Updated clinical guidelines: improving their methods and reporting

Robin Wilhelmus Maria Vernooij

Universitat Autònoma de Barcelona

Faculty of Medicine

Department of Paediatrics, Obstetrics, Gynaecology and Preventive Medicine

PhD programme in Methodology of Biomedical Research and Public Health

Updated clinical guidelines: improving their methods and reporting

DOCTORAL THESIS

Robin Wilhelmus Maria Vernooij

Directors: dr. Pablo Alonso-Coello, dr. Laura Martínez García

Tutor: dr. Xavier Bonfill i Cosp

Barcelona, December 2017

Universitat Autònoma de Barcelona

Faculty of Medicine

Department of Paediatrics, Obstetrics, Gynaecology and Preventive Medicine

PhD programme in Methodology of Biomedical Research and Public Health

Updated clinical guidelines: improving their methods and reporting

DOCTORAL THESIS

Robin Wilhelmus Maria Vernooij

Barcelona, December 2017

Thesis report as a compendium of publications presented by Robin Wilhelmus Maria Vernooij to apply for a PhD from the Autonomous University of Barcelona and conducted under the direction of dr. Pablo Alonso Coello and dr. Laura Martínez García.

ACKNOWLEDGEMENTS

Finalising this thesis is the result of inspiration and guidance from many sources. Far too many people to mention by name have made this thesis possible, yet below I would like to express my gratitude towards the persons most involved in this journey.

Firstly, I would like to express my profound gratitude to the person who inspired me in starting this thesis: dr. **Pablo Alonso-Coello**. You have pushed me towards greater success on a daily basis and you did not allow me to get away easy. Thanks for helping me navigate my journey through this thesis and for creating so many opportunities along the way.

Secondly, without dr. **Laura Martínez García** this thesis would have never been finished. You have taught me much more than only conducting proper research, but as well to look more critically at own work and not take the easy road. I am confident that I will benefit the rest of my career from the professional, organisational, and academic skills you have taught me over the years.

The third person I would like to thank made it possible that I was able to work at the Iberoamerican Cochrane centre for 2.5 years and consequently allowed me to learn from his experience, skills, and capabilities. Dr. **Xavier Bonfill i Cosp**, I would like to thank you for all the opportunities you have offered me to improve my academic skills, gain experiences, and grow my academic network.

Thanks to all the co-authors that helped me with publishing the articles included in this thesis: **Juliana, Ivan (Solà), Ivan (D. Florez), Michiel, Melissa,** and all members of the **CheckUp panel**. I am grateful for the fact that you have shared your knowledge and experiences with me and it was a joy and honour working together with you.

To my old colleagues at the Iberoamerican Cochrane Center; thanks for creating such a welcoming and pleasant environment. I especially want to thank **Maria José (Diaz)**, **Meritxell, Andrea**, and **Marta (Gorgues)** for paving the road of finalising this thesis. I have had the honour to work with extremely intelligent researchers during my stay at the Iberoamerican Cochrane Centre and I am very much looking forward to continue collaborating with you all in the near future. Finally, **Hector**, I would like to thank you specifically for all the time you have invested in our friendship and this thesis.

To my current colleagues at the Netherlands Comprehensive Cancer Organisation; thanks for all your involvement, compassion, and support in this thesis. I especially want to thank **Katja** and **Valery** for allowing me to gain experience in both the field of guidelines and cancer-registry population-based studies. I am very blessed to work in such an extremely intelligent, fast-pacing, and friendly environment.

Thanks to my wonderful family and friends in both Spain and the Netherlands. I am very fortunate to have a network of loving and caring people around me **Annelies, Jeroen, Maria Paz, Manolo,** and all friends I see on a regular basis, thanks for being there unconditionally.

Pappa en Mamma: Alles wat ik bereikt heb in de afgelopen jaren en zal gaan doen in de toekomst heb ik aan jullie te danken. Jullie hebben het voor mij mogelijk gemaakt om te leren, te studeren, en te werken in een liefdevolle thuissituatie waar ik altijd mezelf heb kunnen zijn. Dat is alles wat iemand kan wensen. Zonder jullie zou ik niet de persoon zijn die ik nu ben. Mijn dankbaarheid en liefde voor jullie gaat verder dan woorden.

Ana. Everything I have and everything I do is better when you are involved. Life, including finishing this thesis, is just a breeze with you. Thank you for your unfailing support and endless encouragement. Blood, sweat, and tears made me finish this thesis, but your love was the ingredient that made it a success.

TABLE OF CONTENTS

1. ABSTRACT

1.1 Abstract

Introduction

Clinical guidelines (CGs) aim to guide healthcare professionals, patients, and policymakers in decision-making by providing recommendations for a healthcare problem. However, since new evidence is published on a regular basis, CGs may require to be updated in order to guarantee the validity of recommendations. As opposed to the methodology for developing *de novo* CGs, there is scarce guidance available for the updating process of CGs and little is known about the methodology that CG institutions use to maintain the validity of their CGs.

Objectives

The objectives of this thesis are: 1) to identify and describe the updating guidance available in CG methodological handbooks, 2) to develop a checklist for the reporting of updated CGs, and 3) to assess the completeness of reporting of updated contemporary CGs.

Methods

For the first study, we conducted a systematic review of CG methodological handbooks searching in MEDLINE, the Guidelines International Network (G-I-N), and the US National Guidelines Clearinghouse. Two authors independently selected evidence and extracted data. We used descriptive statistics and a narrative synthesis to analyse the extracted data.

For the second study, we developed a reporting instrument for the updating process of CGs. This tool was constructed through a multi-step development process that included an assessment of updated CGs, semi-structured interviews with key informants, a Delphi

consensus survey, a single-round survey with CG methodologists, and semi-structured interviews with CG users.

For the third study, we systematically assessed the reporting of the updating process in updated CGs published in 2015. To be eligible, CGs had to be developed by a professional society, report a systematic review of the evidence, and contain at least one recommendation. Three reviewers independently applied the reporting instrument developed in the second study to the included CGs.

Results

For the first study, we included 35 handbooks. Little guidance for updating CGs was identified. Most handbooks focused mainly on providing guidance for developing CGs *de novo*. The majority of the handbooks did not provide guidance for the literature search, evidence selection, quality assessment, evidence synthesis, or external review during the updating process.

In the second study, we developed the Checklist for the Reporting of Updated Guidelines (CheckUp), which includes 16 items regarding: 1) the presentation of an updated guideline, 2) editorial independence, and 3) the methodology of the updating process. We also developed and explanation and elaboration document for CheckUp with the goal to facilitate the potential users.

In the last study, we included 60 updated CGs. The median overall score with CheckUp on a 10-point scale was 6.3 (range 3.1 to 10). The presentation and justification items at recommendation level and the methods for external review and implementing changes in practice were poorly reported.

Conclusions

The guidance available for the updating of CGs and the reporting of updated CGs is suboptimal. CheckUp is the first reporting instrument in the CG enterprise with a focus on the updating process. CheckUp can be used to assess the completeness of reporting of the updating process in updated CGs, and also guide CG panels by providing methodological and reporting principles that should be incorporated into the updating process.

1.2 Resumen

Título:

Guías de práctica clínica actualizadas: mejora de los métodos y la presentación.

Antecedentes:

Las guías de práctica clínica (GPC) tienen como objetivo orientar en la toma de decisiones a profesionales de la salud, pacientes y responsables de elaborar políticas sanitarias, mediante recomendaciones para un problema de salud. Sin embargo, dado que regularmente se publica nueva evidencia, la actualización de las GPC es necesaria para garantizar la validez de las recomendaciones. A diferencia de lo que ocurre con la metodología para elaborar GPC de novo, existen pocas guías disponibles sobre el proceso de actualización de las GPC y los conocimientos sobre los métodos que utilizan las instituciones para mantener la validez de sus GPC son limitados.

Objectivos:

Los objetivos de esta tesis son 1) identificar y describir las guías de actualización incluidas en los manuales metodológicos de las GPC, 2) desarrollar una lista de verificación para informar sobre el proceso de actualización de las GPC actualizadas y 3) evaluar la exhaustividad de la información sobre el proceso de actualización en las GPC actualizadas mediante el uso de la lista de verificación.

Métodos:

En el primer estudio realizamos una revisión sistemática de los manuales metodológicos de las GPC mediante una búsqueda en MEDLINE, en la base de datos de *Guidelines International Network* (G-I-N) y en la base de datos de *US National Guidelines Clearinghouse*. Dos autores

seleccionaron la evidencia y extrajeron los datos de forma independiente. Para analizar los datos, utilizamos variables estadísticas descriptivas y una síntesis narrativa.

En el segundo estudio desarrollamos un instrumento para informar sobre el proceso de actualización de las GPC. Esta herramienta se elaboró siguiendo un proceso con múltiples etapas que incluyó una evaluación de GPC actualizadas, entrevistas semiestructuradas con expertos en GPC, un cuestionario de consenso Delphi, una encuesta con metodólogos y entrevistas semiestructuradas con usuarios de GPC.

En el tercer estudio evaluamos sistemáticamente la información sobre el proceso de actualización en las GPC actualizadas publicadas en 2015. Las GPC incluidas debían haber sido elaboradas por una sociedad profesional, incluir una revisión sistemática de la evidencia y presentar al menos una recomendación. Tres revisores aplicaron, de forma independiente, la lista de verificación desarrollada en el segundo estudio a las GPC.

Resultados:

En el primer estudio incluimos 35 manuales. Se identificaron pocas guías para la actualización de las GPC. La mayoría de los manuales se centraban principalmente en proporcionar guías para elaborar GPC *de novo*. La mayoría de los manuales no proporcionaron guías para la búsqueda bibliográfica, la selección de la evidencia, la evaluación de la calidad, la síntesis de la evidencia ni la revisión externa durante el proceso de actualización.

En el segundo estudio desarrollamos la lista de verificación para la publicación de GPC actualizadas (CheckUp), que incluye 16 ítems sobre 1) la presentación de una GPC actualizada, 2) la independencia editorial y 3) la metodología del proceso de actualización.

Además, desarrollamos un documento adicional de explicación y elaboración para facilitar su utilización a los potenciales usuarios.

En el último estudio incluimos 60 GPC actualizadas. La mediana de puntuación global con el CheckUp, en una escala de 10 puntos, fue de 6,3 (rango 3,1 a 10). La información fue limitada en relación con los ítems presentación y justificación a nivel de recomendación, métodos para la revisión externa e implementación de modificaciones.

Conclusiones:

Las guías disponibles para actualizar las GPC y la información de las GPC actualizadas son subóptimas. El CheckUp es el primer instrumento en el ámbito de las GPC que se centra en el proceso de actualización. El CheckUp puede utilizarse para evaluar la exhaustividad de la información sobre el proceso de actualización en GPC actualizadas y también para guiar a los grupos de trabajo de las GPC, ya que proporciona los estándares metodológicos y la información que deberían incorporar al proceso de actualización.

1.3 Resum

Títol:

Guies de pràctica clínica actualitzades: millora dels mètodes i la presentació.

Antecedents:

Les guies de pràctica clínica (GPC) tenen com a objectiu orientar en la presa de decisions els professionals de la salut, els pacients i els responsables d'elaborar polítiques sanitàries, mitjançant recomanacions per a un problema de salut. No obstant això, com que regularment es publica nova evidencia, l'actualització de les GPC es necessària per garantir la validesa de les recomanacions. A diferència del que passa amb la metodologia per desenvolupar GPC *de novo*, gairebé no hi ha guies disponibles sobre el procés d'actualització de GPC i els coneixements sobre els mètodes de les institucions per mantenir la validesa de les seves GPC són limitats.

Objectius:

Els objectius d'aquesta tesi són: 1) identificar i descriure les guies d'actualització recollides en manuals metodològics de les GPC, 2) desenvolupar una llista de verificació per a informar sobre el procés d'actualització de les GPC actualitzades i 3) avaluar l'exhaustivitat de la informació sobre el procés d'actualització en les GPC actualitzades mitjançant l'ús de la llista de verificació.

Mètodes:

En el primer estudi vam realitzar una revisió sistemàtica de manuals metodològics de les GPC mitjançant una cerca a MEDLINE, a la base de dades de la *Guidelines International Network* i a la base de dades de l'*US National Guidelines Clearinghouse*. Dos autors van seleccionar l'evidència i

van extreure les dades de forma independent. Per analitzar les dades, vam utilitzar variables estadístiques descriptives i una síntesi narrativa.

En el segon estudi vam desenvolupar un instrument per a informar sobre el procés d'actualització de les GPC. Aquesta eina es va elaborar seguint un procés amb múltiples etapes que va incloure: una avaluació de GPC actualitzades, entrevistes semiestructurades amb experts en GPC, un qüestionari de consens Delphi, una enquesta amb experts en metodologia i entrevistes semiestructurades amb usuaris de GPC.

En el tercer estudi vam avaluar sistemàticament la informació del procés d'actualització en les GPC actualitzades publicades al 2015. Les GPC incloses les havia d'haver elaborat una societat professional, havien d'incloure una revisió sistemàtica de l'evidència i contenir almenys una recomanació. Tres revisors van aplicar, de forma independent, la llista de verificació desenvolupada en el segon estudi a les GPC.

Resultats:

En el primer estudi vam incloure 35 manuals. Es van identificar poques guies per a l'actualització de GPC. La major part dels manuals se centraven principalment en proporcionar guies per elaborar GPC *de novo*. La major part dels manuals no van proporcionar guies per a la cerca bibliogràfica, la selecció de l'evidència, l'avaluació de la qualitat, la síntesi de l'evidència ni la revisió externa durant el procés d'actualització.

En el segon estudi vam desenvolupar la llista de verificació per a la publicació de GPC actualitzades (CheckUp), que inclou 16 ítems sobre: 1) la presentació d'una GPC actualitzada, 2) la independència editorial i 3) la metodologia del procés d'actualització. A més, vam desenvolupar un document addicional d'explicació i elaboració per facilitar la utilització als potencials usuaris.

En l'últim estudi vam incloure 60 GPC actualitzades. La mediana de puntuació global amb CheckUp, en una escala de 10 punts, va ser de 6,3 (rang 3,1 a 10). La informació va ser limitada en relació amb els ítems: presentació i justificació a nivell de recomanació, mètodes per a la revisió externa i implementació de modificacions.

Conclusions:

Les guies disponibles per actualitzar les GPC i la informació de les GPC actualitzades són subòptimes. El CheckUp és el primer instrument de l'àmbit de les GPC que se centra en el procés d'actualització. El CheckUp es pot utilitzar per avaluar l'exhaustivitat de la informació sobre el procés d'actualització de les GPC i també per guiar els grups de treball de les GPC, ja que proporciona els estàndards metodològics i la informació que haurien d'incorporar al procés d'actualització.

LIST OF ALL ABBREVIATIONS

- Appraisal of Guidelines for Research and Evaluation (AGREE)
- CG: Clinical Guideline
- CheckUp: Checklist for the Reporting of Updated Guidelines
- CoI: Conflicts of Interest
- CONSORT: Consolidated Standards of Reporting Trials
- EtD: Evidence to Decision
- G-I-N: Guidelines International Network
- GRADE: Grading of Recommendations Assessment, Development and Evaluation
- IOM: Institute of Medicine
- KDIGO: Kidney Disease: Improving Global Outcomes
- MAGIC: Making GRADE the Irresistible Choice
- MeSH: Medical Subject Headings
- NICE: National Institute for Health and Care Excellence
- PICO: Patient, Intervention, Comparison, Outcome
- PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analysis
- RCT: Randomised Controlled Trial
- RIGHT: Reporting Items for practice Guidelines in HealThcare
- SR: Systematic Review
- STARD: Standards for Reporting of Diagnostic Accuracy Studies.
- STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
- WHO: World Health Organisation

2. INTRODUCTION

2.1 The problem of information overload

The use of up-to-date evidence is considered crucial in healthcare decision-making to ensure adequate patient care [1-3]. Due to an overwhelming volume of newly available evidence [4-8], the identification and assessment of new evidence is a time-consuming task for clinicians. Implementing the latest evidence into healthcare practice has proven to be extremely challenging [9].

One potential solution to keep clinicians up-to-date about the latest developments is the use of health decision-making tools, such as systematic reviews (SRs), that select, synthesise, and appraise the quality of the available evidence. SRs summarise the results of studies on a specific clinical question (e.g. the effectiveness of a healthcare interventions, the diagnostic test accuracy of a test modality, cost-effectiveness, or values and preferences) [10]. Different study designs might be included in a SR depending on the type of clinical question. For example, where SRs about the therapeutic effectiveness of an intervention will likely prioritise randomised controlled trials (RCTs), SRs about the of patients' values and preferences will likely search for observational studies and qualitative research [10].

Another health decision-making tool that became very popular among clinicians to remain up-to-date are clinical guidelines (CGs). The similarity between CGs and SRs is that they both aim to select, synthesise, and assess the quality of the studies to generate an evidence profile for answering a clinical question. In other words, CGs often perform or use a SR to synthesise the results of previously published evidence and to evaluate the quality of evidence. However, CGs and SRs differ in several aspects. Firstly, SRs do not intend to provide best-practice recommendations for healthcare professionals, where CGs do [10]. CGs are characterised by

a multipronged approach in a certain clinical area, a methodological rigour of translating evidence into recommendations, and the involvement of many different stakeholders in a CG panel [11,12]. Additionally, the amount of newly published SRs is overwhelming and are widely scattered among different journals [13]. Therefore, CGs are considered the preferred health decision-making tool, with recommendations for best-practice, to keep clinicians up-to-date [11].

2.2 Definition of clinical guidelines

2.2.1 Definition and objectives

There is no consensus in the CG enterprise regarding the definition of CGs (**Table 1**). However, most definitions focus on the importance of including recommendations that are informed by a SR of the evidence aiming to improve the quality of healthcare [11, 14-16]. The most commonly used definition of CGs is from the Institute of Medicine (IOM) from 2011: "*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*" [11].

Table 1: Definition of clinical guidelines by international institutions responsible for developing and updating clinical guidelines.

Institution	CG definition	CG objective
American College of Chest Physicians (ACCP) [16]	Systematically developed set of recommendations, algorithms, and other information to assist health-care decision making in specific clinical circumstances.	To assist healthcare decision making
American Urological Association (AUA) [17]	Evidence based guidance with an explicit clinical scope and purpose.	No objective stated in the definition
Institute of Medicine (IOM) [11]	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.	To optimise patient care

National Institute for Health and Care Excellence (NICE) [15]	NICE guidelines make evidence-based recommendations on a wide range of topics, from preventing and managing specific conditions, improving health, and managing medicines in different settings, to providing social care and support to adults and children, safe staffing, and planning broader services and interventions to improve the health of communities. They aim to promote individualised care and integrated care (for example, by covering transitions between children's and adult services and between health and social care).	To promote individualised and integrated care
New Zealand Guidelines Group (NZGG) [18]	Guidelines provide guidance in decision making at each level of interaction; between health professionals and consumers, between purchaser and provider, and between 'funder' and 'purchaser'.	To provide guidance in decision making
World Health Organisation (WHO) [14]	Any document containing recommendations about health interventions, whether these are clinical, public health or policy recommendations.	No objective stated in the definition

CGs have multiple purposes and can serve to: 1) guide healthcare professionals in evidence-based medicine decision-making to improve the quality of healthcare [19,20], 2) reduce unwanted clinical practice variation [21-24], 3) improve the use of resources in healthcare practice [25], 4) summarise research findings to make clinical decision-making more transparent, 5) identify gaps in knowledge [26], and 6) provide guidance for consumers and inform or empower patients [26]. Therefore, CGs might facilitate healthcare professionals in making clinical decisions, identify quality improvement efforts, and prioritising new research initiatives which might lead to better health outcomes, less ineffective treatments, and greater consistency of care [27].

2.2.3 Clinical guideline popularity and quality

CGs have become a popular resource for healthcare professionals. Their publication of CGs has increased steeply over the last three decades (**Figure 1**), with over 23,000 references in PubMed as of December 2017.

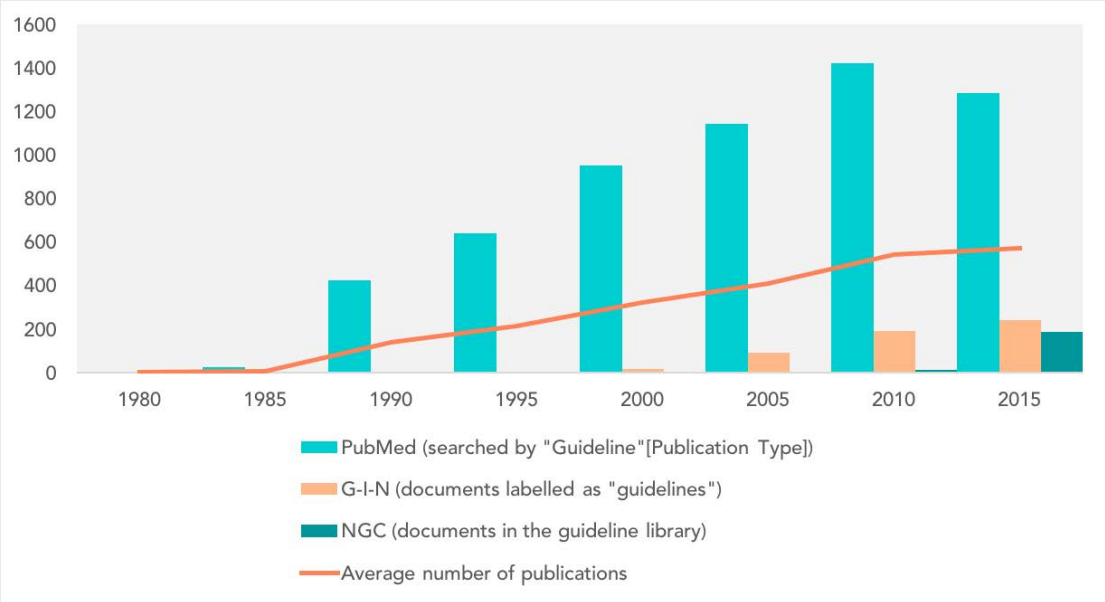


Figure 1: Number of published clinical guidelines.

