/AHA Task Force on Practice Guidelines.* 2010: https://my.americanheart.org/idc/groups/ahamah-public/@wcm/@sop/documents/downloadable/ucm_319826.pdf.

25. Shea, B.J., et al., *External validation of a measurement tool to assess systematic reviews (AMSTAR).* PloS one, 2007. **2**(12): p. e1350.

26. Whiting, P., et al., *ROBIS: A new tool to assess risk of bias in systematic reviews was developed.* J Clin Epidemiol, 2016. **69**: p. 225-34.

27. Tso, G.J., et al., *Automating Guidelines for Clinical Decision Support: Knowledge Engineering and Implementation.* Annual Symposium proceedings. AMIA Symposium, 2017. **2016**: p. 1189-1198.

28. McDonagh, M.S., et al., *Methods for the drug effectiveness review project.* BMC Med Res Methodol, 2012. **12**: p. 140.

29. Whiting, P.F., et al., *QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.* Ann Intern Med, 2011. **155**(8): p. 529-36.

30. Peterson, J., et al., *The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.* Ottawa: Ottawa Hospital Research Institute, 2011.

31. Shea, B.J., et al., *AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both.* bmj, 2017. **358**: p. j4008.

32. World Health Organization, *WHO guidelines on integrated care for older people (ICOPE).* WHO. World Health Organization, Geneva, 2017.

33. National, G.A.U., . *Cystic Fibrosis: Diagnosis and management.* 2017a, National Institute for Health and Care Excellence:
https://www.ncbi.nlm.nih.gov/books/NBK464183/pdf/Bookshelf_NBK464183.pdf.

34. National Guideline Alliance (UK), *Eating Disorders: Recognition and Treatment. NICE Guideline, No. 69.* 2017b, National Institute for Health and Care Excellence: London, UK.

35. Saxby, N., et al., *Nutrition guidelines for cystic fibrosis in Australia and New Zealand.* Sydney: thoracic society of Australia and New Zealand, 2017.

36. Djulbegovic, B. and G. Guyatt, *Evidence vs Consensus in Clinical Practice Guidelines.* JAMA, 2019. **322**(8): p. 725-726.

37. Shaneyfelt, T.M., M.F. Mayo-Smith, and J. Rothwangl, *Are guidelines following guidelines?: The methodological quality of clinical practice guidelines in the peer-reviewed medical literature.* Jama, 1999. **281**(20): p. 1900-1905.

38. Schunemann, H.J., et al., *GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT.* J Clin Epidemiol, 2017. **81**: p. 101-110.

39. NICE, *How we develop NICE guidelines.* https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-guidelines/how-we-develop-nice-guidelines. 2020.

40. Infectious Diseases Society of America, *Clinical Practice Guidelines Development: Training and Resources.* https://www.idsociety.org/es/practice-guideline/clinical-practice-guidelines-development-training-and-resources/

41. Fanaroff, A.C., et al., *Levels of Evidence Supporting American College of Cardiology/American Heart Association and European Society of Cardiology Guidelines, 2008-2018.* Jama, 2019. **321**(11): p. 1069-1080.

42. Meyer, C., et al., *Scientific evidence underlying the American College of Gastroenterology's clinical practice guidelines.* PloS one, 2018. **13**(10).

43. Venkatesh, A.K., et al., *Systematic review of emergency medicine clinical practice guidelines: Implications for research and policy.* PloS one, 2017. **12**(6).

44. Wright, J.D., et al., *Scientific evidence underlying the American College of Obstetricians and Gynecologists' practice bulletins.* Obstetrics & Gynecology, 2011. **118**(3): p. 505-512.

45. Joanna Briggs Institute, *Joanna Briggs Institute reviewers' manual: 2014 edition.* Australia: The Joanna Briggs Institute, 2014.

46. Satya-Murti, S. and J. Lockhart, *Recognizing and reducing cognitive bias in clinical and forensic neurology.* Neurology. Clinical practice, 2015. **5**(5): p. 389-396.

47. Chalmers, I., et al., *How to increase value and reduce waste when research priorities are set.* The Lancet, 2014. **383**(9912): p. 156-165.

48. Chalmers, I. and P. Glasziou, *Avoidable waste in the production and reporting of research evidence.* The Lancet, 2009. **374**(9683): p. 86-89.

49. Clarke, M., S. Hopewell, and I. Chalmers, *Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting.* The Lancet, 2010. **376**(9734): p. 20-21.

50. Robinson, K.A. and S.N. Goodman, *A systematic examination of the citation of prior research in reports of randomized, controlled trials.* Annals of internal medicine, 2011. **154**(1): p. 50-55.

51. WHO, *WHO handbook for guideline development – 2nd edition.* https://www.who.int/publications/guidelines/handbook_2nd_ed.pdf. 2014: Geneva, Switzerland.

52. Bunn, F., et al., *The impact of Cochrane Reviews: a mixed-methods evaluation of outputs from Cochrane Review Groups supported by the National Institute for Health Research.* Health Technol Assess, 2015. **19**(28): p. 1-99, v-vi.

53. Prayle, A.P., et al., *Do guidelines for treating chest disease in children use Cochrane Reviews effectively? A systematic review.* Thorax, 2018. **73**(7): p. 670-673.

54. Silagy, C.A., L.F. Stead, and T. Lancaster, *Use of systematic reviews in clinical practice guidelines: case study of smoking cessation.* Bmj, 2001. **323**(7317): p. 833-6.

55. Lunny, C., et al., *Overviews of reviews incompletely report methods for handling overlapping, discordant and problematic data.* Journal of clinical epidemiology, 2019: p. 69-85.

56. Lunny, C., et al., *Evidence map of studies evaluating methods for conducting, interpreting and reporting overviews of systematic reviews of interventions: rationale and design.* Syst Rev, 2016. **5**: p. 4.

57. Lunny, C., et al., *Toward a comprehensive evidence map of overview of systematic review methods: paper 1-purpose, eligibility, search and data extraction.* Syst Rev, 2017. **6**(1): p. 231.

58. Lunny, C., et al., *Toward a comprehensive evidence map of overview of systematic review methods: paper 2-risk of bias assessment; synthesis, presentation and summary of the findings; and assessment of the certainty of the evidence.* Syst Rev, 2018. **7**(1): p. 159.

59. Pollock, A., *Cochrane overviews: how can we optimize their impact on evidence-based rehabilitation?* European journal of physical and rehabilitation medicine, 2019. **55**(3): p. 395-410.

60. Mhaskar, R., et al., *Critical appraisal skills are essential to informed decision-making.* Indian journal of sexually transmitted diseases and AIDS, 2009. **30**(2): p. 112-119.

61. Schulz, K., D. Altman, and D. Moher, *Factors that limit the quality, number and progress of randomised controlled trials.* BMC Med, 2010. **8**(18): p. 7015-8.

62. Egger, M., et al., *How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study.* Health Technol Assess, 2003. **7**(1): p. 1-76.

63. Petticrew, M., *Why certain systematic reviews reach uncertain conclusions.* BMJ : British Medical Journal, 2003. **326**(7392): p. 756-758.

64. Turner, R.M., S.M. Bird, and J.P. Higgins, *The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews.* PloS one, 2013. **8**(3).

65. Rosenfeld, R.M. and R.N. Shiffman, *Clinical practice guideline development manual: a quality-driven approach for translating evidence into action.* Otolaryngology–head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery, 2009. **140**(6 Suppl 1): p. S1-S43.

66. Moher, D., A. Booth, and L. Stewart, *How to reduce unnecessary duplication: use PROSPERO.* BJOG: An International Journal of Obstetrics & Gynaecology, 2014. **121**(7): p. 784-786.

67. Sideri, S., S.N. Papageorgiou, and T. Eliades, *Registration in the international prospective register of systematic reviews (PROSPERO) of systematic review protocols was associated with increased review quality.* Journal of clinical epidemiology, 2018. **100**: p. 103-110.

68. Stewart, L., D. Moher, and P. Shekelle, *Why prospective registration of systematic reviews makes sense.* 2012, BioMed Central.