2.3 Clinical guideline development

Although the development process of *de novo* CGs differs among institutions [14,15,28,29], it generally includes the following steps: 1) prioritisation process, 2) defining the scope and purpose, 3) convening a guidance panel, 4) formulating the clinical questions, 5) systematically searching, selecting, synthesising, and assessing the evidence, 6) formulating recommendations, and 7) external review (**Figure 2**).

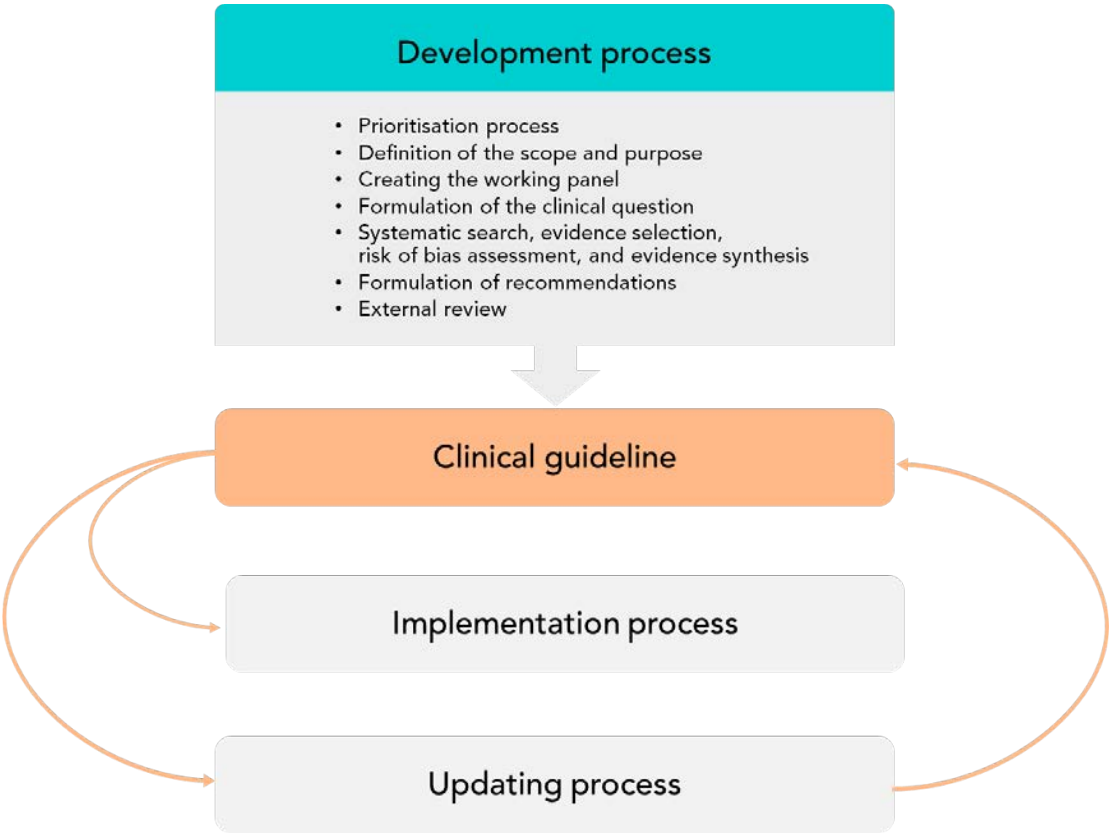


Figure 2: Clinical guideline life-cycle including the development process, implementation process, and updating process.

2.3.1. Prioritisation process

Due to limited resources and capacity, CG developers often need to decide which CG topic has the highest priority for developing to ensure that resources are invested in documents that are most relevant to different stakeholders. Priority should be given to CGs that are most likely to improve health, equity, and efficient use of healthcare resource, always based on a systematic and transparent process, including consultation with relevant stakeholders [14,30,31]. One example of CGs with a high priority for development was the WHO rapid advice CG for the therapeutic management of infectious disease outbreaks published in 2007 [14,32,33].

2.3.2. Definition of the purpose and objectives

The CG panel need to define the population and setting that will be covered in the CG, describe what interventions, diagnostic tests, or what information regarding cost-effectiveness, prognosis, or values of preferences will be considered [15,34].

2.3.3. Panel composition

The CG panel should consist of professionals with varying disciplines, including healthcare professionals with clinical expertise and methodologists qualified in CG development. In addition, active participation of patients or patient representatives is encouraged [35,36]. The conflicts of interest (CoI) of all panel members, including employment or reimbursement from the public or private sector [37,38], professional loyalties [39], or intellectual CoI [40,41], should be disclosed [42-44].

2.3.4. Formulation of the clinical questions

The clinical questions of a CG can concern the effectiveness of an intervention, the diagnostic test accuracy of a test modality, cost-effectiveness, prognosis, patients' values and preferences, or epidemiology. Structured clinical questions are commonly constructed using a PICO format (i.e. P: population/patient, I: intervention/indicator: C: comparator/control, and O: outcomes) [45]. Formulating the clinical questions *a priori* facilitates setting the boundaries of the development process and provides a framework for designing the literature searches. It also informs the planning and completion of the evidence review while guiding the development of recommendations [15].

2.3.5. Systematic search, evidence selection, risk of bias assessment, and evidence synthesis

A SR should be performed to examine the evidence about different aspects of the clinical question of interest (e.g. therapeutic effectiveness, diagnostic test accuracy, cost-effectiveness, values and preferences). This should include a search strategy with a combination of text-words and subject headings (e.g. Medical Subject Headings (MeSH) terms, EMtree) [46,47], covering multiple literature databases (e.g. MEDLINE, EMBASE, or Cochrane library) [48,49]. Consequently, the identified references should be screened on their relevance to the clinical question by at least two reviewers [10].

After all relevant evidence has been identified, the quality of evidence (certainty or confidence in the evidence) should be assessed. The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach is a systematic and explicit methodology to achieve this goal [50-52]. The GRADE approach allows for four levels of quality of evidence assessment, ranging from high to very low quality of evidence. There are five criteria for downgrading the quality of evidence: 1) risk of bias [53], 2) inconsistency [54], 3) indirectness [55], 4) imprecision [56], and 5) publication bias [57]. Factors that allow rating up the quality of evidence include the strength of association in observational studies, existence of a dose-response gradient, or consideration of potential confounding factors [58].

2.3.6. Formulation of recommendations

Heterogeneity has been identified among CG institutions regarding the methodology for formulating recommendations [59]. Nevertheless, recommendations should be formulated taking into account the results of the search strategy, the evidence profile, the quantitative effect estimates (if applicable), and the level of evidence [60]. Recently the GRADE working group has developed Evidence to Decision (EtD) frameworks to help CG panels in using the evidence profile in a structured and transparent way to inform decisions in the context of

clinical recommendations. The EtD framework might facilitate the CG panel to ensure consideration of key criteria that determine the strength and direction of the recommendation, including the priority of the problem, resource requirements of the intervention, impact on health equity, and feasibility of implementation of the intervention [61,62].

2.3.7. External review

An external multidisciplinary group of stakeholders should be invited to provide feedback on the content of the CG, including all recommendations and the used methodology. The external review aims to search stakeholder views on factors that will foster or hinder CG implementation and potential equity impact [15].

2.4 Clinical guidelines updating

Due to the fact that new studies emerge on a regular basis [4], CGs should be kept up-to-date to ensure the CGs validity. Updating CGs is considered an essential component of high-quality CGs [11,12]. The goal of the updating process is to minimise the time-gap between the publication of new evidence and its translation into clinical practice by a systematic approach for identifying and assessing new evidence that was not included in the previous version of the CG [63]. However, as opposed to the field of developing *de novo* CGs, little empirical research has been conducted on the updating process of CGs [63]. An urgent need for standards and methodology for the CG updating process that could be used by international CG institutions has been identified by previous research [64].

2.4.1 Time of validity

Time of validity is defined as the time-interval between the publication of a CG and the identification of at least one signal for updating [65]. The signals for updating include, for

example, the emergence of new scientific knowledge regarding new treatment options or diagnostic tests modalities, changes in the cost-effectiveness, or changes in the patients' values and preferences [15,66,67].

Several studies have investigated the time-interval of the validity of CGs or recommendations [67-71]. The study by Shekelle *et al.* (2001), evaluating the time of validity of 17 CGs from the Agency for Healthcare Research and Quality, found that 90% of the CGs were still valid after 3.6 years while 50% were valid after 5.8 years [71]. Similarly, in the study of Alderson *et al.* (2014), with 134 CGs from NICE, 86% of the CGs were still valid three years after publication [68]. Regarding the time of validity of recommendations, Lyratzopoulos *et al.* (2012) found a mean period of validity of 5.3 years for 11 CGs [70], whereas Martínez García *et al.* (2014), including 113 recommendations from 4 CGs, found that 81.3% (92/113) of the recommendations were still valid three years after publication [67].

2.4.2 Clinical guidelines updating process

The CG updating process should be triggered once a signal for updating has been identified. The updating process of CGs is defined as "*an iterative set of processes with a systematic and explicit methodology that involves identifying and reviewing new evidence that had not been included in the current version of a CG*" [65]. The updating process involves 1) prioritising which CG, recommendation, or clinical question to update firstly, 2) identifying new relevant evidence; 3) assessing whether the new evidence has an impact on the current recommendations, and; 4) reviewing and, if necessary, modifying recommendations [63,65,72] (**Figure 3**).

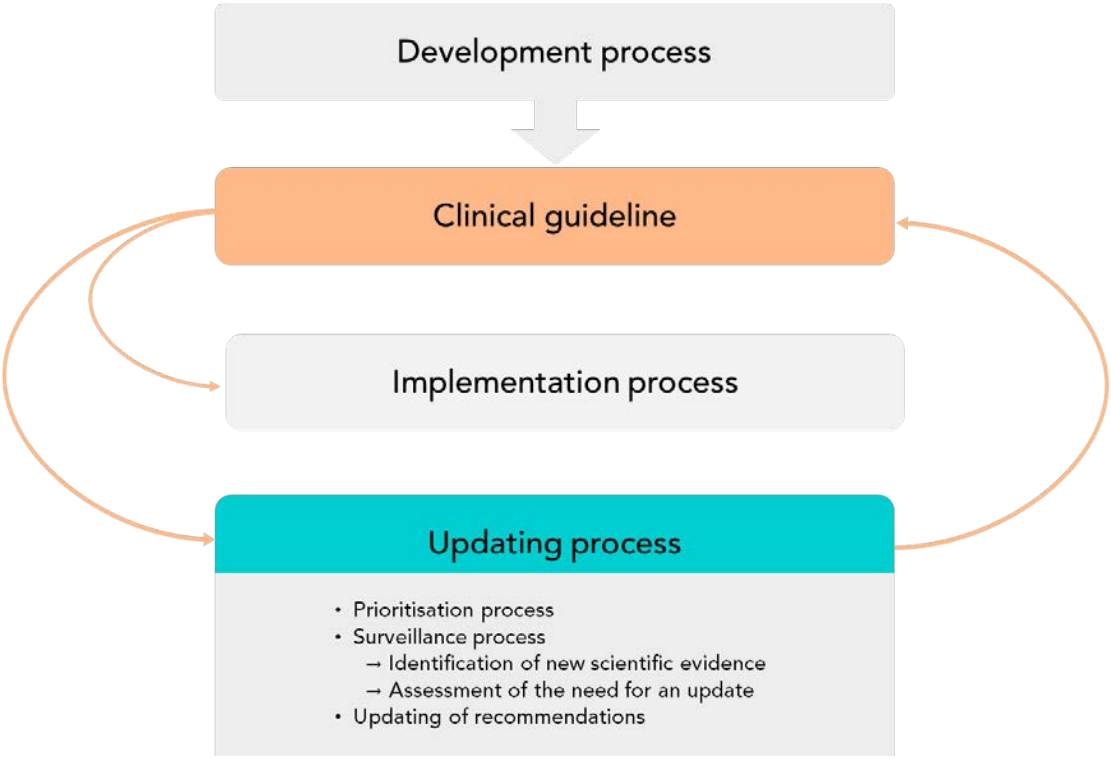


Figure 3: The updating process of clinical guidelines.

2.4.2.1 Prioritisation process

Similar to the CG development process, the aim of the prioritisation process in updating CGs is to ensure that the resources are directed to the CGs, recommendations, or clinical questions that are most relevant to stakeholders and have the highest priority for updating [73,74]. The prioritisation process needs to consider the volume of new research and the balance between updating and developing CGs *de novo* within the portfolio of a CG institution [75]. Prioritisation can be considered at the portfolio level, i.e. deciding which CGs should be updated over others [73], or at CG level, i.e. deciding which sections, clinical questions, and/or recommendations within a CG should be prioritised for updating [74,76,77]. A variety of prioritisation approaches in updating CGs were identified by the SR of Martínez García *et al.* (2017) [74]. These generally involve a pragmatic assessment (using a questionnaire with different prioritisation criteria) or a formal assessment (using a step-by-step algorithm including literature searches to identify new evidence), followed by a classification

of the priority to update assigned to a CG or section within a CG (ranging from low to high priority for updating) [74].

2.4.2.2 Surveillance process

2.4.2.2.1 Identification of new scientific evidence

After the prioritisation process, a systematic search needs to be conducted to identify new evidence after the publication of the previous version of the CG. While often the original search strategy is used again [63], an organisation might decide to apply a restrictive search strategy, including a combination of MeSH terms and text-words from the original systematic search. Restrictive searches are considered an efficient and feasible method that is less resource-intensive than adopting the original search strategy [78]. The study of Martínez García *et al.* (2015), including 4 CGs with 249 recommendations, concluded that with a restrictive search strategy, 68% less references are identified, however, 90% of the key recommendations (i.e. relevant references that might trigger an update due to their impact on the PICO, quality of evidence, the direction of the recommendation, or strength of the recommendation) were still identified [78]. Furthermore, Gartlehener *et al.* (2004) compared a comprehensive search strategy versus the restrictive approach of Shekelle *et al.* (2001) in which SRs, editorials, CGs, and commentaries are sought in high-impact journals [71,79]. The restrictive approach proved to be an efficient and acceptable method to assess the need for an update [79].

2.4.2.2.2 Assessment of the need for an update

The impact of the newly identified evidence on the validity of the recommendations should be assessed [67,71,73]. Whenever new relevant evidence has been found, including key references that might induce significant changes in the clinical questions (i.e. PICO question) or in the formulation of the recommendations (e.g. due to changes in the quality of evidence, balance between benefit and harms, values and preferences, or use of resources), the CG

panel might confirm the need for an update [67]. The study of Martínez García *et al.* (2014) proposed a nine-stage strategy to assess the validity of recommendations, in which the first five stages include a systematic search strategy, survey with clinical experts, and evidence selection. The last four stages consist of evaluating the impact of the new references on the recommendation in relation to the PICO question, quality of evidence, or direction and strength of the recommendation. Consequently, the recommendations should be classified as either in need of updating or, alternatively, the CG panel can decide to confirm the CGs validity [67].

2.4.2.3 Updating process

Although no gold standard exists for the updating process of CGs, it shares several aspects with the methodology of *de novo* development (i.e. systematic literature search, evidence selection, and data extraction). In the updating process, the CG panel might decide to introduce changes to the clinical questions (i.e. add new, modify previous, or delete previous PICO questions) or recommendations (i.e. add new, delete, or modify existing recommendations). Similarly, in the updating process, the CG panel can decide to modify the direction and strength of the recommendations, as well as the quality of evidence, balance between benefit and harms, values and preferences, and use of resources [63,67,80,81]. The updating process is also an opportunity to incorporate changes in the methodology used or improvements in the edition of the manuscript (e.g. different layout or linguistic modifications).

2.4.2.3.1 Living clinical guidelines

The terminology of the concept 'living CGs' is heterogeneous, with other terms as 'continuous updating' or 'dynamic CGs' being used commonly [82]. Continuous updating suggests a more frequent and repeating assessment of new evidence compared to traditional updating, with a fixed time-interval (i.e. updating every three or five years) [82]. Recently, Akl *et al.* (2017)

suggests that for the successful implementation of 'living CGs', the unit to update needs to be individual recommendations, as opposed to the whole CG. Therefore, 'living CGs' are based on the perception of updating individual recommendations as soon as possible after new relevant evidence becomes available [83].

The methodology of 'living CGs' has been evaluated in several studies [84-86]. A continuous surveillance and updating strategy, including regular searches of peer-reviewed literature and meeting proceedings, interpretation of the new evidence, revision of the recommendations, and evidence alerts in oncology was assessed by Johnston *et al.* (2003). Although the continuous updating process was considered feasible, it was resource intensive [84]. Additionally, a panel of CGs methodology experts published recommendations for a continuous strategy to keep the CGs of the Kidney Disease: Improving Global Outcomes (KDIGO) current. Integrated electronic platforms were recommended to facilitate the dynamic updating strategy [85]. Finally, the study of Martínez García *et al.* (2017) concluded that continuous and restricted literature search strategies are a feasible approach requiring, nonetheless, long-term substantial resources [86].

2.5 Reporting in evidence-based medicine

Published articles in scientific journals are considered the main output of research and the primary means of sharing new knowledge. Therefore, reporting the methodology and subsequent results is crucial to inform and allow the reader to assess the quality of the study while ensuring accuracy and transparency of the results [87-89]. If the methodology is poorly reported, the dissemination of new knowledge might be ineffective [90]. Suboptimal reporting might limit the applicability of the study and mislead patients and healthcare professionals in implementing the results into clinical practice [91]. There is multiple evidence that the reporting of scientific evidence is suboptimal for different study designs [92-96].

2.5.1 Assessment of the reporting process

Given that adequate reporting is considered important, several instruments are developed to assess the reporting of the methodology and findings of different study designs [97]. Considerable reporting instruments are identified for multiple study designs [97]. The most widely used checklist is CONSORT, the "*Consolidated Standards of Reporting Trials*" for reporting RCTs [98]. Other checklists have been designed for other study designs including the "*Strengthening the Reporting of Observational Studies in Epidemiology*" (STROBE) for observational studies [99], the "*Standards for Reporting of Diagnostic Accuracy Studies*" (STARD) for diagnostic test accuracy studies [100], and the "*Preferred Reporting Items for Systematic Reviews and Meta-analysis*" (PRISMA) for the reporting of SRs [101]. The use of these reporting instruments is associated with improved quality of reporting [102.103].

2.5.2 Tools to evaluate clinical guidelines

There are three tools available for assessing the reporting of the development process of CGs: 1) the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument [34], including the AGREE Reporting Checklist [104]; 2) the Guidelines International Network (G-I-N) McMaster Guideline Development checklist [105]; and 3) the Reporting Items for practice Guidelines in HealThcare (RIGHT) statement [106].

2.5.2.1 AGREE II tool

The AGREE II instrument is developed as a tool to address variability in CGs quality. This instrument assesses the methodological rigour and transparency of the CGs development process [34]. The AGREE II instrument includes 23 items that target various aspects of CG reporting organised in six domains: 1) scope and purpose, 2) stakeholder involvement, 3) rigour of development, 4) clarity of presentation, 5) applicability, and 6) editorial independence. Besides assessing the reporting of the development process of CGs, AGREE II can serve as a methodological strategy when the CG panel is planning the development

process [34]. Recently, the AGREE II items have been reformed in the AGREE Reporting Checklist [104].

2.5.2.2 G-I-N McMaster Guideline Development Checklist

The G-I-N McMaster Guideline Development Checklist (Guidelines 2.0) was recently developed to start, improve, and evaluate the CG development process [105]. This checklist includes a description of all the different CG development process steps [105]. Adherence to the checklist might ensure that the CG panel covers all key items of the development process and increase the likelihood of achieving higher scores when evaluated with other CG quality instruments, such as the AGREE II [105].

2.5.2.3 The RIGHT statement

The Reporting Items for practice Guidelines in HealThcare (RIGHT) statement was recently published and includes 22 items regarding the development process of CGs, ranging from relevant background information, evidence surveillance, recommendation development, quality assurance, CoI disclosure, among others [106]. The RIGHT statement has the potential to assist the CG panel in reporting the development process and, consequently, provide CG users a clear reporting of the development process.