69. Soumerai, S.B., D. Starr, and S.R. Majumdar, *How do you know which health care effectiveness research you can trust? A guide to study design for the perplexed.* Preventing chronic disease, 2015. **12**.

70. Shea, B., et al., *A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals.* Eval Health Prof, 2002. **25**(1): p. 116-29.

71. Petticrew, M., et al., *Quality of Cochrane reviews. Quality of Cochrane reviews is better than that of non-Cochrane reviews.* Bmj, 2002. **324**(7336): p. 545.

72. Eccleston, C., et al., *Growing pain: striving for quality, relevance and applicability in Cochrane Reviews.* Cochrane Database Syst Rev, 2013(5): p. Ed000059.

Declarations

All authors have met the ICMJE criteria for authorship by having substantial contributions to the conception or design of the work; and have approved the final version to be published; and have agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Patient and Public Involvement subsection: Patients and the public have not been involved in this research.

Appendices

Appendix 1: Search Strategies

Date searched: January 8, 2019

Epistemonikos

Dates searched: January 1, 2017 to December 31, 2018

Limit: Broad syntheses

Turning Research Into Practice (TRIP)

Dates searched: January 1, 2017 to December 31, 2018

Limit: None

Appendix 2: Data extraction form

A data extraction form was developed in Microsoft Excel (2013). Ten CPGs were independently extracted by six authors and then discussed to come to consensus about definitions and to calibrate the coding.

The form is located in the raw data file on repository Open Science Framework (https://osf.io/8rxnp/ with DOI 10.17605/OSF.IO/8RXNP).

Appendix 3: Included Clinical practice guidelines

1. Addington J, Addington D, Abidi S, Raedler T, Remington G. Canadian Treatment Guidelines for Individuals at Clinical High Risk of Psychosis. Can J Psychiatry. 2017;62(9):656-61.

2. Allan GM, Ramji J, Perry D, Ton J, Beahm NP, Crisp N, et al. Simplified guideline for prescribing medical cannabinoids in primary care. Can Fam Physician. 2018;64(2):111-20.

3. Badhwar V, Rankin JS, Damiano RJ, Jr., Gillinov AM, Bakaeen FG, Edgerton JR, et al. The Society of Thoracic Surgeons 2017 Clinical Practice Guidelines for the Surgical Treatment of Atrial Fibrillation. Ann Thorac Surg. 2017;103(1):329-41.

4. Baumgartner H, Falk V, Bax JJ, De Bonis M, Hamm C, Holm PJ, et al. 2017 ESC/EACTS Guidelines for the management of valvular heart disease. Eur Heart J. 2017;38(36):2739-91.

5. Braga LH, Lorenzo AJ, Romao RLP. Canadian Urological Association-Pediatric Urologists of Canada (CUA-PUC) guideline for the diagnosis, management, and followup of cryptorchidism. Can Urol Assoc J. 2017;11(7):E251-e60.

6. British Columbia Centre for Substance Abuse (BCCSU). A Guideline for the Clinical Management of Opioid Use Disorder. https://www.bccsu.ca/wp-content/uploads/2017/06/BC-OUD-Guidelines_June2017.pdf. BC, Canada: BC Ministry of Health; 2017.

7. British Columbia Centre for Substance Abuse (BCCSU). A guideline for the clinical management of opioid use disorder. In: Committee POUDTG, editor. Vancouver, Canada. https://www.bccsu.ca/wp-content/uploads/2017/06/BC-OUD-Guidelines_June2017.pdf: British Columbia Centre on Substance Use; 2017.

8. Brooks AJ, Smith PJ, Cohen R, Collins P, Douds A, Forbes V, et al. UK guideline on transition of adolescent and young persons with chronic digestive diseases from paediatric to adult care. Gut. 2017;66(6):988-1000.

9. Chang SS, Bochner BH, Chou R, Dreicer R, Kamat AM, Lerner SP, et al. Treatment of Non-Metastatic Muscle-Invasive Bladder Cancer: AUA/ASCO/ASTRO/SUO Guideline. J Urol. 2017;198(3):552-9.

10. Chin J, Rumble RB, Kollmeier M, Heath E, Efstathiou J, Dorff T, et al. Brachytherapy for Patients With Prostate Cancer: American Society of Clinical Oncology/Cancer Care Ontario Joint Guideline Update. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(15):1737-43.

11. Corcos J, Przydacz M, Campeau L, Gray G, Hickling D, Honeine C, et al. CUA guideline on adult overactive bladder. Can Urol Assoc J. 2017;11(5):E142-e73.

12. Dreyling M, Campo E, Hermine O, Jerkeman M, Le Gouill S, Rule S, et al. Newly diagnosed and relapsed mantle cell lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2017;28(suppl_4):iv62-iv71.

13. Duceppe E, Parlow J, MacDonald P, Lyons K, McMullen M, Srinathan S, et al. Canadian Cardiovascular Society Guidelines on Perioperative Cardiac Risk Assessment and Management for Patients Who Undergo Noncardiac Surgery. Can J Cardiol. 2017;33(1):17-32.

14. Duggal M, Tong HJ, Al-Ansary M, Twati W, Day PF, Nazzal H. Interventions for the endodontic management of non-vital traumatised immature permanent anterior teeth in children and adolescents: a systematic review of the evidence and guidelines of the European Academy of Paediatric Dentistry. Eur Arch Paediatr Dent. 2017;18(3):139-51.

15. European Society for Medical Oncology (ESMO) Guidelines Committee. eUpdate–Renal Cell Carcinoma Treatment Recommendations. European Society for Medical Oncology; 2017a.

16. European Society for Medical Oncology (ESMO) Guidelines Committee. eUpdate– Chronic Lymphocytic Leukaemia Treatment Recommendations. Lugano, Switzerland: ESMO. https://www.esmo.org/Guidelines/Haematological-Malignancies/Chronic-Lymphocytic-Leukaemia/eUpdate-Treatment-Recommendations; 2017b.

17. Finelli A, Ismaila N, Bro B, Durack J, Eggener S, Evans A, et al. Management of small renal masses: American Society of Clinical Oncology clinical practice guideline. 2017;35(6):668-80.

18. Franz MJ, MacLeod J, Evert A, Brown C, Gradwell E, Handu D, et al. Academy of Nutrition and Dietetics Nutrition Practice Guideline for Type 1 and Type 2 Diabetes in Adults: Systematic Review of Evidence for Medical Nutrition Therapy Effectiveness and Recommendations for Integration into the Nutrition Care Process. J Acad Nutr Diet. 2017;117(10):1659-79.

19. Frost JL, Campos-Outcalt D, Hoelting D. Pharmacologic management of newly detected atrial fibrillation. April 2017. https://www.aafp.org/dam/AAFP/documents/patient_care/clinical_recommendations/a-fib-guideline.pdf: American Academy of Family Physicians; 2017.

20. Fung KFK, Eason E. No. 133-Prevention of Rh Alloimmunization. Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC. 2018;40(1):e1-e10.

21. Gordon C, Amissah-Arthur MB, Gayed M, Brown S, Bruce IN, D'Cruz D, et al. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults: Executive Summary. Rheumatology (Oxford). 2018;57(1):14-8.

22. Hadji P, Aapro MS, Body JJ, Gnant M, Brandi ML, Reginster JY, et al. Management of Aromatase Inhibitor-Associated Bone Loss (AIBL) in postmenopausal women with hormone sensitive breast cancer: Joint position statement of the IOF, CABS, ECTS, IEG, ESCEO IMS, and SIOG. J Bone Oncol. 2017;7:1-12.

23. Hadjipanayis CG, Carlson ML, Link MJ, Rayan TA, Parish J, Atkins T, et al. Congress of Neurological Surgeons Systematic Review and Evidence-Based Guidelines on Surgical Resection for the Treatment of Patients With Vestibular Schwannomas. Neurosurgery. 2018;82(2):E40-e3.

24. Hanley J, McKernan A, Creagh MD, Classey S, McLaughlin P, Goddard N, et al. Guidelines for the management of acute joint bleeds and chronic synovitis in haemophilia: A United Kingdom Haemophilia Centre Doctors' Organisation (UKHCDO) guideline. Haemophilia. 2017;23(4):511-20.