2.5.3 Updating process in the clinical guideline reporting instruments

Although ensuring the CGs validity is considered an important quality standard of CGs [11,12], as far as we know, there is no reporting instrument available that assess the reporting of the updating process and provides guidance in reporting in updated CGs. While the AGREE II, G-I-N McMaster Guideline Development checklist, and the RIGHT statement focus on the development process of *de novo* CGs, they do ignore the specific methodology aspects related to the updating process of CGs [34,105,106]. Although there are some

methodological similarities between developing *de novo* CGs and updating CGs, the updating process should be considered a completely different step in the life-cycle of CGs (**Figure 3**). Where the AGREE II instrument includes one item about the updating process (i.e. *"A procedure for updating the guideline is provided."*) [34], and the G-I-N McMaster Guideline Development checklist includes five items about updating CGs: 1) a procedure for updating, 2) the person responsible for monitoring the literature, 3) conditions that will determine a partial or full update, 4) CG panel responsible for updating, and 5) funding for updating the CG [105], the RIGHT statement does not include any item about the updating process [106].

2.6 Justification

2.6.1 Justification of the research topic of the thesis

Evidence-based medicine is in a constant state of evolution and practice-changing studies are published on a daily basis [4,5]. Using the most recent evidence in healthcare decision-making is considered crucial in improving patient care [1-3]. Therefore, in order to be useful in clinical decision-making, CGs need to remain up-to-date to ensure the validity of the recommendations. Some of the consequences of outdated CGs might be loss of confidence of end users or inadequate or inefficient healthcare practice [29,107,108].

An adequate updating process minimises the time gap between the publication of new relevant evidence and incorporating it in the CG [63]. The updating process of CGs is considered an iterative process with a systematic and rigorous process, which includes the identification and assessment of new evidence, in which the CG should be modified if evidence is identified that is considered to have an impact on the current CG [63,65,72]. Whereas nowadays the field of developing CGs *de novo* is relatively well established on the basis of several methodological advancements, including the development of the GRADE approach [50], the publication of quality standards by both the IOM and G-I-N [11,12], and the availability of several reporting instruments [34,105,106], the field of CG updating has

been lagging behind. This is surprising since in the CG enterprise, the focus has switched from developing CGs *de novo* to keeping existing CGs up-to-date, however, there is scarce guidance available for the updating process and CG institutions might experience troubles keeping their CG portfolio up-to-date, therefore, more investigation and guidance for the updating process is needed [63,64].

2.6.2 Justification of the publications

2.6.2.1 Publication I

There is little research available regarding the optimal methodology to operationalise the updating process of CGs and the practices for updating CGs across institutions [63,64,109,110]. Several institutions responsible for updating CGs, have described their methodology for the updating process in methodological handbooks. However, no systematic summary or synthesis of the guidance of the updating process included in these handbooks has been conducted. We, therefore, conducted a SR to identify and describe the updating process guidance available in the CGs methodological handbooks.

2.6.2.2 Publication II

As described by Djulbegovic and Guyatt (2017), there is a clear need for the development of standards for the reporting of clinical research to improve the field evidence-based medicine [111]. Optimal reporting of methods and results is considered crucial to allow users to assess the certainty of the evidence and to translate the results into practice-changing initiatives [87-89]. Several tools are available for assessing the reporting of the development process of new CGs [34,105,106]. However, no such tool exists for assessing the reporting in CG updating. Given that the updating process requires different methodological considerations and unique communication procedures [112], we developed a reporting checklist for this purpose.

2.6.2.3 Publication III

Reporting the methodology and subsequent findings is crucial to inform and allow the reader to assess the certainty of evidence and poor reporting might interfere adequate dissemination of new knowledge. However, the reporting of the updating process has not been systematically examined and after developing the reporting checklist in the second study, no systematic assessment of the actual reporting of the updating process in a sample of updated contemporary CGs has been conducted to date. These results can provide insight in the current reporting of the updating process and identify potential room for improvement.

2.6.2.4. Appendices

Additionally, six articles that are related to the updating process of CGs are included in the appendix of this thesis: 1) explanation and elaboration article of the reporting instrument (published as an appendix to second study of this paper) [113], 2) an evaluation of two search strategies to identify the need to update [78], 3) a SR about the methods used to prioritise CGs for updating [74], 4) a protocol regarding the development of an instrument for the prioritisation in updating CGs [114], 5) a glossary with the domains, terms, definition, and synonyms related to the updating process of CGs [65], and 6) a cohort study to investigate a continuous surveillance and updating strategy in one CG [86]. Furthermore, the translations of the reporting checklist are included in the appendices: 1) the translation into Spanish, 2) the translation into Chinese, and 3) the translation into Dutch.

3. OBJECTIVES

- To identify and describe the updating guidance available in CG methodological handbooks.
- To develop a checklist for the reporting of updated CGs.
- To assess the completeness of reporting of updated contemporary CGs.

4. METHODS

4.1 Article 1. "Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks."

4.1.1 Design

Systematic review of methodological handbooks that provide guidance on the updating process of CGs.

4.1.2 Information sources and search strategy

We conducted a systematic search in September 2013 in MEDLINE (accessed through PubMed, from 1966 onwards), the G-I-N library, and the National Guidelines Clearinghouse. Additionally, we searched the websites of the institutions that reported to use a methodological handbook in a previous international survey conducted by our research group [64].

4.1.3 Inclusion criteria

We included methodological handbooks that provided guidance on the CG updating process. Handbooks that exclusively report methodologies for developing *de novo* CGs were excluded. No limitations in the language or publication status of the handbooks were applied.

4.1.4 Study selection and data extraction

Two authors independently assessed the eligibility of the identified references by reviewing the titles and abstracts, followed by an evaluation of the full-text articles to determine the eligibility. Disagreements were resolved by consensus, and if necessary, with the help of a third author. We extracted the following details of the included handbooks: the main

characteristics of the handbooks and CG institutions; the group responsible; the strategy for identifying new evidence, the methodology for assessing the need for an update; and the methodology of the updating process, including the literature search, evidence selection, evidence assessment, evidence synthesis, external review, and the dissemination of the updated CG.

4.1.5 Data analysis

We applied descriptive statistics to analyse the extracted data, including absolute frequencies and proportions. We analysed the data using SPSS, version 18.0 (SPSS INC., Chicago, IL, USA).

4.2 Article 2: "Reporting Items for Updated Clinical Guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp)."

4.2.1 Design

We performed a multi-step development process including: 1) development panel selection; 2) generation of an initial checklist; 3) optimisation of the checklist, including an assessment of contemporary updated CGs, semi-structured interviews, a Delphi consensus survey, external review with CGs methodologists, and external review with CG users; and 4) approval of the final checklist. The characteristics of the development process are presented in Table 2 (**Table 2**). Before the development process started, a core group, including the main authors, was established to design the protocol and provide operational advice.

Table 2: Characteristics of the multistep development process.

Step	Objective	Study design	Participants	Result
Panel selection	To assemble a panel that can provide expert advice during the development process and participate in the Delphi survey.	We convened the development panel comprising of individuals with relevant experience in CG development and/or SR/CG research methodology.	Invited panel participants were identified based on a review of authors in the CG enterprise, as well as the AGREE trust and G-I-N members.	The development panel was convened.
Generation of the initial checklist	To generate an initial version of the checklist.	We developed an initial list of items, including explanation and examples, through brainstorming and discussion, taking into account: 1) available research evidence regarding updating CGs, 2) the AGREE II instrument, and 3) the core group experience.	The core group.	An initial version of the checklist.
Optimisation of the checklist				

Assessment of existing updated CGs	To verify the used terminology in updated CGs, and to identify missing items as a first step to explore its face validity.	We searched updated CGs that were: 1) developed by CG institutions, 2) published in English or Spanish, and 3) published between 2011 and 2013 in the G-I-N library or the National Guideline Clearinghouse. Two reviewers applied the checklist, solving disagreements by consensus. The core group discussed the results and refined the initial list of items.	The core group.	Optimisation of the checklist.
Semi-structured interviews	To refine the checklist and to identify missing items in the current version of the checklist.	We conducted semi-structured interviews until data saturation was achieved. In each interview, participants were asked about their experiences and challenges in updating CGs. Subsequently, the interviewees were prompted by the interviewer to reflect on the strengths, weaknesses, missing concepts, and redundancies in the current version of the checklist. The core group discussed the results and refined the checklist.	A convenience sample of participants with: 1) experience in updating CG, and 2) fluency in English, were identified by contacting professionals associated to G-I-N or researchers in the CG enterprise.	Optimisation of the checklist.
Delphi consensus survey	To assess the inclusion, comprehensiveness, clarity, and coverage of each item and to identify potentially additional items for the checklist.	We asked the participants of the Delphi survey to rate whether the items should be included in the checklist. Additionally, for each item, participants were asked whether their perceptions of the: 1) completeness, 2) usability, and 3) quality of a CG would be influenced if the item was reported. We included a free text box for suggestions to modify the items, explanations, or examples. We calculated the median score for the inclusion, completeness, usability, and quality of each item and classified into: 1) items with a median score of 0 to 3 points, which were excluded without further evaluation; 2) items with a median score of 4 to 5 points or with substantial comments that required important revision, which were modified and further tested; and 3) items with a median	Development panel.	Optimisation of the checklist.

		score of 6 to 7 points and without substantial comments, which were included and not evaluated further in the following rounds. One reviewer analysed the results and suggested potential solutions. The core group discussed the results and refined the list of items accordingly. We continued the Delphi survey with additional rounds until consensus regarding the inclusion or exclusion of the items was reached, and no more relevant comments were provided.		
External review	To explore the usability of the checklist and the importance of reporting the included items.	We conducted a survey using a seven-point scale and asked participants to rate the usability of each item, and their confidence in an updated CG if the item was reported. A free text option was included for suggestions to modify the items, explanation, or examples. We calculated the median score for usability and confidence for each item. One reviewer analysed the quantitative and qualitative results, and suggested potential solutions. The core group discussed the results and potential solutions and refined the list of items accordingly.	CG methodologists who had experience in updating CGs. All G-I-N institutional members were invited to participate in the external review.	Optimisation of the checklist.
External review	To explore the usability of the checklist and the importance of reporting the included items.	We conducted semi-structured interviews until data saturation. For each interview, we asked the participants whether reporting the item in an updated CG would increase their confidence in the CG and prompted the participants to consider missing concepts, redundancy, and usability of the checklist. The core group discussed the results and refined the list of items.	CG users who were: 1) health care professionals and 2) located in Canada, Spain, or the Netherlands. We identified the participants with the help of the panel members.	Optimisation of the checklist.
Approval of the final checklist	To finalise and approve the checklist.	In this workshop, we asked the participants whether the checklist was deemed adequate for the assessment of the reporting in updated CGs. Additionally, we asked the participants to give an overall impression of the checklist. The core	Participants of a workshop at the 2015 G-I-N conference.	Final version of the checklist.

	group discussed the results and agreed on the final list of items.		

4.3 Article 3: "Updated clinical guidelines experience major reporting limitations".

4.3.1 Design

Systematic assessment of the reporting in updated CGs.

4.3.2 Information sources and search strategy

We conducted a systematic search in August 2016 in MEDLINE (accessed through PubMed), the G-I-N library, and the National Guidelines Clearinghouse.

4.3.3 Inclusion criteria

We included all updated CGs published in 2015 which met the following criteria: 1) developed by a CG institution, 2) included a search strategy using at least one bibliographic database, 3) reported at least one recommendation, 4) is an updated version of a CG (including a reference to a previous version of the CG), and 5) published in English.

4.3.4 Study selection and data extraction

Two reviewers independently screened the titles and abstracts to identify potentially eligible CGs for inclusion. We obtained the full-text articles of the included references for further assessment. Disagreements were initially solved by consensus and, if necessary, with the help of a third reviewer. Three reviewers independently evaluated each CG with the reporting checklist developed in article 2, and whenever the included CGs referred to supplemental documents (e.g. methodological manuals or appendices), these documents were also reviewed.

4.3.5 Data analysis

We calculated the median (including range) of the items, domains, and overall scores and converted the domain and overall scores to a 10-point scale (**Table 3**). To identify potential predictors, we conducted a multiple linear regression to test whether the overall score (dependent variable) differed between the CG institutions country, type of organisation, objective of the CG, and CG topic (independent variables). Finally, we calculated the intraclass correlation coefficient (ICC) with its 95% confidence interval (CI), as an indicator of overall agreement between the three reviewers for each item according to the scale proposed by Landis and Koch [115]. We accepted a p-value of less than 0.05 as significant and performed all analyses using SPSS version 22.0 (SPSS Inc., Chicago, Illinois, United States).

Table 3: Calculation of the item, domain, and overall scores.

Score	Calculation	Numeric variables
Item	$\textbf{Item} = \sum (Item\ n = yes)\ CGs$	Absolute frequencies and proportions
Domain*	$Domain\ n = \frac{\sum (Item\ n = yes)\ in\ domain\ n}{n\ items\ in\ domain} \times \mathbf{10}$	Median and range
Overall*	$Overall = \frac{\sum (Item\ n = yes)\ in\ CheckUp}{n\ of\ items\ of\ CheckUp\ assessed} \times \mathbf{10}$	Median and range

* 10-point scale (10 as the best possible score).

5. RESULTS

5.1 Article 1: "Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks."

5.1.1 Summary of the results.

In total we included 35 handbooks, mostly developed by public European institutions [116]. The majority of the included handbooks do not provide adequate updating guidance. Where most handbooks focus on the development process of *de novo* CGs, we identified one handbooks that solely described the updating process. Approximately one third of the handbooks describe the CG panel responsible for updating the CGs. Guidance for identifying new relevant evidence is generally poorly described by the included handbooks, with only 31% (n=11) of the handbooks providing guidance on how to identify new evidence. The time-interval between the development of CGs and updating (or two updates) is described in 71% (n=25) of the handbooks, with two to three years being the most frequent period of time. The methodology for assessing the need for an update is described in 23% (n=8) of the included handbooks. Similarly, the majority of the included handbooks do not provide guidance on how to conduct literature searches, evidence selection, assessment, synthesis, or the external review of their guidelines that are updated. Solely three (9%) handbooks provide guidance on how to report and disseminate an updated CG [116].

5.1.2 Copy of article 1.

IMPLEMENTATION SCIENCE

SYSTEMATIC REVIEW Open Access

Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks

Robin WM Vernooij[1,2], Andrea Juliana Sanabria[1], Ivan Solà[1], Pablo Alonso-Coello[1*] and Laura Martínez García[1]

Abstract

Background: Updating clinical practice guidelines (CPGs) is a crucial process for maintaining the validity of recommendations. Methodological handbooks should provide guidance on both developing and updating CPGs. However, little is known about the updating guidance provided by these handbooks.

Methods: We conducted a systematic review to identify and describe the updating guidance provided by CPG methodological handbooks and included handbooks that provide updating guidance for CPGs. We searched in the Guidelines International Network library, US National Guidelines Clearinghouse and MEDLINE (PubMed) from 1966 to September 2013. Two authors independently selected the handbooks and extracted the data. We used descriptive statistics to analyze the extracted data and conducted a narrative synthesis.

Results: We included 35 handbooks. Most handbooks (97.1%) focus mainly on developing CPGs, including variable degrees of information about updating. Guidance on identifying new evidence and the methodology of assessing the need for an update is described in 11 (31.4%) and eight handbooks (22.8%), respectively. The period of time between two updates is described in 25 handbooks (71.4%), two to three years being the most frequent (40.0%). The majority of handbooks do not provide guidance for the literature search, evidence selection, assessment, synthesis, and external review of the updating process.

Conclusions: Guidance for updating CPGs is poorly described in methodological handbooks. This guidance should be more rigorous and explicit. This could lead to a more optimal updating process, and, ultimately to valid trustworthy guidelines.

Keywords: Clinical practice guidelines, Evidence-based medicine, Handbooks, Methodology, Systematic review

Background

Clinical practice guidelines (CPGs) intend to patient care by providing recommendations about the benefits and downsides of best practice in healthcare [1]. If adequately implemented, CPGs have the potential of reducing variability and translating scientific research into clinical practice and consequently improve the quality and safety of healthcare [2-4].

However, scientific knowledge is in constant change; therefore CPGs need to be updated regularly to maintain validity [5]. The obsolescence of a CPG might occur because of new scientific research, including the development of new technologies in treatment and diagnosis alternatives, economic differences, or changes in values and preferences [6,7]. Generally, an updating process consists of three components: the identification of new evidence, the assessment of the need to update, and the formulation of new or modified recommendations [5,8-11]. Some authors suggest that an update is generally required after three to five years; however, little research has been undertaken so far [8,12,13].

Several institutions responsible for developing CPGs drafted their own methodological handbooks including methodology for developing and updating their CPGs. Some of these handbooks are very influential and often used in smaller organizations [6,14]. Even though the methodology developed greatly over the last years, the quality of CPGs is lagging behind [1,15,16]. A lack of compliance with state of the art methodology for developing CPGs has been found, and hence the methodological quality of CPGs remained very similar over the

* Correspondence: palonso@santpau.cat
[1]Iberoamerican Cochrane Centre, Institute of Biomedical Research (IIB Sant Pau), C/ Sant Antoni Maria Claret 167, Barcelona 08025, Spain
Full list of author information is available at the end of the article

last two decades [17,18]. Little is known about the guidance for updating CPGs included in these handbooks [19,20]. Therefore, we systematically reviewed CPGs methodological handbooks to identify and describe the methodological guidance about updating.

Methods

Search strategy

We conducted a systematic search in September 2013 in MEDLINE (via PubMed, from 1966 onwards), using a combination of free text terms (Clinical Practice Guidelines, Clinical Guidelines, Guidelines, Methodolog*, Handbook*). The search strategy is available as supplementary data (Additional file 1). In addition, we searched: the database of the Guidelines International Network (http://www.g-i-n.net); the US National Guidelines Clearinghouse database (http://www.guidelines.gov); and the website of institutions that reported to use a methodological handbook in a previous international survey conducted by our group [12]. If necessary, we contacted organizations to obtain the handbooks.

Eligibility criteria

We included methodological handbooks that provide guidance on the updating process of CPGs. Handbooks that exclusively report methodologies for developing *de novo* guidelines were excluded. We included handbooks regardless of their language or publication status. When necessary, the handbook was translated.

Study selection

Two authors (RV, AJS) independently selected potential handbooks by reviewing titles and abstracts, and finally full text for a more detailed evaluation. Disagreements were initially resolved by consensus, and if necessary, with the help of a third author (PA-C).

Data extraction

Based on our previous experiences concerning updating, including an international survey [12] a systematic review [8] and additional relevant literature [5,6,9-11,14] we developed, reviewed, and piloted iteratively a case report form (CRF). After consensus, the following items are included in the CRF: characteristics of the handbook and institution, group responsible for updating CPGs, strategy for identifying new evidence, methodology for assessing the need for an update, methods for the literature search, evidence selection, evidence assessment, evidence synthesis, external review, and for the edition and dissemination of the updated CPG. The CRF can be made available upon request.

Two authors (RV, AJS) extracted independently the data of the handbooks accepted for inclusion. Disagreements were initially resolved by consensus, and if necessary, with the help of a third author (PA-C). While extracting the data, we considered a strategy to be specific if the handbook included a detailed methodology, enabling the reader to conduct the suggested strategy. We considered a non-specific strategy if not enough methodological guidance is provided to facilitate an adequate approach.

Data analysis

We used descriptive statistics to analyze the extracted data. We calculated absolute frequencies and proportions for all items. In addition, we conducted a narrative synthesis. Data analysis was performed using SPSS statistical software, version 18.0 (SPSS INC., Chicago, IL, USA). By consensus of two authors (RV, AJS), we collected relevant quotations within the themes included in the handbooks and provide these in the free text area.

Results

Handbooks selection

We screened the titles and abstracts of 1,992 references (Figure 1). We selected 94 articles for full-text review. Thirty-eight articles were excluded because they were not methodological handbooks. Additionally, ten handbooks were excluded because they exclusively focused on developing *de novo* CPGs. We could not locate eight articles and one article was a summary of an included handbook. Two handbooks were excluded because a more recent version was included. Additional file 2 provides an overview of the excluded documents. Finally, we included thirty-five handbooks (Additional file 3) [5,6,14,21-52].

Handbooks characteristics

In total, 48.6% of the included handbooks are developed by institutions based in Europe [5,6,14,21-34] mostly being public institutions (57.1%) (Table 1) [5,6,14,22-26,28,31,35-43]. One handbook (2.9%) addresses specifically the methodology of updating CPGs [5]; the others (97.1%) focus mainly on developing *de novo* CPGs, and include variable degrees of information about updating [6,14,21-52]. Fourteen handbooks (40.0%) are published between 2005 and 2010 [5,21,23,26,30,32,34,39,40,43,44,46,48,50].

Updating group

The persons responsible for updating the CPG are specified in twelve handbooks (34.3%). Seven handbooks (20.0%) state that the updating group should have a similar structure to the group that contributed to developing the CPG [6,14,23,30,37,44,45]. Four handbooks (11.4%) state that the group, responsible for updating the CPG, should be tailored to the new scope of the guideline [5,38,39,41].