25. Ho VP, Patel NJ, Bokhari F, Madbak FG, Hambley JE, Yon JR, et al. Management of adult pancreatic injuries: A practice management guideline from the Eastern Association for the Surgery of Trauma. J

Trauma Acute Care Surg. 2017;82(1):185-99.

26. Isaacs CG, Kistler, C., Hunold, K.M., Pereira, G.F., Buchbinder, M., Weaver, M.A., Doherty, S., Knott, J., Bennetts, S., Jazayeri, M., Huckson, S. VA/DoD clinical practice guideline for opioid therapy for chronic pain. Department of Veterans Affairs. Department of Defense. https://pami.emergency.med.jax.ufl.edu/wordpress/files/2019/07/Basic-of-Pain-References-07012019.pdf; 2017.

27. KDIGO Update Work Group. KDIGO 2017 Clinical Practice Guideline Update for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). Kidney Int Suppl. 2017;7(1):1-59.

28. Kusumoto FM, Bailey KR, Chaouki AS, Deshmukh AJ, Gautam S, Kim RJ, et al. Systematic review for the 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. Heart Rhythm. 2018;15(10):e253-e74.

29. Lutz S, Balboni, T., Jones, J., Lo, S., Petit, J., Rich, S. E., Wong, R., Hahn, C. Palliative radiation therapy for bone metastases: Update of an ASTRO Evidence-Based Guideline. Pract Radiat Oncol. 2017;7(1):4-12.

30. Millington GWM, Collins A, Lovell CR, Leslie TA, Yong ASW, Morgan JD, et al. British Association of Dermatologists' guidelines for the investigation and management of generalized pruritus in adults without an underlying dermatosis, 2018. Br J Dermatol. 2018;178(1):34-60.

31. Moreau P, San Miguel J, Sonneveld P, Mateos MV, Zamagni E, Avet-Loiseau H, et al. Multiple myeloma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2017;28 Suppl 4:iv52-iv61.

32. Mowery NT, Bruns BR, MacNew HG, Agarwal S, Enniss TM, Khan M, et al. Surgical management of pancreatic necrosis: A practice management guideline from the Eastern Association for the Surgery of Trauma. J Trauma Acute Care Surg. 2017;83(2):316-27.

33. Mulhall JP, Trost LW, Brannigan RE, Kurtz EG, Redmon JB, Chiles KA, et al. Evaluation and Management of Testosterone Deficiency: AUA Guideline. J Urol. 2018;200(2):423-32.

34. National GAU, . Cystic Fibrosis: Diagnosis and management. https://www.ncbi.nlm.nih.gov/books/NBK464183/pdf/Bookshelf_NBK464183.pdf: National Institute for Health and Care Excellence; 2017a.

35. National Guideline Alliance (UK). Eating Disorders: Recognition and Treatment. NICE Guideline, No. 69. London, UK: National Institute for Health and Care Excellence; 2017b.

36. Panel on Antiretroviral Therapy and Medical Management of Children Living with HIV. Guidelines for the use of antiretroviral agents in pediatric HIV infection. 2017. US Office of AIDS Research Advisory Council (OARAC). https://aidsinfo.nih.gov/guidelines/html/2/pediatric-arv/45/whats-new-in-the-guidelines

https://aidsinfo.nih.gov/contentfiles/lvguidelines/PediatricGuidelines.pdf 2017.

37. Practice Committee of the American Society for Reproductive Medicine. Role of metformin for ovulation induction in infertile patients with polycystic ovary syndrome (PCOS): a guideline. Fertil Steril. 2017;108(3):426-41.

38. Qaseem A, Barry MJ, Humphrey LL, Forciea MA. Oral Pharmacologic Treatment of Type 2 Diabetes Mellitus: A Clinical Practice Guideline Update From the American College of Physicians. Ann Intern Med. 2017;166(4):279-90.