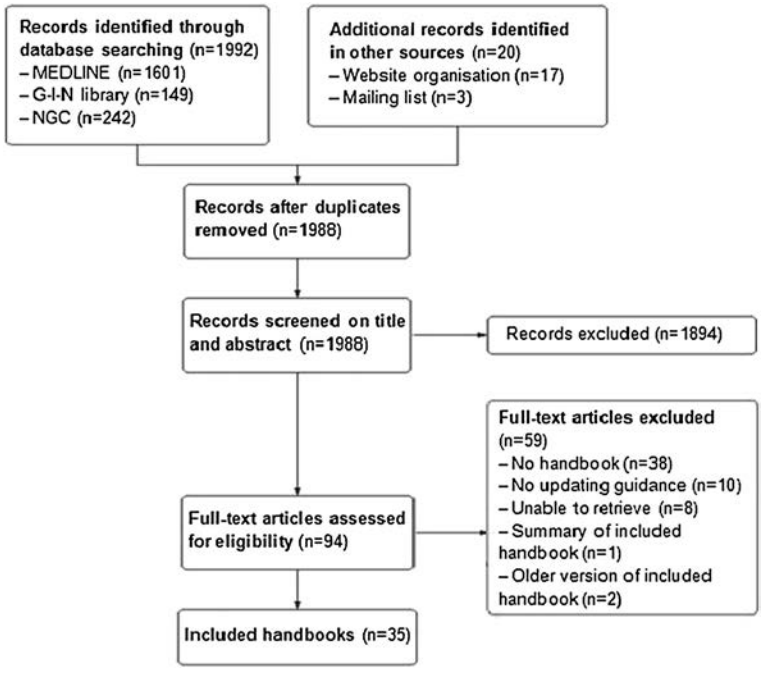


Figure 1 Flow chart of the screening literature process.

Time between updates

Twenty-five (71.4%) of the included handbooks recommend a time frame between publishing a CPG and commencing an updating process (Table 2), with two to three years being the most frequently recommended (40.0%) [5,6,14,22,27,28,30-32,37,39,41,45,46]. Furthermore, three handbooks (8.6%) suggest a time frame of less than one year [33,34,44], and eight handbooks (22.9%) include a four to five year time frame [24,36,38,42,43,47-49].

Identification of new relevant evidence

Eleven handbooks (31.4%) provide guidance on how to identify new relevant evidence. Of these eleven handbooks, six (17.1%) suggest using opinions or experiences from experts, users, or members of the original development group for identifying new relevant evidence [5,14,23,37,43,46]. Five handbooks (14.3%) provide guidance on conducting limited searches to identify new relevant evidence [5,37-39,47]. Furthermore, two handbooks (5.7%) propose the editorial board to have periodic meetings to discuss topics with experts [32,33]. One handbook (2.9%) suggests collecting alerts to identify newly published articles [5]. Externally reviewing the CPG by experts, who were not involved in developing the CPGs, is recommended by one handbook (2.9%) [47]. Two other handbooks (5.7%) provide a 'non-specific strategy' and only emphasize the importance of identifying new relevant evidence (Table 2) [23,28]. Figure 2 shows examples of relevant passages included in the handbooks.

Assessment of the need for an update

The methodology of assessing the need for an update is described in eight handbooks (22.8%). Six of them (17.1%) give guidance on how to assess the importance and relevance of the new evidence, the disagreement between the new evidence and current recommendations, and whether the new knowledge is not yet included [5,6,23,38,43,49]. Two handbooks (5.7%) recommend expert judgment to assess the need for an update [38,40]. Producing and regularly updating evidence summaries and assessing the need for an update with these summaries are described in one handbook (2.9%) (Figure 2) [32].

Updating recommendations

Eight handbooks (22.9%) provide guidance on what type of update is required in specific situations, by making a distinction between partial or full updates (Table 2) [5,6,14, 33,37,38,43,44].

Guidance for conducting a literature search strategy is included in seventeen handbooks (48.6%). Eight of them (22.8%) include guidance to adjust the original search strategy [5,6,14,24,26,27,37,43]. Four handbooks (11.4%) provide guidance on what kind of evidence to search for, including evidence based guidelines, health technology assessments, systematic reviews, and randomized controlled trials [14,27,38,41]. Two handbooks (5.7%) recommend to include a medical librarian or research officer in the team to conduct the literature searches [41,48]. Using multiple databases, *e.g.*, MEDLINE and Cochrane Library, in the search strategy is recommended by two handbooks

Table 1 Characteristics of institutions and handbooks

Institution characteristics	n	(%)
Continent		
Europe	17	48.6
North America	12	34.3
Oceania	4	11.4
International	2	5.7
Type of organization		
Public institution	20	57.1
Scientific society	9	25.7
Private organism	3	8.6
Other (Federal institute, NGO)	3	8.6
Number of years developing guidelines		
≤10 years	10	28.6
10 – 20 years	19	54.3
>20 years	6	17.1
Number of guidelines published		
≤5 per year	22	62.9
>5 per year	8	22.9
Unknown	5	14.3
Handbook characteristics		
Type of handbook		
Development CPG handbook	34	97.1
Update CPG handbook	1	2.9
Publication date		
Before the year 2004	8	22.9
Between 2005 – 2010	14	40.0
Between 2011 – 2013	8	22.9
Unknown	5	14.3

Table 2 Guidance reported in the included handbooks

Group responsible for updating CPG	n	(%)
Are the participants in the updating group specified?		
Yes	12	34.3
No	23	65.7
What members do the updating group consist of?		
Similar to the development team	7	20.0
Updating group specifically defined	4	11.4
Not defined	24	68.6
Identification of new evidence		
Time frame for updating		
≤1 year	3	8.6
2-3 years	14	40.0
4-5 years	8	22.9
No specific time frame indicated	10	28.6
Identification of new evidence		
Specific strategy	9	25.7
Non specific strategy	2	5.7
Not defined	24	68.6
Assessment of the need for an update		
Assessment of the need for an update		
Specific strategy	8	22.8
Not defined	27	77.1
Updating strategy		
Distinction between different updates (partial / full)		
Yes	8	22.9
No	27	77.1
Literature search		
Specific strategy	11	31.4
Similar to the development process	6	17.1
No strategy defined	18	51.4
Evidence selection		
Specific strategy	3	8.6
Similar to the development process	8	22.9
Not defined	24	68.6
Evidence assessment		
Specific strategy	5	14.3
Similar to the development process	8	22.9
Not defined	22	62.9
Evidence synthesis		
Specific strategy	3	8.6
Similar to the development process	5	14.3
Not defined	27	77.1

(5.7%) [41,43]. Furthermore, six handbooks (17.1%) suggest using the original strategy used for the development of the original guideline (Table 2, Figure 2) [23,28,34,40,44,50].

Eleven handbooks (31.4%) provide guidance for selecting adequate evidence in the updating process. Three handbooks (8.6%) provide specific guidance on how to discard irrelevant information [5,14,44]. Eight handbooks (22.9%) refer the reader to the development process for guidance on evidence selection [6,27,28,34,37,38,48,50].

Guidance for evidence assessment is provided in thirteen handbooks (37.1%). The assessment of the available evidence on the consistency, directness, validity or reliability is described in four handbooks (11.4%) [14,37,43,48]. Using critical appraisal frameworks, like OstFLCritica, is recommended in one handbook (2.9%) (Figure 2) [5]. Eight handbooks (22.9%) recommend the same original development strategy [6,23,27,28,34,38,44,50].

Table 2 Guidance reported in the included handbooks *(Continued)*

External review		
Specific strategy	5	14.3
Similar to development process	6	17.1
Non specific strategy	2	5.7
Not defined	22	62.9
Edition and dissemination		
Indication of changes		
Specific strategy	5	14.3
Not defined	30	85.7
Dissemination of the updated CPG		
Specific strategy	3	8.6
Not defined	32	91.4

Similarly, guidance for the evidence synthesis is described in eight handbooks (22.9%). Three handbooks (8.6%) recommend producing evidence tables including the characteristics of included studies, quality of randomized trials, results for continuous outcomes, and results for dichotomous outcomes [14,43,48]. Moreover, five handbooks (14.3%) direct the reader to the section with guidance for evidence synthesis used for developing *de novo* CPGs [5,6,34,44,50].

Guidance for an external review of the updated CPG is described in thirteen handbooks (37.1%). Five handbooks (14.3%) describe the process of external reviewing the updated CPG by multiple external reviewers [37,43,45,47,48]. Furthermore, two handbooks (5.7%) provides 'non-specific guidance' for conducting an external review of the updated CPG [28,38]. Six handbooks (17.1%) refer to the guidance described in the section of developing *de novo* CPGs [5,6,27,34,44,50].

Edition and dissemination

Two handbooks (5.7%) suggest to post a notification on the website of the institution whenever the need for an update is confirmed [28,29]. Five handbooks (14.3%) include a specific strategy for indicating the changes made in the update (Table 2, Figure 2). These handbooks recommended actions to identify the main changes in the update without any difficulty, including a table of updated evidence, summary reports, or highlight the updated parts in the text with a red font [5,32,33,37,47].

Three handbooks (8.6%) provide guidance on how to publish and disseminate the updated CPG. All three of them include methods to disseminate the updated CPG as widely as possible by publishing in relevant indexed journals [5], disseminate within the patient organization of the specific disease [48], or working together with

Identification of relevant new evidence:

- *"¨All comments received on published SIGN guidelines, or information on important new evidence in the field, or evidence of impacts on equality groups are fed back to the guideline development group, either for immediate response or for more detailed consideration on review of the guideline."*[14]
- *¨The editorial board meets once a month, and at every meeting, one speciality or a group of topics are discussed with 1-3 top experts on the field invited to attend.¨*[32]

Assessment of the need for an update:

- *¨The editorial team produces and updates evidence summaries continuously, and whenever the evidence summaries give rise to updates to the guidelines, the guidelines are updated.¨*[32]
- *¨At this point, the group should determine the extent of the update required. In addition, the composition of the group should be reassessed based on the planned extent of the update.¨*[38]

Updating process:

- *¨An update search is carried out looking for evidence based guidelines, HTAs, and systematic reviews produced since publication of the last version of a guideline. These searches are based on the key questions and search strategies used in the original guidelines.¨*[14]
- *¨Use of critical appraisal files like: OstFLCrítica, the free-access critical appraisal files IT application of Osteba.¨*[5]

Edition and dissemination:

- *"In EBM Guidelines, updated content appears in red font for 6 months after the update was made."*[32]
- "*Publish the changes using different methods: publishing in relevant indexed journals and/or the journals of the societies involved, indexing in their own website or in other international sources like the National Guideline Clearinghouse (NGC)."*[5]

Figure 2 Box of relevant comments.

public and private partners to reach specific groups and individuals [43].

Discussion

We systematically reviewed 35 methodological handbooks that provide some type of guidance on the updating process of CPGs. Our results show that overall the updating guidance is poorly described. Crucial elements in identifying new evidence, the assessment for the need for an update and the updating strategy itself, are generally lacking or include solely a reference to the development process. Our findings resonate with previous findings that suggest that there is a need for rigorous international guidance for updating CPGs [8,14].

Figure 3 summarizes an updating process framework for CPGs based on a previous systematic review from our group and the results of the present study [8]. The process of updating a CPG starts with assembling a group responsible for updating the CPG. However, we found that the majority of the institutions (65.7%) do not include any information about this first step. There is no clear consensus on who should participate in an updating process and, consequently different organizations use different strategies, depending on the characteristics of the organization and type of update. An updating working group, should consist of individuals with a background in methodology and experts in the field of interest, just as the original guideline group [5]. New developments in the clinical area, such as new technologies, might require including additional members with different expertise.

The actual updating process starts with identifying new relevant evidence. Currently, the period between the last publication of the CPG and starting the updating process (time frame) is frequently determined at the time of publication. The majority of the handbooks (62.9%) include a fixed time frame from two to five years, consistent with the results of previous research by Shekelle et al. [13]. This study including a sample of 17 guidelines, estimated that approximately one-half of the CPGs will be outdated after 5.8 years (95% CI: 5.0 – 6.6), and 10% are obsolete after 3.6 years (95% CI: 2.6 – 4.6) [13]. However, these average estimates can be misleading as CPG deteriorating speed is highly topic-specific, with some fields with rapid developments requiring more frequent surveillance for new evidence than others. Suboptimal time frames are likely to result in guidelines becoming obsolete or inefficient use of resources.

After identifying new relevant evidence, an assessment of the effect of this new evidence should be conducted, determining the need for an update [5,9-11]. We believe that this process is best conceptualized as a two-stage process because these are two independent stages with identifying possible new relevant evidence as first step, and, subsequently, deciding whether the identified evidence this evidence alters the validity of the current

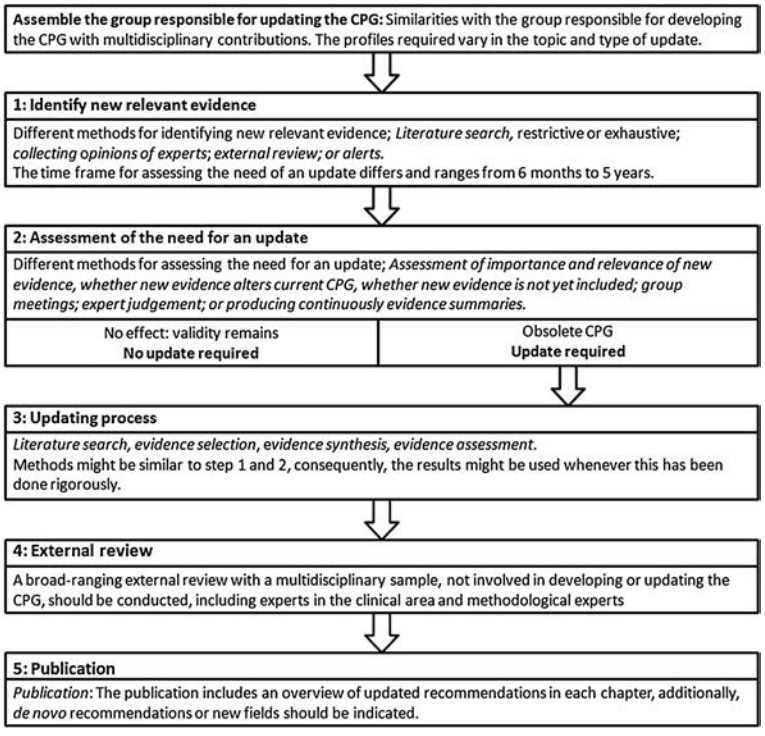


Figure 3 The updating process of CPGs.

recommendations as second step. However, at the moment, formal explicit procedures for assessing the need for an update are not available, with most of the included handbooks (77.1%) not providing explicit methods for assessing the need for an update.

When the need for an update is confirmed, the new evidence has to be incorporated in the current recommendations. However, less than one-half of the included handbooks state specific methods for this process. Previous studies suggest a model of assessing the need for an update using expert opinion, focused literature reviews, and consensus meeting [11,13]. A reference to the development process, often included in the evaluated handbooks, is not enough because the aim of any update should be to incorporate new evidence in the context of previous recommendations. More specific methods should be included in the handbooks.

A further problem is that several institutions use different terminology and consequently bring further confusion. Some institutions use the term 'monitoring' for the identification of new evidence and assessment of the need for an update, often within an abridged time frame [5,14,32,33,37,43,44,52]. In addition, the term 'dynamic updating' and 'living guideline' is used indistinctively, suggesting that CPGs are updated promptly and are always up-to-date [14,40,46]. Nevertheless, none of these handbooks provide guidance for conducting these processes and there is no consensus on when a guideline starts being dynamic or can be considered as a living guideline (Figure 3). We suggest avoiding these terms because it solely reflects the aspect of time between two versions. In Figure 3, we include a proposal regarding consistent terminology. Further research and consensus is needed in the international community about coherent terminology.

Our study is, as far as we know, the first study to examine the guidance about the updating process provided by CPG methodological handbooks. Our work has several strengths. We conducted a systematic and exhaustive search that included main databases, clearinghouses, and several institutions identified by a previous survey [12]. In addition, we contacted several organizations to retrieve non-published handbooks; therefore we believe that we included most of the existing handbooks. We independently performed eligibility and data extraction with a CRF developed and piloted by a group with extensive experience in the field.

Our study, however, might be subject to some limitations. It is possible that, after our extensive literature search, we did not identify all available handbooks because some are not indexed nor published, and only used for in-house purposes. However, unpublished handbooks are likely to be of lower quality. If this is the case, it would imply that we overestimated the quality of the updating guidance, further strengthening our conclusions. Finally, the reported methods in handbooks might not reflect the actual updating in CPGs. However, we believe that this is unlikely given previous results of our international survey with CPG developers [12].

Conclusion

Our work shows that updating guidance included in CPGs methodological handbooks is overall of poor quality. CPGs developers should provide more explicit and rigorous guidance and standardize the terminology used. This could, consequently, lead to a more optimal updating process of CPGs, and ultimately, to valid trustworthy guidelines.

Additional files

Additional file 1: Search strategy (September 16, 2013).

Additional file 2: List of excluded studies after full-text evaluation [in alphabetic order].

Additional file 3: Included handbooks [ordered by organisation].

Abbreviations

CPGs: Clinical practice guidelines; CRF: Case report form.

Competing interests

PA-C is an author of one of the included handbooks. For this reason, other authors completed data extraction for this handbook.

Authors' contribution

Conceiving the review: PA-C, LM. Design of the study: PA-C, LM, RV, AJS. Undertaking searches: IS, RV. Screening and extracting data: RV, AJS. Writing the review: RV, AJS, PA-C. Comment and editing of review drafts: all authors. All authors read and approved the final manuscript.

Authors' information

RV is a doctoral candidate at the Paediatrics, Obstetrics and Gynaecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain.

Funding

Pablo Alonso-Coello is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (CP09/00137). Andrea Juliana Sanabria and Laura Martínez García are funded by a Río Hortega research contract from the Instituto de Salud Carlos III (CM12/00168 and CM11/00035 respectively).

Author details

[1]Iberoamerican Cochrane Centre, Institute of Biomedical Research (IIB Sant Pau), C/ Sant Antoni Maria Claret 167, Barcelona 08025, Spain. [2]Department of Health Sciences, Faculty of Earth and Life Sciences, VU University, Amsterdam, The Netherlands.

Received: 12 June 2013 Accepted: 19 December 2013
Published: 2 January 2014

References

1. IOM (Institute of Medicine): *Clinical Practice Guideline We Can Trust.* Washington, DC: The National Academies Press; 2011.
2. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L, Wensing M, Dijkstra R, Donaldson C: **Effectiveness and efficiency of guideline dissemination and implementation strategies.** *Health Technol Assess* 2004, **8**(6):1–72.

3. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J: **Clinical guidelines: potential benefits, limitations and harms of clinical guidelines.** *BMJ* 1999, **318**(7182):527–530.
4. European Observatory on Health Systems and Policies: *Clinical guidelines for chronic conditions in the European Union.* Geneva: World Health Organisation; 2013.
5. Working Group on CPG Updates: *Updating Clinical Practice Guidelines in the Spanish National Health System: Methodology Handbook.* Madrid: National Plan for the National Health System of the Spanish Ministry for Health and Social Policy; Aragon Health Sciences Institute (I+CS); 2009.
6. National Institute for Health and Clinical Excellence: *The guidelines manual. London: National Institute for Health and Clinical Excellence*; 2012.
7. Shekelle P, Woolf S, Grimshaw JM, Schünemann HJ, Eccles MP: **Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development.** *Implement Sci* 2012, **7**:62.
8. Martínez Garcia L, Arévalo-Rodríguez I, Solà I, Haynes RB, Vandvik PO, Alonso-Coello P, Updating Guidelines Working Group: **Strategies for monitoring and updating clinical practice guidelines: a systematic review.** *Implement Sci* 2012, **7**(1):109.
9. Shekelle P, Eccles MP, Grimshaw JM, Woolf SH: **When should clinical guidelines be updated?** *BMJ* 2001, **323**(7305):155–157.
10. Banzi R, Cinquini M, Liberati A, Moschetti I, Pecoraro V, Tagliabue L, Moja L: **Speed of updating online evidence based point of care summaries: prospective cohort analysis.** *BMJ* 2011, **343**:d5856.
11. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, Whitener L, Voisin CE, Sutton S: **Assessing the need to update prevention guidelines: a comparison of two methods.** *Int J Qual Health Care* 2004, **16**(5):399–406.
12. Alonso-Coello P, Martínez García L, Carrasco Gimeno JM, Solà I, Qureshi S, Burgers JS, Díaz del Campo P, Estrada MD, Gracia J, Mengual J, Rico Iturrioz R, Rotaeche del Campo R, Salcedo-Fernandez F: **The updating of clinical practice guidelines: insights from an international survey.** *Implement Sci* 2011, **6**:107.
13. Shekelle P, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, Woolf SH: **Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated?** *JAMA* 2001, **286**(12):1461–1467.
14. Scottish Intercollegiate Guidelines Network: *SIGN 50: A guideline developer's handbook.* Edinburgh: Scottish Intercollegiate Guidelines Network; 2011.
15. Brouwers M, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L, AGREE Next Steps Consortium: **AGREE II: advancing guideline development, reporting and evaluation in healthcare.** *CMAJ* 2010, **182**:839–842.
16. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A: **GRADE guidelines: a new series of articles in the journal of clinical epidemiology.** *J Clin Epidemiol* 2011, **64**(4):380–382.
17. Kung J, Miller RR, Mackowiak PA: **Failure of clinical practice guidelines to meet institute of medicine standards: two more decades of little, if any, progress.** *Arch Intern Med* 2012, **172**(21):1628–1633.
18. Alonso-Coello P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, Tort S, Bonfill X, Burgers J, Schunemann H: **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):e58.
19. Turner T, Misso M, Harris C, Green S: **Development of evidence-based clinical practice guidelines (CPGs): comparing approaches.** *Implement Sci* 2008, **3**:45.
20. Ansari S, Rashidian A: **Guidelines for guidelines: are they up to the task? A comparative assessment of clinical practice guideline development handbooks.** *PLoS One* 2012, **7**(11):e49864.
21. de Santé HA: *Élaboration de recommandations de bonne pratique.* Saint-Denis: Haute Autorité de Santé; 2010.
22. Centrul Național de Studii Medicina Familiei: *Metodologie elaborarii ghidului de practica.* http://www.ghidurimedicale.ro/index.php?option=com_content&task=view&id=19&Itemid=27. (accessed 1 August 2012).
23. Kwaliteitsinstituut voor de Gezondheidszorg CBO: *Evidence-based Richtlijnontwikkeling Handleiding voor werkgroepleden.* Utrecht: Kwaliteitsinstituut voor de Gezondheidszorg; 2007.
24. Van Der WP, Mead J: *Framework for Clinical Guideline Development in Physiotherapy.* Limassol: European Region of the World Confederation for Physical Therapy; 2004.
25. Agència d'Avaluació de Tecnologia i Recerca Mèdiques: *Guies de pràctica clínica.* http://www.gencat.cat/salut/depsan/units/aatrm/html/ca/dir252/doc7841.html (accessed 3 August 2012).
26. Schubert I, Lelgemann M, Kirchner H, von Ferber C, von Ferber L, Ollenschläger G: *Handbuch zur Entwicklung regionaler Leitlinien.* Berlin: Ärztliches Zentrum für Qualität in der Medizin; 2006.
27. Arzneimittelkommission der deutschen Ärzteschaft: *Leitfaden für die Erstellung von Therapieempfehlungen.* Berlin: Arzneimittelkommission der deutschen Ärzteschaft; 2011.
28. Bundesärztekammer: *National Disease Management Guidelines.* Berlin: Bundesärztekammer; 2002.
29. Dumonceau JM, Hassan C, Riphaus A, Ponchon T: **European society of gastrointestinal endoscopy (ESGE) guideline development policy.** *Endoscopy* 2012, **44**(6):626–629.
30. Iorio A, Ageno W, Cosmi B, Imberti D, Lussana F, Siragusa S, Tormene D, Tosetto A, Cattaneo M: **Objectives and methodology: guidelines of the Italian society for haemostasis and thrombosis (SISET).** *Thromb Res* 2009, **124**(5):1–5.
31. Ärztliche Zentralstelle Qualitätssicherung: *National Disease Management Guidelines: Method Report.* Köln: Ärztliche Zentralstelle Qualitätssicherung; 2002.
32. Kunnamo I: *Preface: What is Evidence-Based Medicine Guidelines*, Duodecim Medical Publications Ltd.; 2008.
33. Society DFM: *Submitted NHS Evidence Accreditation Application.* Helsinki: Duodecim Finnish Medical Society; 2001.
34. Chevalier P, de Sutter A, Dirven K, Paulus D, Peremans L, van Royen P, van Welde A: *Algemeen Stramien voor de Ontwikkeling van Aanbevelingen van Goede Medische Praktijkvoering.* Antwerpen: Domus Medical Flemish College of General Practitioners; 2007.
35. World Health Organization: *Guidelines for WHO Guidelines.* Geneva: World Health Organization; 2003.
36. National Health and Medical Research Council: *A guide to the development, implementation and evaluation of clinical practice guidelines.* Canberra: National Health and Medical Research Council; 1998.
37. American Society of Clinical Oncology: *American Society of Clinical Oncology Guideline Procedures Manual.* Alexandria: American Society of Clinical Oncology; 2001.
38. Rosenfeld RM, Shiffman RN, Robertson P: **Clinical practice guideline development manual, third edition: a quality-driven approach for translating evidence into action.** *Otolaryngol Head Neck Surg* 2013, **148**(1 Suppl):S1–S55.
39. Harris JS, Sinnott PL, Holland JP, Ording J, Turkelson C, Weiss M, Hegmann KT: **Methodology to update the practice recommendations in the American college of occupational and environmental medicine's occupational medicine practice guidelines, Second Edition.** *J Occup Environ Med* 2008, **50**(3):282–295.
40. Qaseem A, Snow V, Owens DK, Shekelle P: **The development of clinical practice guidelines and guidance statements of the American college of physicians: summary of methods.** *Ann Intern Med* 2010, **153**(3):194–199.
41. American Urological Association Education and Research: *Overview: Standard Operating Procedures.* Linthicum: American Urological Association; 2011.
42. Guidelines and Protocols Advisory Committee: *GPAC Handbook.* Victoria: Guidelines and Protocols Advisory Committee; 2011.
43. U.S. Preventive Services Task Force: *U.S. Preventive Services Task Force Procedure Manual.* Rockville: Agency for Healthcare Research and Quality; 2008.
44. American College of Cardiology Foundation and American Heart Association: *Methodology Manual and Policies From the ACCF/AHA Task Force on Practice Guidelines.* Washington DC: American College of Cardiology Foundation; 2010.
45. Joanna Briggs Institute Synthesis Science Unit: *Best Practice Information Sheet (BPIS) Procedures.* Adelaide: Joanna Briggs Institute.
46. Canadian Medical Association: *Handbook on Clinical Practice Guidelines.* Ottawa: Canadian Medical Association; 2007.
47. Therapeutic Guidelines Limited: *How Therapeutic Guidelines are produced.* http://www.tg.org.au/?sectionid=81 (accessed 3 August 2012).
48. Caring for Australasians with Renal Impairment: *A Guide for Writers.* Concord: Caring for Australasians with Renal Impairment; 2009.
49. New Zealand Guidelines Group: *Handbook for the Preparation of Explicit Evidence-base Clinical Practice Guidelines.* Wellington: New Zealand Guidelines Group; 2001.

50. Gupta S, Bhattacharyya OK, Brouwers MC, Estey EA, Harrison MB, Hernandez P, Palda VA, Boulet LP: **Canadian thoracic society: presenting a new process for clinical practice guideline production.** *Can Respir J* 2009, **16**(6):62–68.
51. American College of Chest Physicians: *Methodology.* http://www.chestnet.org/Guidelines-and-Resources/Guidelines-and-Consensus-Statements/Methodology (accessed 3 August 2012).
52. Murad MH, Montori VM, Sidawy AN, Ascher E, Meissner MH, Chaikof EL, Gloviczki P: **Guideline methodology of the society for vascular surgery including the experience with the GRADE framework.** *J Vasc Surg* 2011, **53**(5):1375–1380.

doi:10.1186/1748-5908-9-3
Cite this article as: Vernooij *et al.*: Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. *Implementation Science* 2014 **9**:3.

5.2 Article 2: "Reporting Items for Updated Clinical Guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp)"

5.2.1 Summary of the results.

The initial version of the checklist, that was developed through brainstorming and considering the relevant literature, included 13 items in the following 7 domains: updating rationale, scope and purpose of the updated CG, participants in the updating panel, CoI, updating methodology, differentiating original and new information, and the reasons for changes in the recommendations [117].

In the assessment of updated CGs with the initials version of the checklist, no new additional items emerged. Consequently, we conducted 13 semi-structured interviews, at which point data-saturation was reached. As a result, we modified five items and added four new items. Afterwards, with a revised version of the checklist, we conducted a Delphi consensus survey, in which all members of the development panel (n=33) were invited to participate. In the first round, the participants provided substantial feedback on various items, their explanation, and the accompanying examples. Afterwards, in the second and third round of the Delphi survey, respectively, the amount of feedback decreased and in the third round general consensus was reached regarding all items, explanations, and examples.

In the external review, 53 CG methodologists participated and provided comments that improved mostly the writing style of the items, explanations, and examples. However, no substantial modifications were made. Finally, for the last part of the external review, we conducted semi-structured interviews with 10 CG users, when data-saturation was reached and here neither new items nor modifications were proposed. At this phase, the Checklist for the Reporting of Updated Guidelines (CheckUp) was completed and the final version was approved at the G-I-N 2015 conference by the participants of the workshop regarding

updating CGs. CheckUp includes 16 items in three domains: 1) presentation, 2) editorial independence, and 3) the methodology of the updating process (**Table 4**) [117].

Table 4: Final version of CheckUp

Item	Assessment	Reported on page number	Notes
1. The updated version can be distinguished from the previous version of the clinical guideline.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
2. The rationale for updating the clinical guideline is reported.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
3. Changes in the scope and purpose between the updated and previous version are described and justified.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
4. The sections reviewed in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
5. Recommendations are clearly presented and labelled as new, modified, or not changed. Deleted recommendations are clearly noted.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
6. Changes in recommendations are reported and	☐ Yes		

justified.	☐ No ☐ Unclear ☐ Not applicable
7. The panel participants in the updated version are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
8. Disclosures of interests of the group responsible for the updated version are recorded.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
9. The role of the funding body for the updated version is identified and described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
10. The methods used for searching and identifying new evidence in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
11. The methods used for evidence selection in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
12. The methods used to assess the quality of the included evidence in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
13. The methods used for the evidence synthesis in	☐ Yes

the updating process are described.	☐ No ☐ Unclear ☐ Not applicable
14. The methods used for externally reviewing the updated version are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
15. The methods and plan for implementing the changes of the updated version in practice are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable
16. The plan and methods for updating the new version in the future are reported.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable

5.2.2 Copy of article 2.

PLOS | MEDICINE

GUIDELINES AND GUIDANCE

Reporting Items for Updated Clinical Guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp)

Robin W. M. Vernooij[1], Pablo Alonso-Coello[1,2,3]*, Melissa Brouwers[4], Laura Martínez García[1], CheckUp Panel[¶]

1 Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain, **2** CIBER of Epidemiology and Public Health (CIBERESP), Spain, **3** Clinical Epidemiology and Biostatistics Department, McMaster University, Hamilton, Canada, **4** Program in Evidence-based Care, Cancer Care Ontario, Hamilton, Canada

¶ Membership of the CheckUp Panel is provided in the Acknowledgments.
* palonso@santpau.cat

Citation: Vernooij RWM, Alonso-Coello P, Brouwers M, Martínez García L, CheckUp Panel (2017) Reporting Items for Updated Clinical Guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp). PLoS Med 14(1): e1002207. doi:10.1371/journal.pmed.1002207

Published: January 10, 2017

Funding: This project is funded by research grants from the Instituto de Salud Carlos III through the project "PI14/02006" (co-funded by European Regional Development Fund/European Social Fund "Investing in your future"). LMG is funded by a Juan Rodés contract from the Instituto de Salud Carlos III (JR15/00004). PAC is supported by a Miguel Servet investigator contract from the Instituto de Salud Carlos III (CPII15/0034). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: AGREE, Appraisal of Guidelines for Research and Evaluation; CheckUp, Checklist for the Reporting of Updated Guidelines; EQUATOR,

Abstract

Background

Scientific knowledge is in constant development. Consequently, regular review to assure the trustworthiness of clinical guidelines is required. However, there is still a lack of preferred reporting items of the updating process in updated clinical guidelines. The present article describes the development process of the Checklist for the Reporting of Updated Guidelines (CheckUp).

Methods and Findings

We developed an initial list of items based on an overview of research evidence on clinical guideline updating, the Appraisal of Guidelines for Research and Evaluation (AGREE) II Instrument, and the advice of the CheckUp panel ($n = 33$ professionals). A multistep process was used to refine this list, including an assessment of ten existing updated clinical guidelines, interviews with key informants (response rate: 54.2%; 13/24), a three-round Delphi consensus survey with the CheckUp panel (33 participants), and an external review with clinical guideline methodologists (response rate: 90%; 53/59) and users (response rate: 55.6%; 10/18). CheckUp includes 16 items that address (1) the presentation of an updated guideline, (2) editorial independence, and (3) the methodology of the updating process. In this article, we present the methodology to develop CheckUp and include as a supplementary file an explanation and elaboration document.

Conclusions

CheckUp can be used to evaluate the completeness of reporting in updated guidelines and as a tool to inform guideline developers about reporting requirements. Editors may request its completion from guideline authors when submitting updated guidelines for publication. Adherence to CheckUp will likely enhance the comprehensiveness and transparency of

Enhancing the Quality and Transparency of Health Research; G-I-N, Guidelines International Network; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; IOM, Institute of Medicine.

Provenance: Not commissioned; externally peer-reviewed.

clinical guideline updating for the benefit of patients and the public, health care professionals, and other relevant stakeholders.

Background

Trustworthy clinical guidelines aim to assist decision making by providing recommendations that are informed by the best available evidence and include an assessment of the benefits and harms of alternative care options [1,2]. Because of the continuous emergence of new research evidence (i.e., changes in available interventions, effects, or cost) [3], appropriate updating to maintain the trustworthiness of clinical guidelines is challenging since it requires regular surveillance and reviewing of the new evidence [4,5].

Updating clinical guidelines is a process that includes different stages: (1) prioritisation of candidate guidelines or recommendations to update [6], (2) identification of new scientific evidence [3,6–8], (3) assessment of the need to update [3,6,9], (4) updating the recommendations [6,10–12], and (5) publication of the updated guideline [6,13]. However, there is no consensus about what is the optimal methodology to operationalise each of these steps or how to report on the process [5,14,15,16]; the available guidance from guideline institutions is suboptimal [17,18].

Trustworthiness standards for guidelines have been published by both the Institute of Medicine (IOM) and the Guidelines International Network (G-I-N) [1,2]. Additionally, instruments are available for assessing the quality of clinical guidelines, such as the Appraisal of Guidelines for Research and Evaluation (AGREE) II Instrument [19], while others, such as the GIN-McMaster Guideline Development Checklist [20], support developing and implementing trustworthy clinical guidelines. However, guideline updating requires some different methodological considerations and unique communication procedures. Currently, none of the existing tools address these issues.

To address this gap, in a partnership of the Iberoamerican Cochrane Centre (www.cochrane.org), the AGREE Collaboration (www.agreetrust.org), and the G-I-N Updating Guidelines Working Group (www.g-i-n.net/working-groups/updating-guidelines), we have developed the Checklist for the Reporting of Updated Guidelines (CheckUp). This article about CheckUp is targeted at guideline developers and users of guidelines. In the article, we present the methodology of the development process and the final checklist. In a supplementary file, we present explanations and examples for each item (S1 Appendix).

Methodology

For reporting the development process of CheckUp, we followed Enhancing the QUAlity and Transparency Of health Research(EQUATOR) and Moher's criteria [21,22]. The development of CheckUp consisted of four phases: (1) panel selection, (2) generation of the initial checklist, (3) optimisation of the checklist, and (4) approval of the final checklist (Fig 1).

Panel Selection

To advise on the development of the CheckUp, a panel comprising individuals with relevant experience in clinical guideline development and updating and/or in systematic reviews/guidelines research methodology was convened. Invited panel participants were identified based on a review of the main authors in the field, as well as the AGREE Trust (www.agreetrust.org) and

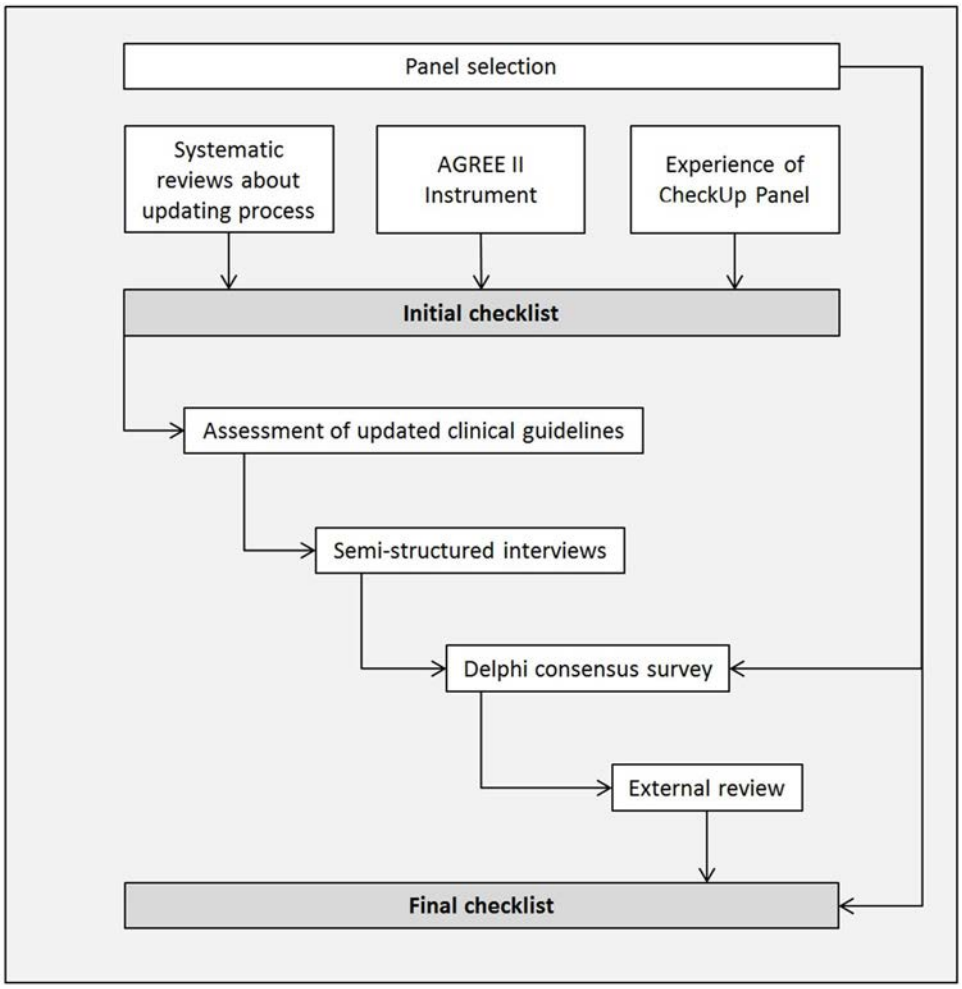


Fig 1. Checklist development process. Abbreviation: AGREE, Appraisal of Guidelines for Research and Evaluation.

doi:10.1371/journal.pmed.1002207.g001

the G-I-N (www.g-i-n.net) members. The purpose of the panel was to provide expert advice during the development process and to participate in the Delphi survey. A core group (RWMV, PAC, MB, and LMG) was established to design the protocol and provide more time-sensitive and operational advice.

Generation of the Initial Checklist

The core group first developed an initial list of items—including explanation and examples—through brainstorming and discussion, taking into account (1) research evidence in the field [15,16,18], (2) the AGREE II Instrument [19], and (3) the panel experience. We used three core updating publications as a starting point from which the initial version of the checklist was generated [15,16,18]. These studies include an overview of the available guidance from guideline methodological handbooks [18], a systematic review of the published methodological research [16], and an international survey about the experiences of the main guideline developers [15].

Optimisation of the Checklist

We optimised the initial checklist through a multistep process that included an assessment of existing updated clinical guidelines, semistructured interviews, a Delphi consensus survey, and an external review with clinical guideline methodologists and users (Fig 1).

Assessment of updated clinical guidelines. We piloted the initial checklist among a convenience sample of updated clinical guidelines to assess the used terminology, to identify missing items, and as a first step to explore its face validity. We included updated clinical guidelines that were (1) developed by G-I-N members, (2) published in English or Spanish, and (3) published between 2011 and 2013. We searched the G-I-N library (www.g-i-n.net) and the National Guideline Clearinghouse (www.guidelines.gov). Two reviewers (RWMV and SPS) applied the checklist, solving disagreements by consensus. The core group discussed the results and refined the initial list of items.

Semistructured interviews. To refine the checklist and to identify missing items, we conducted semistructured interviews with clinical guideline experts. We chose a convenience sample of participants, outside the CheckUp panel, with (1) experience in updating clinical guidelines, defined as having participated in the updating process of at least one clinical guideline over the past year, and (2) fluency in English. We identified the participants by contacting professionals associated with G-I-N or researchers in the field. When someone did not respond or could not participate, a new person was recruited. We continued to recruit participants and collect data until data saturation was achieved.

In each interview, participants were asked about their experiences and challenges in updating clinical guidelines. Subsequently, the participant was prompted by the interviewer (RWMV) to reflect on the strengths, weaknesses, missing concepts, and redundancies in the checklist. The interviews were audiotaped, and key themes were identified. The core group discussed the results and refined the list of items.

Delphi consensus survey. We reviewed the refined list of items through a Delphi consensus survey [23], with all members of the CheckUp panel. The Delphi participants assessed the inclusion, comprehensiveness, clarity, and coverage of each item and tried to identify potentially additional items for the checklist.

Using a seven-point Likert scale (one meaning strongly disagree and seven meaning strongly agree) [24], we asked participants to rate whether the item should be included in the checklist. For each item, participants were asked whether their perceptions of (1) the completeness, (2) the usability, and (3) the quality of a clinical guideline would be influenced if the item was reported. We included a free text box for suggestions to modify the items, the explanation, or the examples. We used online software to design the survey and to collect the responses (www.surveymonkey.com).