39. Registered Nurses Association of Ontario (RNAO). Integrating tobacco interventions into daily practice. Toronto (ON): Registered Nurses Association of Ontario (RNAO). https://www.guidelinecentral.com/summaries/integrating-tobacco-interventions-into-daily-practice/#section-442; 2017.

40. Roberts G, Pfaar O, Akdis CA, Ansotegui IJ, Durham SR, Gerth van Wijk R, et al. EAACI Guidelines on Allergen Immunotherapy: Allergic rhinoconjunctivitis. Allergy. 2018;73(4):765-98.

41. Roselló S, Blasco I, García Fabregat L, Cervantes A, Jordan K. Management of infusion reactions to systemic anticancer therapy: ESMO Clinical Practice Guidelines. Ann Oncol. 2018;29(Suppl 4):iv260.

42. Rosen R, Vandenplas Y, Singendonk M, Cabana M, DiLorenzo C, Gottrand F, et al. Pediatric Gastroesophageal Reflux Clinical Practice Guidelines: Joint Recommendations of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition and the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition. Journal of pediatric gastroenterology and nutrition. 2018;66(3):516-54.

43. Sateia MJ, Buysse DJ, Krystal AD, Neubauer DN, Heald JL. Clinical Practice Guideline for the Pharmacologic Treatment of Chronic Insomnia in Adults: An American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med. 2017;13(2):307-49.

44. Saxby N, Painter C, Kench A, King S, Crowder T, van der Haak N. Nutrition guidelines for cystic fibrosis in Australia and New Zealand. Sydney: thoracic society of Australia and New Zealand. 2017.

45. Schmidt-Hieber M, Bierwirth J, Buchheidt D, Cornely OA, Hentrich M, Maschmeyer G, et al. Diagnosis and management of gastrointestinal complications in adult cancer patients: 2017 updated evidence-based guidelines of the Infectious Diseases Working Party (AGIHO) of the German Society of Hematology and Medical Oncology (DGHO). Annals of hematology. 2018;97(1):31-49.

46. Tan DHS, Hull MW, Yoong D, Tremblay C, O'Byrne P, Thomas R, et al. Canadian guideline on HIV pre-exposure prophylaxis and nonoccupational postexposure prophylaxis. Cmaj. 2017;189(47):E1448-e58.

47. Van Poznak C, Somerfield MR, Barlow WE, Biermann JS, Bosserman LD, Clemons MJ, et al. Role of Bone-Modifying Agents in Metastatic Breast Cancer: An American Society of Clinical Oncology-Cancer Care Ontario Focused Guideline Update. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(35):3978-86.

48. Virgo KS, Basch E, Loblaw DA, Oliver TK, Rumble RB, Carducci MA, et al. Second-Line Hormonal Therapy for Men With Chemotherapy-Naive, Castration-Resistant Prostate Cancer: American Society of Clinical Oncology Provisional Clinical Opinion. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(17):1952-64.

49. Wessells H, Angermeier KW, Elliott S, Gonzalez CM, Kodama R, Peterson AC, et al. Male Urethral Stricture: American Urological Association Guideline. J Urol. 2017;197(1):182-90.

50. World Health Organization. WHO guidelines on integrated care for older people (ICOPE). WHO World Health Organization, Geneva. 2017.

Appendix 4: Table of associations

Table of chi-square test of independence results

| Predictor variable	Dependent variable	Proportion yes	Chi squared results	P value
GRADE compliant	Funder (all)		9.491	0.050
	· Funder: government	12%	4.48	0.034
	· Funder: Medical society or association	12%	0.09	0.768
	· Funder: pharmaceutical industry	0%	6.17	0.013
	· No funding	4%	0.02	0.880
	· Funding source not reported	8%	0.81	0.368
	Scope (narrow, broad)	22%	1.4	0.20
	Conflict of Interest	34%	0.18	0.07
	Affiliation with the pharmaceutical industry	18%	2.4	0.10
	Continent (all)3		6.0	0.05
	· International	6%	5.674	0.017
	· Europe	8%	1.2357	0.266
	· North America	22%	0.014468	0.904
Systematic process used	Funder (all)2		3.60	0.463
	· Government	8%	0.7811	0.377
	· Medical society or association	10%	0.23042	0.631
	· Pharmaceutical industry	2%	2.2009	0.138
	· No funding	6%	1.0152	0.314
	· Not reported	6%	0.1324	0.716
	Scope (narrow, broad)	16%	0	1.0
	Affiliation with the pharmaceutical industry	18%	0.6134	0.433
	Conflict of Interest	32%	0.9804	0.322
	Continent (all)3		7.551	0.023
	· International	6%	6.782	0.009
	· Europe	10%	0.006082	0.938
	· North America	14%	2.5888	0.108