We calculated the median score for inclusion, completeness, usability, and quality for each item and classified them into (1) items with a median score of 0 to 3 points, which were excluded; (2) items with a median score of 4 to 5 points or with substantial comments that needed important revision, which were retained, modified, and further tested; and (3) items with a median score of 6 to 7 points and without substantial comments, which were included and not evaluated further in the following rounds.

One reviewer (RWMV) analysed the quantitative and qualitative results and suggested potential solutions. The core group discussed the results and potential solutions and refined the list of items accordingly. We continued with additional rounds until consensus for inclusion or exclusion was reached and no more relevant comments were provided.

External review with clinical guideline methodologists. To evaluate the usability of the checklist, we conducted a survey with clinical guideline methodologists who had experience in

updating clinical guidelines, as measured by having participated in the updating process of at least one clinical guideline over the past year. We also invited all of the G-I-N institutional member contacts to participate in the external review. If the contact person was not able to participate, we asked them to provide contact details of another expert working at the same institution.

Using a seven-point scale (one meaning strongly disagree and seven meaning strongly agree), we asked participants to rate the usability of each item and its influence on the confidence in an updated clinical guideline if the item was reported. A free text option was included for suggestions to modify the items, the explanation, or the examples. We used online software to design the survey and to collect the responses (www.surveymonkey.com).

We calculated the median score for usability and confidence for each item. One reviewer (RWMV) analysed the quantitative and qualitative results and suggested potential solutions. The core group discussed the results and potential solutions and refined the list of items accordingly.

External review with clinical guideline users. We conducted semistructured interviews with clinical guideline users to evaluate the usability of the checklist. We engaged individuals who were (1) health care professionals who used clinical guidelines in clinical practice and (2) located in Canada, Spain, or the Netherlands. We identified the participants with the help of the panel members. When someone did not respond or could not participate, a new person was recruited. We continued to recruit participants and collect data until the information was repeated and no new information emerged (data saturation).

For each interview, participants were asked whether reporting of the item in an updated clinical guideline would increase their confidence in the guideline. The participant and interviewer (RWMV) reviewed the checklist, and the participants were prompted to consider missing concepts, redundancy, and the usability of the checklist. The interviews were audiotaped, and key themes were identified. The core group discussed the results and refined the list of items.

Approval of the final checklist. The checklist was presented and discussed in a workshop at the 2015 G-I-N Conference in Amsterdam. In this workshop, we asked the participants whether the checklist was deemed adequate for assessment of the updated clinical guidelines [25]. We also asked the participants to give an overall impression of the checklist. The core group discussed the results and agreed on the final list of items.

Results

CheckUp Panel

Fifty-six potential individuals were invited to be part of the CheckUp Panel. In total, 33 professionals, 20 males and 13 females, (response rate: 58.9%, 33/56) confirmed their participation (17 from Europe, 9 from South America, 5 from North America, and 2 from Oceania). The primary role of the panellists was health care researcher (60.6%; 20/33), guideline developer (30.3%; 10/33), and clinical guideline user (9.1%; 3/33).

Generation of the Initial Checklist

The initial checklist included 13 items within the following domains: updating rationale, scope and purpose of the updated clinical guideline, participants in the updating panel, conflicts of interest, updating methodology, differentiating original and new information, and reasons for the changes in the recommendations.

Optimisation of the Checklist

Assessment of updated clinical guidelines. Initially, we assessed a convenience sample of ten updated clinical guidelines from the G-I-N library with the initial checklist [26–35]. The items more frequently reported in the included clinical guidelines were related to the literature search strategy (60%), the composition of the panel (50%), and the external review (40%). The other checklist items (e.g., assessment for the need of updating, evidence selection, rationale for updating, and rationale for changes) were reported in less than 20% of the included clinical guidelines. No additional concepts or items emerged (Table 1).

Semistructured interviews. We conducted semistructured interviews with clinical guideline developers (5 from Europe, 5 from North America, and 3 from South America). In total, we interviewed 13 participants, at which point saturation was reached. As a result, we modified five items and added four new ones. The modifications were related to (a) differences in the objectives, purpose, or aim between the original and updated version; (b) the identification of new evidence; (c) the rationale for changing recommendations; and (d) the funding. The new items were related to (a) the scope of the update (partial or complete), (b) the target audience, (c) the changes in the recommendations, and (d) the plans and methodology reported to update the clinical guideline (Table 1).

Delphi consensus survey. All the members of the CheckUp panel (n = 33) were invited to participate in the Delphi consensus survey. Twenty-seven (82%) members participated in the first Delphi round, thirty-one members (93.9%) in the second Delphi round, and all (100%) members in the third and final round.

In the first round, the participants provided substantial feedback on various items, their explanation, and the accompanying examples. This feedback triggered modification in the order of the items, phrasing of the items, explanations, and examples (Table 1). All items met the inclusion criteria, and no major comments were reported. The median score for whether the participants believed that an updated clinical guideline would be more complete, usable, and of higher quality whenever the item was reported was six for all questions.

Table 1. CheckUp: Stages of the optimisation process (objective, sample, and results by optimisation processes).

Stage	Objectives	Sample (n)	Main Results
1. Assessment of updated clinical guidelines	• Assess whether the terminology was consistent between the checklist and updated clinical guidelines • Identify items that were lacking in the checklist.	• Updated clinical guidelines (10).	• No items were modified. • No new items were added.
2. Semistructured interviews	• Explore challenges and issues regarding the clinical guideline updating process. • Identify items that were lacking in the checklist.	• Clinical guideline updating process experts (13).	• Five items had major modifications. • Four new items were added.
3. Delphi consensus survey	• Assess the inclusion, comprehensiveness, clarity, and coverage of each item. • Identify items that were lacking in the checklist.	• Clinical guideline updating process experts (33) (CheckUp panel).	• All items, explanations, and examples had minor modifications regarding writing style. • Two items were combined. • No new items were added.
4. External review	• Evaluate the usability of the checklist.	• Clinical guideline methodologists (53).	• All items, explanations, and examples had minor modifications regarding writing style.
		• Clinical guideline users (10).	• No items were modified.

doi:10.1371/journal.pmed.1002207.t001

Table 2. Results of the Delphi survey (third round).

Item	Median Score for Inclusion[a] (0–7)*	Median Score for Completeness[b] (0–7)*	Median Score for Usability[c] (0–7)*	Median Score for Quality[d] (0–7)*
1. Distinguishing the updated and original version.	7	6	6	6
2. Reviewed and changed sections.	6	6	6	5
3. Presentation of new, modified, or not changed recommendations.	7	6	6	6
4. Working group of the updating process.	6	6	5	6
5. Rationale for updating the guideline.	6	6	5	6
6. Differences in the scope and purpose between the updated and the original guideline.	6	6	6	5.5
7. Reporting and justification of changes in the recommendations.	6	6.5	6	6
8. Methods for searching and identifying new evidence.	7	7	6	7
9. Methods for evidence selection.	6.5	7	6	7
10. Methods to assess the quality of the included evidence.	7	6	6	6
11. Methods for evidence synthesis.	6	6	6	6
12. Methods and plan for implementing the changes.	6	6	6	6
13. Methods for external reviewing.	6	6	5.5	6
14. Plan and methods for updating in the future.	6	6	5	6
15. Conflicts of interests of the updating group.	7	7	6	7
16. Role of the funding body.	6.5	6.5	6	6.5

[a]Should this item be included in the reporting checklist for updated clinical guidelines?;

[b]Whether this information was present or not would influence your perceptions of the completeness of the reporting of an updated clinical guideline;

[c]Whether this information was present or not would influence your perceptions of the usability of an updated clinical guideline;

[d]Whether this information was present or not would influence your perceptions of the quality of an updated clinical guideline.

* Seven-point Likert scale (0 meaning ¨strongly disagree¨ and 7 meaning ¨strongly agree¨).

doi:10.1371/journal.pmed.1002207.t002

In the second round (n = 17 items), the amount of feedback was substantially smaller than in the first round (Table 1). We merged some items, and the checklist was reduced to 16 items. Again, all items met the inclusion criteria, and no major comments were reported. The median score for whether the participants believed the clinical guideline would be more complete, usable, and of higher quality was 7.0, 6.0, and 6.0, respectively.

In the third and last round, a general consensus was reached for all items, explanations, and examples. The median score for item inclusion, completeness, usability, and quality was $\geq$6 in all items (except for two items with a median score of 5.5 in the usability and quality question) (Table 2).

External review. *External review with clinical guideline methodologists.* We conducted a survey with 53 clinical guideline methodologists (53/59, response rate 90%). The median scores of usability and confidence for each item were $\geq$6 (Table 3). Participants provided comments that improved the writing style of the items, explanations, and examples (Table 1).

External review with clinical guideline users. We had conducted semistructured interviews with 10 clinical guideline users (3 from Spain, 2 from the Netherlands, and 5 from Canada) when saturation was reached. All participants acknowledged that all items were useful to evaluate the reporting of the updating process in updated clinical guidelines. Neither new items nor modifications were proposed (Table 1).

Table 3. Results of the external review with clinical guideline methodologists.

Item	Median Score for Usability[a] (0–7)*	Median Score for Confidence[b] (0–7)*
1. The updated version is distinguished from the previous version of the guideline.	6	6
2. The sections reviewed in the updating process are described.	7	6.5
3. The recommendations are clearly presented and labelled as new, modified, or no change. Deleted recommendations are clearly noted.	6	6
4. The panel participants in the updated version are described.	6	6
5. The rationale for updating the guideline is reported.	6	6
6. Changes in the scope and purpose between the updated and original version are described and justified.	6.5	6
7. Changes in the original recommendations are reported and justified.	6	6
8. The methods used for searching and identifying new evidence in the updating process are described.	7	7
9. The methods used for evidence selection in the updating process are described.	7	7
10. The methods used to assess the quality of the included evidence in the updating process are described.	7	7
11. The methods used for the evidence synthesis in the updating process are described.	6	6
12. The methods and plan for implementing the changes of the updated version in practice are described.	5	5
13. The methods used for externally reviewing the updated version are described.	6	6
14. The plan and methods for updating the new version in the future are reported.	6	6
15. The conflicts of interests of the group responsible for the updated version are recorded.	7	7
16. The role of the funding body for the updated guideline is identified and described.	7	7

[a]This item is useful to evaluate an updated clinical guideline;

[b]I have more confidence in an updated clinical guideline if this item is reported.

* Seven-point scale (0 meaning ¨strongly disagree¨ and 7 meaning ¨strongly agree¨).

doi:10.1371/journal.pmed.1002207.t003

Final Checklist

The checklist includes 16 items that can be broadly categorised into three themes: (1) presentation (e.g., clinical guideline sections and recommendations), (2) editorial independence (e.g., the working group and funding), and (3) the methodology used (e.g., search strategy and evidence synthesis) (Table 4). Those attending the presentation of the checklist workshop at the G-I-N 2015 conference reviewed and agreed with the final version of the checklist.

Discussion

We developed CheckUp through a comprehensive development process, including the use of systematic reviews, assessment of updated clinical guidelines, and engagement of the international guideline community through semistructured interviews, a Delphi consensus survey, and an external review.

Main Findings

Across the different processes, an alignment and consensus of opinion emerged between what was documented in the literature and the expectations of clinical guideline developers, users, and researchers in regards to what information ought to be reported in an updated clinical guideline. CheckUp includes 16 items regarding the presentation of the updated clinical guideline, editorial independence, and the methodology used in the clinical guideline updating process.

Table 4. Final version of CheckUp.

Item	Assessment	Reported on Page Number	Notes
1. The updated version can be distinguished from the previous version of the clinical guideline.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
2. The rationale for updating the clinical guideline is reported.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
3. Changes in the scope and purpose between the updated and previous version are described and justified.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
4. The sections reviewed in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
5. Recommendations are clearly presented and labelled as new, modified, or not changed. Deleted recommendations are clearly noted.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
6. Changes in recommendations are reported and justified.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
7. The panel participants in the updated version are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
8. Disclosures of interests of the group responsible for the updated version are recorded.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		

(*Continued*)

Table 4. (*Continued*)

Item	Assessment	Reported on Page Number	Notes
9. The role of the funding body for the updated version is identified and described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
10. The methods used for searching and identifying new evidence in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
11. The methods used for evidence selection in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
12. The methods used to assess the quality of the included evidence in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
13. The methods used for the evidence synthesis in the updating process are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
14. The methods used for externally reviewing the updated version are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
15. The methods and plan for implementing the changes of the updated version in practice are described.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		
16. The plan and methods for updating the new version in the future are reported.	☐ Yes ☐ No ☐ Unclear ☐ Not applicable		

doi:10.1371/journal.pmed.1002207.t004

CheckUp was primarily developed to evaluate the completeness of reporting in updated guidelines. Additionally, the tool can inform guideline developers about strategies for updating clinical guidelines and their reporting requirements. An explanation and elaboration article for the CheckUp is published as a supporting information article (S1 Appendix). CheckUp can be used in several ways. The checklist can provide guidance to developers who update clinical guidelines, by providing methodological principles that should be incorporated into the updating process, as well as strategies for reporting this information. The checklist can be applied by users or appraisers of clinical guidelines to assess whether updated clinical guidelines align with the CheckUp items. We suggest that a minimum of two reviewers assess the reporting of the guideline updating process independently, with the help of a third reviewer if there is a need of reaching consensus.

Our Results in the Context of Previous Research

Updating is a crucial part of maintaining the trustworthiness of clinical guidelines [1,2]. Since clinical guidelines have a limited lifespan, updating clinical guidelines is crucial to maintain the validity of the recommendations [3,9,36,37]. Although the importance of regular updating has been recognised and clinical guidelines may have an "expiration date," little research has been conducted in the field so far [13,15–18]. Published standards for trustworthy guidelines require the description of updating plans [1,2]; however, these standards do not provide specific guidance about the detailed reporting of the updating process of guidelines.

Strengths and Limitations

Our CheckUp proposal has several strengths. For the development process, we systematically reviewed the evidence and followed EQUATOR and Moher's criteria [21,22]. Also, by applying a formal consensus method (Delphi survey) and collecting experts' opinions (semistructured interviews and external reviews), we reached a fair understanding of clinical guideline methodologists' and users' perceptions about the updating of clinical guidelines. Finally, there was fairly strong overall consensus during the development of CheckUp.

Our study has some limitations. We used consensus methods and convenience samples of clinical guideline stakeholders. However, across the different processes, an alignment and consensus of opinion emerged on what clinical guideline developers, users, and researchers expect to see reported in updated clinical guidelines. Another potential limitation is that CheckUp includes some items that may partially overlap with some items that are present in other instruments [19,20]; however, we think this is a minor limitation as CheckUp differs for the most part and has a very specific and differentiated goal. Finally, we did not collect potential conflicts of interest in our panel.

Implications for Practice and Research

CheckUp can be used for multiple purposes. Firstly, guideline developers can use it both for the reporting of their guidelines and to plan their updating processes. Guideline users can assess the reporting of updated guidelines. Editors may request its completion from guideline authors. CheckUp provides an overall picture of how complete the updating process is reported in updated clinical guidelines. Being a reporting checklist, CheckUp does not evaluate the quality of the updating processes, as there are no gold standards for this process. Currently, the G-I-N Updating Guidelines Working Group (http://www.g-i-n.net/working-groups/updating-guidelines) is undertaking an analysis of current guideline updating methods worldwide. From this work, strategies or advice might come on how we might assess guideline updating quality.

There are currently no gold standards for guideline updating methodology. Nonetheless, updating is key to ensuring trustworthy, implementable, and clinically relevant recommendations. Current guideline evaluation tools or guideline method resources (e.g., AGREE II, Grading of Recommendations Assessment, Development, and Evaluation (GRADE), IOM Standards, and the like) are not simply transferable to the conceptual requirements of an updated guideline. CheckUp addresses the gap: it has been supported by our study participants and is a resource that complements (rather than competes with) the other high-quality tools available in the guideline enterprise.

Further rigorous research in updating clinical guidelines is warranted, and we invite users to comment on the items and the usability of CheckUp. It would be important to assess the impact of CheckUp in the updating clinical guideline field over the next few years [16]. When dynamic or living guidelines become a reality, [38] some adaptation of CheckUp could potentially be necessary. Finally, the G-I-N Updating Guidelines Working Group will continue to play a key role in this work and in moving forward the updating agenda in the clinical guideline enterprise.

Supporting Information

S1 Appendix. CheckUp: Explanation and elaboration of a checklist for the reporting of updating clinical guidelines.
(DOCX)

Acknowledgments

Robin W. M. Vernooij is a doctoral candidate at the Paediatrics, Obstetrics and Gynaecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain. The authors thank Sandra Pequeño Saco (Iberoamerican Cochrane Centre) for her support in assessing updated clinical guidelines. The authors also thank María Victoria Leo for her help editing the manuscript.

Contributors. Vernooij RWM, Alonso-Coello P, Brouwers M, Martínez García L, Florez ID, Iorio A, James R, Sanabria AJ, Selva A, Shekelle PG, and Vandvik PO are members of the G-I-N Updating Guidelines Working Group. The members of the CheckUp panel are as follows: Vernooij RWM, Alonso-Coello P, Brouwers M, Martínez García L, Ada L, Alemán A, Arévalo-Rodriguez I, Becker M, Burgers JS, Chan W, Delvaux N, Duggan G, Enciso Olivera CO, Etxeandia-Ikobaltzeta I, Florez ID, Follmann M, Gartlehner G, Iorio A, James R, Jones SL, Kotzeva A, Lloyd M, López Gallegos D, Louro-González A, Marin Leon I, Martí-Carvajal A, Meerpohl JJ, Pardo R, Rojas-Reyes MX, Rotaeche R, Sanabria AJ, Selva A, Shekelle PG, Sierra Matamoros FA, van de Velde S, Vandvik PO, and Willett S.

References

1. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. Ann Intern Med. 2011; 154(11): 774–775. doi: 10.7326/0003-4819-154-11-201106070-00011 PMID: 21646561
2. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P, et al. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med. 2012; 156(7):525–31. doi: 10.7326/0003-4819-156-7-201204030-00009 PMID: 22473437
3. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001; 286(12): 1461–1467. PMID: 11572738
4. Grupo de trabajo sobre actualización de GPC. Actualización de Guías de Práctica Clínica en el Sistema Nacional de Salud. Manual Metodológico. Plan de Calidad para el Sistema Nacional de Salud del

Ministerio de Sanidad y Política Social. Instituto Aragonés dológico. de Ciencias de la Salud-I+CS; 2009. Guías de Práctica Clínica en el SNS: I+CS N° 2007/02-01

5. Shekelle PG. Updating practice guidelines. JAMA. 2014; 311(20): 2072–2073. doi: 10.1001/jama.2014.4950 PMID: 24867011
6. Agbassi C, Messersmith H, McNair S, Brouwers M. Priority-based initiative for updating existing evidence-based clinical practice guidelines: the results of two iterations. J Clin Epidemiol. 2014; 67(12): 1335–1342. doi: 10.1016/j.jclinepi.2014.06.013 PMID: 25216900
7. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Assessing the need to update prevention guidelines: a comparison of two methods. Int J Qual Health Care. 2004; 16(5): 399–406. doi: 10.1093/intqhc/mzh081 PMID: 15375101
8. Martínez García L, Sanabria AJ, Araya I, Lawson J, Solà I, Vernooij RW, et al. Efficiency of pragmatic search strategies to update clinical guidelines recommendations. BMC Med Res Methodol. 2015; 15: 57. doi: 10.1186/s12874-015-0058-2 PMID: 26227021
9. Martínez García L, Sanabria AJ, García Alvarez E, Trujillo-Martín MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ. 2014; 186 (16): 1211–1219. doi: 10.1503/cmaj.140547 PMID: 25200758
10. Eccles M, Rousseau N, Freemantle N. Updating evidence-based clinical guidelines. J Health Serv Res Policy. 2002; 7(2): 98–103. PMID: 11934374
11. Johnston ME, Brouwers MC, Browman GP. Keeping cancer guidelines current: results of a comprehensive prospective literature monitoring strategy for twenty clinical practice guidelines. Int J Technol Assess Health Care. 2003; 19(4): 646–655. PMID: 15095770
12. Parmelli E, Papini D, Moja L, Bandieri E, Belfiglio M, Ciccone G, et al. Updating clinical recommendations for breast, colorectal and lung cancer treatments: an opportunity to improve methodology and clinical relevance. Ann Oncol. 2011; 22(1): 188–194. doi: 10.1093/annonc/mdq324 PMID: 20605933
13. Martínez García L, McFarlane E, Barnes S, Sanabria AJ, Alonso-Coello P, Alderson P. Updated recommendations: an assessment of NICE clinical guidelines. Implement Sci. 2014; 9: 72. doi: 10.1186/1748-5908-9-72 PMID: 24919856
14. Burgers JS, Cluzeau FA, Hanna SE, Hunt C, Grol R. Characteristics of high-quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. Int J Technol Assess Health Care. 2003; 19(1): 148–157. PMID: 12701947
15. Alonso-Coello P, Martínez García L, Carrasco Gimeno JM, Solà I, Qureshi S, Burgers JS, et al. The updating of clinical practice guidelines: insights from an international survey. Implement Sci. 2011; 6: 107. doi: 10.1186/1748-5908-6-107 PMID: 21914177
16. Martínez García L, Arévalo-Rodríguez I, Solà I, Haynes RB, Vandvik PO, Alonso-Coello P, et al. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci. 2012; 7(1): 109.
17. Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol. 2014; 67(1): 33–45. doi: 10.1016/j.jclinepi.2013.06.021 PMID: 24125894
18. Vernooij RW, Sanabria AJ, Solà I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci. 2014; 9: 3. doi: 10.1186/1748-5908-9-3 PMID: 24383701
19. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010; 182(18): E839–842. doi: 10.1503/cmaj.090449 PMID: 20603348
20. Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ. 2014; 186(3): E123–142. doi: 10.1503/cmaj.131237 PMID: 24344144
21. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med. 2010; 7(2): e1000217. doi: 10.1371/journal.pmed.1000217 PMID: 20169112
22. Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG, et al. Describing reporting guidelines for health research: a systematic review. J Clin Epidemiol. 2011; 64(7): 718–742. doi: 10.1016/j.jclinepi.2010.09.013 PMID: 21216130
23. Hsu C, Sandford B. The delphi technique: making sense of consensus. Pract Assessment Res Eval. 2007; 12: 1–8.
24. Likert R. A technique for the measurement of attitudes. Archives of psychology. 1932.
25. Burgers J, Alonso-Coello P, Vernooij R. Staying alive: methods to keeping guideline programs up to date. In: 12th G-I-N Conference; 2015. http://www.g-i-n.net/conference/12th-conference/2015-scientific-programme.