2: Df=4, 3: Df=2
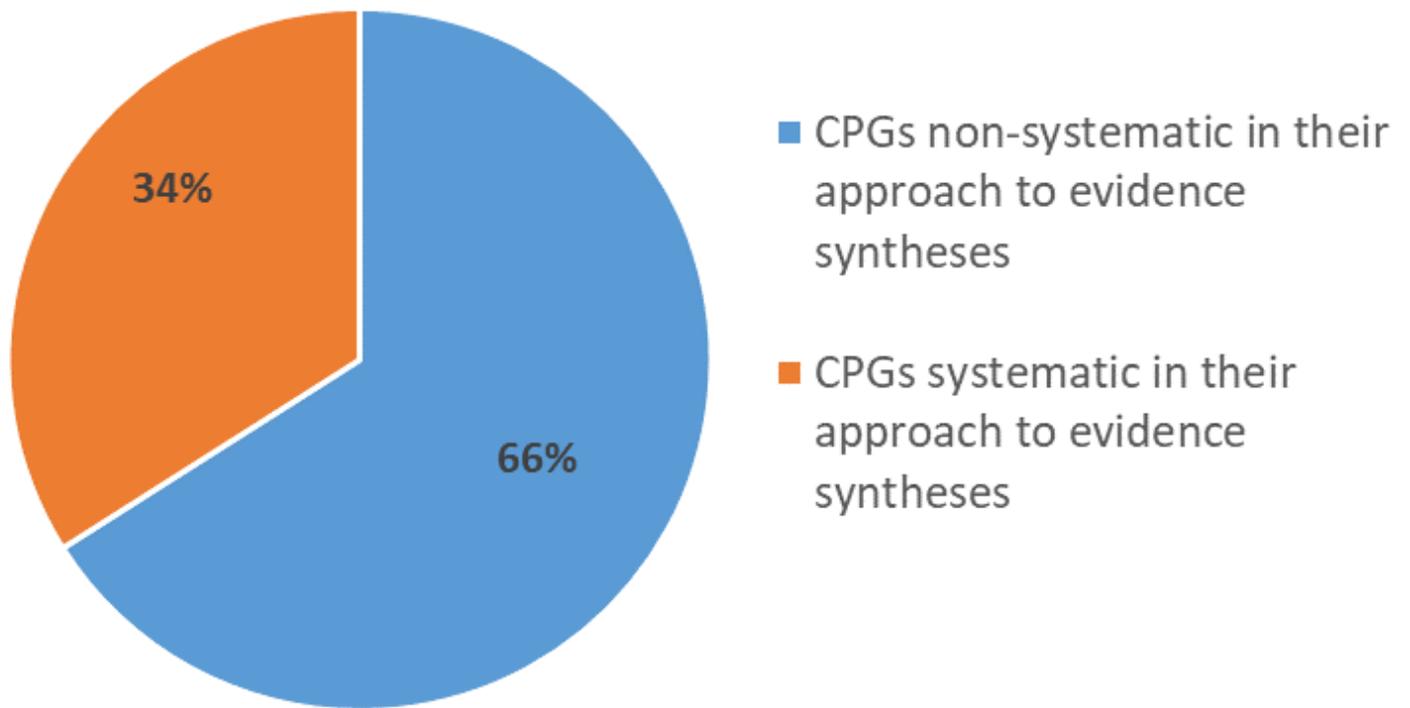
Figures

Figure 2

Systematic or non-systematic process used by CPGs in their approach to evidence synthesis (n=50)