26. Rooke TW, Hirsch AT, Misra S, Sidawy AN, Beckman JA, Findeiss L, et al. Management of patients with peripheral artery disease (compilation of 2005 and 2011 ACCF/AHA Guideline Recommendations): a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013; 61(14): 1555–1570. doi: 10.1016/j.jacc.2013.01.004 PMID: 23473760

27. Qaseem A, Wilt TJ, Weinberger SE, Hanania NA, Criner G, van der Molen T, et al. Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline update from the American College of Physicians, American College of Chest Physicians, American Thoracic Society, and European Respiratory Society. Ann Intern Med. 2011; 155(3): 179–191. doi: 10.7326/0003-4819-155-3-201108020-00008 PMID: 21810710

28. Khatcheressian JL, Hurley P, Bantug E, Esserman LJ, Grunfeld E, Halberg F, et al. Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. J Clin Oncol. 2013; 31(7): 961–965. doi: 10.1200/JCO.2012.45.9859 PMID: 23129741

29. Haas FJLM, de Vries RRP. CBO Richtlijn Bloedtransfusie. 2011.

30. Lougheed MD, Lemiere C, Ducharme FM, Licskai C, Dell SD, Rowe BH, et al. Canadian Thoracic Society 2012 guideline update: diagnosis and management of asthma in preschoolers, children and adults. Can Respir J. 2012; 19(2): 127–164. PMID: 22536582

31. Grupo de trabajo de la guía de práctica clínica sobre dispepsia. Manejo del paciente con dispepsia. Guía de práctica clínica. Barcelona: Asociación Española de Gastroenterología, Sociedad Española de Medicina de Familia y Comunitaria y Centro Cochrane Iberoamericano; Actualización 2011. Programa de Elaboración de Guías de Práctica Clínica en Enfermedades Digestivas, desde la Atención Primaria a la Especializada: 3.

32. National Health and Medical Research Council. Clinical Practice Guideline for the Management of Borderline Personality Disorder. Melbourne: National Health and Medical Research Council; 2012.

33. National Collaborating Centre for Mental Health (UK), Social Care Institute for Excellence (UK). Antisocial Behaviour and Conduct Disorders in Children and Young People: Recognition, Intervention and Management. Leicester (UK): British Psychological Society; 2013.

34. New Zealand Guidelines Group. (2012). Cancer control strategy guidance completion: update of evidence for prostate-specific antigen (PSA) testing in asymptomatic men. Wellington: Ministry of Health.

35. Scottish Intercollegiate Guidelines Network (SIGN). Acute coronary syndromes. Edinburgh: SIGN; 2013. (SIGN publication no. 93). [February 2013]. www.sign.ac.uk

36. Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. Updating clinical practice recommendations: is it worthwhile and when? Int J Technol Assess Health Care. 2012; 28(1): 29–35. doi: 10.1017/S0266462311000675 PMID: 22236808

37. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. J Clin Epidemiol. 2014; 67(1): 52–55. doi: 10.1016/j.jclinepi.2013.07.012 PMID: 24139089

38. Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. Chest. 2013; 144(2): 381–389. doi: 10.1378/chest.13-0746 PMID: 23918106

5.3 Article 3: "Updated clinical guidelines experience major reporting limitations"

5.3.1 Summary of the results.

We screened, in total, 1,465 references on the titles and abstracts and 216 references were included for full-text evaluation [118]. Finally, we included a total of 60 updated CGs, of which the majority was updated by a North-American scientific society or public institution. The majority of the included CGs addressed the management of a specific disease and the most common clinical area was oncology.

Regarding the presentation domain, all CGs could be distinguished from the previous version and more than half reported the rationale for updating, changes in the scope and purpose, and the reviewed sections in the updating process. However, solely 27% (n=16) of the included CGs clearly labelled the recommendations as new, modified, or not changed, and 38% (n=23) of the included CGs justified the changes in the recommendations. The majority of the CGs reported the required information for the editorial independence domain (i.e. panel participants and their CoI), however, half of the included CGs (n=30) reported the entity or the role of the funding body of the updating process. Finally, most of the included CGs reported the methodology for the search strategy, evidence selection, and assessment of the quality of the evidence. However, we found suboptimal reporting in the methods for synthesising the evidence, externally reviewing the CG, implementing the changes of the updated version in practice, and future updating of the CG.

CGs developed by a European (median overall score: 8.1) or international institutions (median overall score: 7.8) had higher scores than North American or Asian institutions (both median overall score: 5.6, p=0.014). However, no differences between types of organisation, CG focus, or CG topic were found in the overall score. Finally, the agreement on the overall score

among the three reviewers on the overall score was excellent (ICC 0.88, 95%-CIU: 0.75 to 0.95).

5.3.2 Copy of article 3.

Vernooij et al. Implementation Science (2017) 12:120
DOI 10.1186/s13012-017-0651-3

Implementation Science

RESEARCH Open Access

Updated clinical guidelines experience major reporting limitations

Robin W.M. Vernooij[1], Laura Martínez García[1*], Ivan Dario Florez[2,3], Laura Hildago Armas[1], Michiel H.F. Poorthuis[4], Melissa Brouwers[2,5] and Pablo Alonso-Coello[1,6]

Abstract

Background: The Checklist for the Reporting of Updated Guidelines (CheckUp) was recently developed. However, so far, no systematic assessment of the reporting of updated clinical guidelines (CGs) exists. We aimed to examine (1) the completeness of reporting the updating process in CGs and (2) the inter-observer reliability of CheckUp.

Methods: We conducted a systematic assessment of the reporting of the updating process in a sample of updated CGs using CheckUp. We performed a systematic search to identify updated CGs published in 2015, developed by a professional society, reporting a systematic review of the evidence, and containing at least one recommendation. Three reviewers independently assessed the CGs with CheckUp (16 items). We calculated the median score per item, per domain, and overall, converting scores to a 10-point scale. Multiple linear regression analyses were used to identify differences according to country, type of organisation, scope, and health topic of updated CGs. We calculated the intraclass coefficient (ICC) and 95% confidence interval (95% CI) for domains and overall score.

Results: We included in total 60 updated CGs. The median domain score on a 10-point scale for presentation was 5.8 (range 1.7 to 10), for editorial independence 8.3 (range 3.3 to 10), and for methodology 5.7 (range 0 to 10). The median overall score on a 10-point scale was 6.3 (range 3.1 to 10). Presentation and justification items at recommendation level (respectively reported by 27 and 38% of the CGs) and the methods used for the external review and implementing changes in practice were particularly poorly reported (both reported by 38% of the CGs). CGs developed by a European or international institution obtained a statistically significant higher overall score compared to North American or Asian institutions ($p = 0.014$). Finally, the agreement among the reviewers on the overall score was excellent (ICC 0.88, 95% CI 0.75 to 0.95).

Conclusions: The reporting of updated CGs varies considerably with significant room for improvement. We recommend using CheckUp to assess the updating process in updated CGs and as a blueprint to inform methods and reporting strategies in updating.

Keywords: Checklist/standards, Guideline [publication type], Publishing/standards

Background

Clinical guidelines (CGs) are defined as 'statements that include recommendations intended to optimise patient care, that are informed by systematic reviews of evidence and an assessment of the benefits and harms of alternative care options' [1]. Scientific knowledge is in constant evolution [2, 3]; therefore, surveillance of the new evidence is required to ensure the trustworthiness of clinical guidelines (CGs) [4–8].

Updating CGs is an iterative process with a systematic and explicit methodology that involves identifying and reviewing new evidence not included in the original version of a CG [9]. The fundamental stages of the updating process are (1) prioritising of CGs and clinical questions [10, 11], (2) identifying of new evidence [8, 12, 13], (3) assessing the impact of the new evidence and decision to update [4, 8], (4) reviewing and—if necessary—modifying the recommendations [14–16], and (5) reporting updated recommendations [17]. Currently, there is no

* Correspondence: laura.martinez.garcia@cochrane.es
[1]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain
Full list of author information is available at the end of the article

consensus about the optimal methodology to maintain CGs up-to-date [11, 18, 19].

The reporting of updated CGs is a process within an updating strategy that communicates users about the methods and changes in an updated CG [9]. So far, there is limited guidance on the reporting of the updating process [19]. To address this gap, we recently developed the Checklist for the Reporting of Updated Guidelines (CheckUp) [20]. The aim of CheckUp is to evaluate the completeness of reporting in updated CGs [20]. CheckUp can be used (1) to inform about strategies for updating CGs and their reporting requirements (CG developers), (2) to assess the reporting of updated CGs (interested CG users), and (3) to complete as a publication requirement of updated CGs (editors of scientific journals that publish CGs) [20]. Although CheckUp has been already included in some methodological handbooks and methodological studies [21, 22], it has not been yet formally implemented.

To our knowledge, updated CGs have not been systematically reviewed to assess the completeness of reporting the updating process. An overview of the current status could be informative for the CG community. Therefore, the objectives of our study were (1) to assess the completeness of reporting the updating process of updated CGs using CheckUp and (2) to explore the inter-observer reliability of CheckUp.

Methods

Study design

We conducted a systematic assessment of the reporting of the updating process in a sample of updated CGs using CheckUp. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline to the extent it was applicable to our study [23].

Information sources and search strategy

We searched in MEDLINE (accessed through PubMed), the G-I-N library (http://www.g-i-n.net), and the National Guidelines Clearinghouse (NGC) (https://www.guidelines.gov) in August 2016 for updated CGs published during 2015. The search strategy can be found in Additional file 1.

Inclusion criteria

We included all updated CGs published in 2015 (as the most recent year prior to publication of CheckUp) which met the following criteria: (1) developed by a professional society, (2) search strategy using at least one bibliographic database, (3) reporting at least one recommendation, (4) updated version of a previous version of the same CG (including a reference to a previous version of the CG), and (5) published in English.

Study selection

Two reviewers (RV, IDF, LHA, or MHFP) independently screened the titles and abstracts to identify potentially eligible references. We obtained the full-text articles of the potentially eligible references for further assessment. Disagreements were solved by consensus and, if necessary, with the help of a third reviewer (LMG).

Data extraction

CheckUp is a checklist consisting of 16 items that examine the reporting of the updating process in updated CGs [20]. CheckUp consists of three domains: (1) presentation of the updated CG (6 items), (2) editorial independence (3 items), and (3) the methodology of the updating process (7 items).

Three reviewers (RV, IDF, LHA, or MHFP) independently evaluated each CG with CheckUp, and whenever the included CGs referred to supplemental documents (e.g. methodological manuals or appendices), these documents were reviewed for additional information.

Furthermore, we collected the following information regarding: (1) the institution that updated the CG (name, country, and type of organisation), (2) the scope of the updated CG (diagnosis, management, prevention, screening, or treatment), and (3) the health topic of the updated CG.

Data analysis

We calculated summary statistics to provide quantitative information about the institution that updated CGs and CheckUp scores. We calculated item scores (absolute frequencies and proportions) by summing up the updated CGs that reported each item. We calculated domain scores (median and range) by summing up all scores of the individual items for each domain: presentation of the updated CG (6 items), editorial independence (3 items), and the methodology of the updating process (7 items). Additionally, we calculated the overall score (median and range) by summing up all scores of the individual items. Both domain scores and total scores were converted to a 10-point scale.

To identify potential predictors, we used multiple linear regression to test whether the overall score (dependent variable) differed between CG institution's country, type of organisation, objective of the CG, and CG topic (independent variables).

We calculated the intraclass coefficient (ICC) with its 95% confidence interval (CI) as an indicator of the overall agreement between the three reviewers for each item. According to the scale proposed by Landis and Koch, the degree of agreement between 0.00 and 0.20 was considered poor, from 0.21 to 0.40 fair, from 0.41 to 0.60 moderate, from 0.61 to 0.80 substantial, and from 0.81 to 1.00 almost perfect [24].

We accepted p values of less than 0.05 as significant. We performed the analyses using SPSS version 22.0 (SPSS Inc., Chicago, IL, USA).

Results

Selection of updated clinical guidelines

The screening process is summarised in a flow diagram (Fig. 1). We initially identified 1465 references and excluded 1249 references after examining their titles and abstracts. We reviewed 216 full-text articles and excluded 156 references (Additional file 2). Finally, we included 60 updated CGs [25–84].

Characteristics of included clinical guidelines

Most institutions responsible for updating the CGs were North American (61.7%; 37/60) and scientific/professional societies (46.7%; 28/60) or public institutions (43.3%; 26/60) (Table 1). In total, 25 (41.7%; 25/60) of the included CGs addressed the management of a specific disease. Other CGs address solely the treatment (25.0%; 15/60), screening (15.0%; 9/60), diagnosis (11.7%; 7/60), or prevention (6.7%; 4/60) of a healthcare problem. The clinical area of the included CGs varied widely, with oncology (26.7%; 16/60) the most common.

Domain scores

Presentation of the updated CG

All of the included updated CGs could be distinguished from their predecessors since this was one of the eligibility criteria. The included CGs often used the term 'update', 'version', or the year of publication (i.e. 2015) in their title (Table 2, Fig. 2).

More than half of the updated CGs included the rationale for updating (61.7%; 37/60), described changes in the scope and purpose between the updated CG and its predecessor (56.7%; 34/60), and reported the reviewed sections (66.7%; 40/60) (Table 2, Fig. 2).

At the recommendation level, 26.7% (16/60) of the included CGs clearly labelled the recommendations as new, modified, or not changed, and 38.3% (23/60) justified the changes. The justifications for changes commonly included a description of the new evidence that triggered the change in the recommendation and the changes between the new and old version of the recommendations (Table 2, Fig. 2).

The median score of the presentation domain on a 10-point scale was 5.8 (range 1.7 to 10), and the agreement among the three reviewers was adequate (ICC 0.854; 95% CI 0.701 to 0.941) (Table 3).

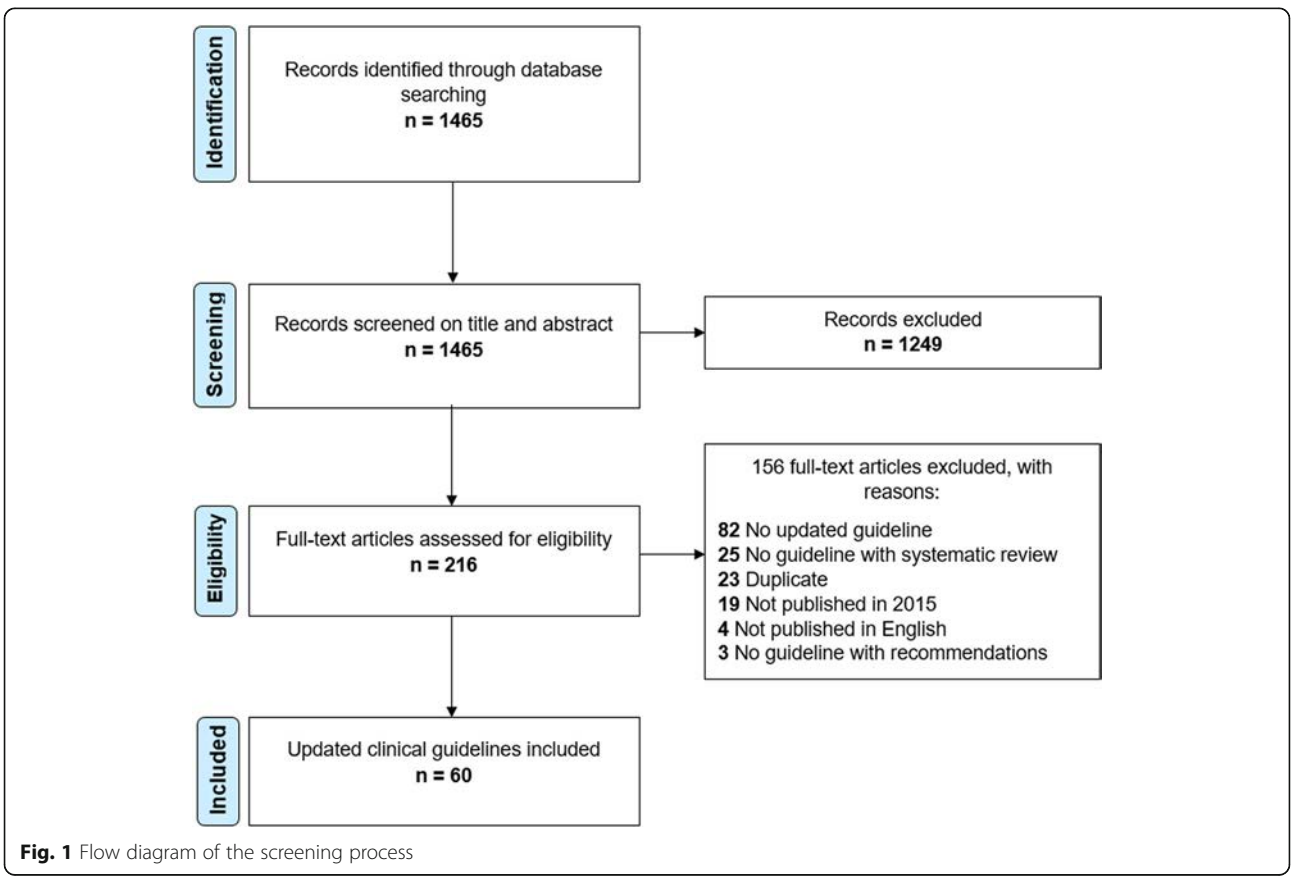


Fig. 1 Flow diagram of the screening process

Table 1 Characteristics of the updated clinical guidelines

	n (%)
Institution	
Country	
– North America	37 (61.7)
– Europe	17 (28.3)
– Asia	4 (6.7)
– International	2 (3.3)
Type of organisation	
– Scientific/professional society	28 (46.7)
– Public institution	26 (43.3)
– Other (Federal institute, NGO)	6 (10.0)
Updated clinical guidelines	
Scope	
– Management	25 (41.7)
– Treatment	15 (25.0)
– Screening	9 (15.0)
– Diagnosis	7 (11.7)
– Prevention	4 (6.7)
Health topic	
– Oncology	16 (26.7)
– Public health	5 (8.3)
– Internal medicine	3 (5.0)
– Mental health	3 (5.0)
– Others	33 (55.0)

Editorial independence

Almost all included CGs described the panel participants in the updated version (95.0%; 57/60) and their respective conflicts of interest (96.7%; 58/60) (Table 2, Fig. 2). However, half of the updated CGs did not report the entity and/or the role of the funding body that financed the updated version (50.0%; 30/60).

The median score of the editorial independence domain on a 10-point scale was 8.3 (range 3.3 to 10), and the agreement among the three reviewers was adequate (ICC 0.724; 95% CI 0.534 to 0.860) (Table 3).

Methodology of the updating process

Most of the included CGs reported the methods used for searching and identifying new evidence (81.7%; 49/60), selecting the evidence (78.3%; 47/60), and assessing the quality of the included evidence (76.7%; 46/60) (Table 2, Fig. 2). However, the methods for synthesising the evidence (46.7%; 28/60), external review (38.3%; 23/60), implementing the changes of the updated version in practice (38.3%; 23/60), or updating the new version (40.0%; 24/60) were reported less often in the included CGs.

The median score of the methodology domain on a 10-point scale was 5.7 (range 0 to 10), and the agreement among the three reviewers was adequate (ICC 0.886; 95% CI 0.771 to 0.952) (Table 3).

Overall score

The median overall score on a 10-point scale was 6.3 (range 3.1 to 10), and the agreement among the three reviewers was adequate (ICC 0.880; 95% CI 0.749 to 0.952) (Table 3).

CGs developed by a European or International institution obtained a higher overall score compared to North American or Asian institutions ($p = 0.014$) (Table 4). No significant differences in the overall score were found between CG differing in the type of organisation, scope, or topic.

Discussion

Main findings

Our study is the first systematic assessment of the reporting of the updating process in updated CGs using CheckUp. The presentation and methodology domains were reported less completely than the editorial independence domain. Particularly, the items regarding the presentation and justification of the updating process at recommendation level and the methods used for evidence synthesis, external review, implementing, and future updating were poorly reported. Both the domains and overall scores of the included CGs were highly variable. We identified only two (3.3%) CGs with a perfect score (10-point overall score 10) [60, 74].

We observed an adequate ICC reliability between the three reviewers. The lowest ICC was found for the editorial independence domain, but the ICC domain score was still considered adequate. This was mainly due to some CGs that reported the panel participants and their conflicts of interest for those that were responsible for updating the CG; however, they failed to report the same information for those who were responsible for developing the preceding CG.

Our results in the context of previous research

Presentation of updated CGs

Previous research showed that there was no clear improvement in the reporting or methodological quality after updating systematic reviews [85]. Similarly, Hasenfield et al. found that updated CGs were of worse methodological quality compared to their previous version [86]. Few studies have evaluated the optimal presentation formats of CGs in general [87, 88]. Similarly, regarding the updating process of CGs, a wide variability in the formats used to present updated recommendations has been reported by our group [17]. In the field of systematic reviews, Newberry et al. [89] evaluated different formats for presenting

Table 2 Item scores

	Updated CGs reporting each item *n* (%)
Presentation of the updated clinical guideline	
Item 1: The updated version can be distinguished from the previous version of the clinical guideline.	60 (100)
Item 2: The rationale for updating the clinical guideline is reported.	37 (61.7)
Item 3: Changes in the scope and purpose between the update and the previous version are described and justified.	34 (56.7)
Item 4: The sections reviewed in the updating process are described.	40 (66.7)
Item 5: Recommendations are clearly presented and labelled as new, modified, or not changed. Deleted recommendations are clearly noted.	16 (26.7)
Item 6: Changes in recommendations are reported and justified.	23 (38.3)
Editorial independence	
Item 7: The panel participants in the updated version are described.	57 (95.0)
Item 8: Disclosures of interest of the group responsible for the updated version are recorded.	58 (96.7)
Item 9: The role of the funding body for the updated version is identified and described.	30 (50.0)
Methodology of the updating process	
Item 10: The methods used for searching and identifying new evidence in the updating process are described.	49 (81.7)
Item 11: The methods used for evidence selection in the updating process are described.	47 (78.3)
Item 12: The methods used to assess the quality of the included evidence in the updating process are described.	46 (76.7)
Item 13: The methods used for evidence synthesis in the updating process are described.	28 (46.7)
Item 14: The methods used for external review of the updated version are described.	23 (38.3)
Item 15: The methods and plan for implementing the changes of the updated version in practice are described.	23 (38.3)
Item 16: The plan and methods for updating the new version in the future are reported.	24 (40.0)

One guideline is rated as not applicable

the results of updated systematic reviews. One of their conclusions was that different interest groups have different information needs. For example, health managers preferred to have access to all data and the analysis of a systematic review (the original and the updated), whereas clinicians prefer a synthesis that clearly shows what has been changed [89].

In our study, we have identified that, in particular, the presentation of updated recommendations is not optimal, with recommendations often not presented or not clearly labelled as new, modified, or not changed. This might confuse readers who might not be able to identify which recommendations are updated and which ones remain identical. Additionally, the modifications conducted in recommendations are often not described nor justified.

Reporting the editorial independence

The same principle regarding editorial independence for developing new CGs should be applied to the updating process [90]. Previous studies, in which the quality of CGs was reviewed with the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument, have observed low scores in the domain of editorial independence [91, 92]. We found similar results for the source of funding. However, most of the included updated CGs in our study reported the panel members and their conflicts of interest.

Reporting the methodology of the updating process

Until now, most of the methodological research regarding the updating process of CGs concerned the identification and assessment of new evidence (described commonly as the surveillance process) [18]. However, the complete updating process, including the presentation and justification of the updating process at recommendation level, has received less attention. CG developers possibly assume that the complete updating process is equal to the development process of the initial CG [19]. This could explain why the items that have a certain overlap with the development process (i.e. search strategy, evidence selection, and quality assessment) have higher scores compared to the updating items that are methodologically different from the development process (i.e. synthesis, external review, implementation of changes, and updating in the future) of the initial CG. Although the methods for developing CGs evolve rapidly [93], the updating process still does not follow this progress correspondingly [18, 19, 94, 95].

Item 1: The updated version can be distinguished from the previous version of the clinical guideline
"Clinical Practice Guideline for the Treatment of Obstructive Sleep Apnea and Snoring with Oral Appliance Therapy: An update for 2015". [22]

Item 2: The rationale for updating the clinical guideline is reported
"The last American College of Emergency Physicians (ACEP) clinical policy addressing the use of IV tPA for acute ischemic stroke was approved in 2012. Since then, changes to the ACEP clinical policies development process have been implemented, the grading form used to rate published research have continued to evolve, and new research articles have been published". [28]

Item 3: Changes in the scope and purpose between the update and the previous version are described and justified
"This document updates the treatment guideline with the reappraisal of previously assessed treatment options and new recommendations for novel agents. Evidence surrounding the clinical management of IPF is rapidly evolving, and it is intended that future iterations of the 2011 guideline dealing with questions related to diagnosis, genetics, and other new questions will be made available promptly."[38]

Item 4: The sections reviewed in the updating process are described
"This is a partial update of the 2011 clinical guideline on Anaemia Management in Chronic Kidney Disease.
The sections new or updated in 2015 are:

- *Guideline development group and scope*
- *Methodology*
- *Diagnostic tests for the prediction of response to iron therapy*
- *Concurrent illness*
- *Iron therapies*
- *Treatment of ESA resistance*

All other sections and recommendations from the 2011 guideline remain unchanged." [69].

Item 5: Recommendations are clearly presented and labelled as new, modified, or not changed. Deleted recommendations are clearly noted
"There is insufficient evidence to recommend for or against the use of pharmacotherapy in the treatment of cannabis use disorder. (Reviewed, new-added)" [49].

Item 6: Changes in recommendations are reported and justified
"All patients who receive highly emetogenic chemotherapy regimens (including anthracycline plus cyclophosphamide) should be offered a three-dug combination of an NK_1 receptor antagonist, a 5-HT_3 receptor antagonist, and dexamethasone. The oral combination of NEPA plus dexamethasone is an additional treatment option in this setting. The remaining recommendations from the 2011 ASCO guideline are unchanged pending a full update. The full set of recommendations is listed in the Bottom Line Box." [76].

Item 7: The panel participants in the updated version are described
Guideline development group members [2006]

Names	*Roles*

Guideline development group members [2011]

Names	*Roles*

Guideline development group members [2015]

Names	*Roles*

[69]

Item 8: Disclosures of interest of the group responsible for the updated version are recorded
Declarations of interest 2015.

GDG Member	Interest declared	Declared when?	Type of interest	Decision taken

Item 9: The role of the funding body for the updated version is identified and described
"This Clinical Practice Guideline was funded exclusively by the American Academy of Orthopaedic Surgeons who received no funding from outside commercial sources to support the development of this document." [23].

Item 10: The methods used for searching and identifying new evidence in the updating process are described
"Published literature was retrieved through searches of Medline and The Cochrane Database from January 1994 to January 2015 using appropriate controlled vocabulary (e.g. contraception, sexuality, sexual health) and key words (e.g. contraception, family planning, hormonal contraception, emergency contraception). Searches were updated on a regular basis and incorporated in the guideline to June 2015. Grey (unpublished) literature was identified through searching the websites of health technology assessment and health technology-related agencies, clinical practice guideline collections, clinical trial registries, and national and international medical speciality societies" [42].

Item 11: The methods used for evidence selection in the updating process are described.
"Studies reporting total and false positives as well as total and false negatives for the use of CT with intravenous contrast and aortography were included for further review." [50].

Item 12: The methods used to assess the quality of the included evidence in the updating process are described.
"For each recommendation, we provided the quality of the supporting evidence. According to GRADE, we classified the quality of evidence into 4 categories: high, moderate, low, and very low. The quality of evidence reflects the extent to which a guideline panel's confidence in an estimate of the effect was adequate to support a particular recommendation" [48].

Item 13: The methods used for evidence synthesis in the updating process are described.
"Evidence statements are summary statements that are presented after the GRADE profiles, summarising the key features of the clinical effectiveness evidence presented. The wording of the evidence statements reflects the certainty or uncertainty in the estimate of effect. The evidence statements are presented by outcome and encompass the following key features of the evidence:

- *the number of studies and the number of participants for a particular outcome*
- *a brief description of the participants*
- *an indication of the direction of effect (if one treatment is beneficial or harmful compared to the other, or whether there is no difference between the 2 tested treatments)."* [69].

Item 14: The methods used for external review of the updated version are described.
"On completion, the draft guideline was reviewed by external reviewers. It was also posted on the MoH Malysia official website for feedback from any interested parties. The draft was finally presented to the Technical Advisory Committee for CPG, and the HTA and CPG Council MoH Malaysia for review and approval." [65].

Item 15: The methods and plan for implementing the changes of the updated version in practice are described.
"The following measures have or will be undertaken to disseminate and aid implementation of the guidelines:

- *E-publication on the BHIVA website and the journal HIV medicine.*
- *Publication in HIV Medicine.*
- *Shortened version detailing concise summary of recommendations.*
- *Shortened version for BHIVA guidelines app.*
- *E-learning module accredited for CME.*
- *Educational slide set to support local and regional educational meetings.*
- *National BHIVA audit programme."* [43].

Item 16: The plan and methods for updating the new version in the future are reported.
"A formal review of the guideline will be conducted at the Annual Provincial Meeting in 2017. If critical new evidence is brought forward before that time, however, the guideline working group members will revise and update the document accordingly." [47].

Fig. 2 Reporting examples of the included updated CGs

Table 3 Domains, overall, and agreement scores

	Score[a] median (range)	Agreement ICC (95% CI)
Domain		
– Presentation of the updated CG	5.8 (1.7–10)	0.854 (0.701–0.941)
– Editorial independence	8.3 (3.3–10)	0.724 (0.534–0.860)
– Methodology of the updating process	5.7 (0–10)	0.886 (0.771–0.952)
Overall	6.3 (3.1–10)	0.880 (0.749–0.952)

[a]10-point scale (10 as the best possible score)
ICC intraclass coefficient, *CI* confidence interval

When updating CGs, developers need to pay special attention to the implementation implications of the changes introduced in updated CGs [96]. This can be done by exploring facilitators and barriers, by developing supporting materials, or by providing audit criteria [97]. Recently, GRADE has published Evidence to Decision frameworks to support developers to systematically consider this aspect and other criteria [98]. As living CGs become more common practice [99], developers will need to assess to what extent more frequent changes in recommendations impact their implementability and optimisation of patient care.

Table 4 Overall scores stratified by characteristics of the updated clinical guidelines

	Overall score[a] median (range)	*p value*
Institution		
Country		
– Europe	8.1 (4.4–10.0)	0.014
– International	7.8 (6.9–8.8)	
– Asia	5.6 (3.8–6.3)	
– North America	5.6 (3.1–8.1)	
Type of organisation		
– Public institution	6.3 (3.1–10.0)	0.617
– Scientific/professional society	6.3 (3.1–8.8)	
– Other (Federal institute, NGO)	4.4 (3.8–8.1)	
Updated clinical guidelines		
Scope		
– Diagnosis	8.1 (5.0–9.4)	0.097
– Prevention	5.6 (4.4–6.3)	
– Management	6.3 (3.1–10.0)	
– Treatment	6.3 (4.4–8.8)	
– Screening	3.8 (3.1–8.1)	
Health topic		
– Mental health	6.9 (5.0–8.1)	0.099
– Oncology	6.3 (3.8–9.4)	
– Internal medicine	6.3 (5.6–8.1)	
– Public health	3.8 (3.1–3.8)	
– Others	6.3 (3.1–10.0)	

[a]10-point scale (10 as the best possible score)

Strengths and limitations

Our study has several strengths. We followed a rigorous and transparent approach and developed a protocol that is available from the authors on request. Additionally, three reviewers independently conducted the assessment of the included CGs and adequate agreement was found.

Our study has some limitations. It is possible that we did not identify all updated CGs that would meet our inclusion criteria due to suboptimal indexing of CGs in biomedical databases, which may limit the representativeness of the results. Additionally, one eligibility criterion was also an item from the checklist, which might have led to the inclusion of more high-quality updated CGs. Consequently, our results might be an overestimate, and the actual reporting be actually worse than our findings.

Implications for practice and research

When CG developers are interested in updating CGs, we suggest firstly assessing the quality of CGs using the AGREE II instrument. After that, we suggest to (1) prioritise the update of high-quality CGs or (2) improve the methodological quality of the CG during the updating process. After the updating process, CG developers can assess the reporting of the updating process using CheckUp. Consequently, when both the AGREE II and CheckUp instruments are properly applied, developers will have a complete and detailed overview of the quality of the developing and updating processes. Afterwards, if applicable, the prioritisation process of updating CGs can be conducted [11].

There is currently no gold standard for updating CGs [18, 19, 94, 95]. Although CheckUp does not evaluate the quality of the updating process, CG developers can use it to inform their updating processes. Additionally, CheckUp can be used by interested CG users to assess whether updated CGs are in alignment with the CheckUp items, and editors of scientific journals that publish updated CGs may request the completion of CheckUp from the CG authors [20].

It would be relevant to monitor the use and the impact of CheckUp in the updating CG field over the next few years, potentially using this study as a baseline evaluation before the publication of CheckUp. Finally, we invite users to comment on the items and the usability of CheckUp contacting the corresponding author of this publication.

We suggest users of CheckUp to assess the reporting of the updating process in updated CGs by at least three calibrate reviewers. We involved three reviewers for convenience to avoid ties. Further examinations of CheckUp are required to determine if the inter-observer agreement between two reviewers would be adequate. Clinical expertise regarding the clinical area of the CG is not required; however, methodological comprehension on the updating process of CGs is highly desirable. To facilitate understanding of the domain scores and overall scores, we have transformed the domain and overall scores to a 10-point scale score.

Conclusions

The reporting of the updating process in updated CGs is suboptimal. Presentation of updated CGs and the methodology of the updating process where areas where more work is needed. We advise CG developers to use CheckUp to improve the reporting of updated CGs. CheckUp can also be used to assess the updating process in updated CGs and as a blueprint that could be used to inform specific updating methods and reporting strategies.

Additional files

Additional file 1: Literature search strategy. (DOCX 27 kb)

Additional file 2: Excluded full text references including reason for exclusion. (DOCX 52 kb)

Abbreviations

AGREE: Appraisal of Guidelines for Research and Evaluation; CG: Clinical guideline; CheckUp: Checklist for the Reporting of Updated Guidelines; CI: Confidence interval; G-I-N: Guidelines International Network; ICC: Interclass Correlations Coefficient

Acknowledgements

Robin W. M. Vernooij is a doctoral candidate at the Paediatrics, Obstetrics and Gynaecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

RV, LM, MB, and PA-C contributed to conceiving the study. All authors designed the study. RV conducted the searches. RV, IDF, LHA, and MHFP contributed to screening and extracting the data. RV, LM, and PA-C contributed to writing the manuscript. All authors contributed to commenting and editing of review drafts. All authors read and approved the final manuscript.

Funding

This project is funded by research grants from the Instituto de Salud Carlos III through the project 'PI14/02006' (co-funded by European Regional Development Fund/European Social Fund 'Investing in your future'). LMG is funded by a Juan Rodés contract from the Instituto de Salud Carlos III (JR15/00004). PAC is supported by a Miguel Servet investigator contract from the Instituto de Salud Carlos III (CPII15/0034). The funders had no role in the study design, data collection and analysis, interpretation of data, and writing the manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain. [2]Department of Health Research Methods, Evidence and Impact; McMaster University, Hamilton, Canada. [3]Department of Pediatrics, University of Antioquia, Medellin, Colombia. [4]University Medical Center Utrecht, Utrecht, The Netherlands. [5]Department of Oncology, McMaster University, Hamilton, Canada. [6]CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain.

Received: 2 June 2017 Accepted: 3 October 2017
Published online: 12 October 2017

References

1. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. Ann Intern Med. 2011;154(11):774–5.
2. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q. 2016;94(3):485–514.
3. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 2010;7:e1000326.
4. Martínez García L, Sanabria AJ, García Alvarez E, Trujillo-Martín MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ. 2014;186(16):1211–9.
5. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. J Clin Epidemiol. 2014;67(1):52–5.
6. Neuman MD, Goldstein JN, Cirullo MA, Schwartz JS. Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations. JAMA. 2014;311(20):2092–100.
7. Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. Updating clinical practice recommendations: is it worthwhile and when? Int J Technol Assess Health Care. 2012;28(1):29–35.
8. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001; 286(12):1461–7.
9. Martínez García L, Pardo-Hernandez H, McFarlane E, Sanabria AJ, Sparrow K, Alonso-Coello P, the G-I-N updating guidelines working group OBO. UpGlossary: Guidance on terminology and definitions for updating clinical guidelines. Cape Town: Global Evidence Summit 2017.
10. Agbassi C, Messersmith H, McNair S, Brouwers M. Priority-based initiative for updating existing evidence-based clinical practice guidelines: the results of two iterations. J Clin Epidemiol. 2014;67(12):1335–42.
11. Martínez García L, Pardo H, Superchi C, Niño de Guzman E, Ballesteros M, Ibargoyen Roteta N, et al. Updating health decision-making tools: a systematic review of prioritisation processes. J Clin Epidemiol. 2017;86: 11–24.
12. Martinez Garcia L, Sanabria AJ, Araya I, Lawson J, Solà I, Vernooij RW, et al. Efficiency of pragmatic search strategies to update clinical guidelines recommendations. BMC Med Res Methodol. 2015;15:57.
13. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Assessing the need to update prevention guidelines: a comparison of two methods. Int J Qual Health Care. 2004;16(5):399–406.

14. Johnston ME, Brouwers MC, Browman GP. Keeping cancer guidelines current: results of a comprehensive prospective literature monitoring strategy for twenty clinical practice guidelines. Int J Technol Assess Health Care. 2003;19(4):646–55.
15. Eccles M, Rousseau N, Freemantle N. Updating evidence-based clinical guidelines. J Health Serv Res Policy. 2002;7(2):98–103.
16. Parmelli E, Papini D, Moja L, Bandieri E, Belfiglio M, Ciccone G, et al. Updating clinical recommendations for breast, colorectal and lung cancer treatments: an opportunity to improve methodology and clinical relevance. Ann Oncol. 2011;22(1):188–94.
17. Martinez Garcia L, McFarlane E, Barnes S, Sanabria AJ, Alonso-Coello P, Alderson P. Updated recommendations: an assessment of NICE clinical guidelines. Implement Sci. 2014;9:72.
18. Martinez Garcia L, Arevalo-Rodriguez I, Sola I, Haynes RB, Vandvik PO, Alonso-Coello P, et al. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci. 7:109.
19. Vernooij RW, Sanabria AJ, Sola I, Alonso-Coello P, Martínez GL. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci. 2014;9:3.
20. Vernooij RW, Alonso-Coello P, Brouwers M, CheckUp Panel. Reporting items for updated clinical guidelines: checklist for the reporting of updated guidelines (CheckUp). PLoS medicine. 2017;14: e1002207.
21. Grupo de trabajo para la actualización del Manual de Elaboración de GPC. Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud. Actualización del Manual Metodológico. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad - GuíaSalud; 2016.
22. Eady EA, Layton AM, Sprakel J, BWM A, Fedorowicz Z, van Zuuren EJ. AGREE II assessments of recent acne treatment guidelines: how well do they reveal trustworthiness as defined by the Institute of Medicine (IOM) criteria? Br J Dermatol. 2017. doi:10.1111/bjd.15777.
23. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
25. American Academy of Dental Sleep Medicine/American Academy of Sleep Medicine. Clinical practice guideline for the treatment of obstructive sleep apnea and snoring with oral appliance therapy: an update for 2015:2015.
26. American Academy of Orthopaedic Surgeons. American Academy of Orthopaedic Surgeons clinical practice guideline on the treatment of pediatric diaphyseal femur fracture; 2015.
27. American Academy of Otolaryngology–Head and Neck Surgery. Clinical practice guideline (update): adult sinusitis; 2015.
28. American Academy of Sleep Medicine. Clinical practice guideline for the treatment of intrinsic circadian rhythm sleep-wake disorders: advanced sleep-wake phase disorder (ASWPD), delayed sleep-wake phase disorder (DSWPD), non-24-hour sleep-wake rhythm disorder (N24SWD), and irregular sleep-wake rhythm disorder (ISWRD). An update for 2015; 2015.
29. American Cancer Society. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society; 2015.
30. American College of Chest Physicians. Somatic cough syndrome (previously referred to as psychogenic cough) and tic cough (previously referred to as habit cough) in adults and children: CHEST guideline and Expert Panel report; 2015.
31. American College of Emergency Physicians. Clinical policy: use of intravenous tissue plasminogen activator for the management of acute ischemic stroke in the emergency department; 2015.
32. American Dental Association. The use of prophylactic antibiotics prior to dental procedures in patients with prosthetic joints: evidence-based clinical practice guideline for dental practitioners—a report of the American Dental Association Council on Scientific Affairs; 2015.
33. American Geriatrics Society. American Geriatrics Society 2015 updated Beers Criteria for potentially inappropriate medication use in older adults; 2015.
34. Stav W. Occupational therapy Occupational therapy practice guidelines for driving and community mobility for older adults. Bethesda (MD)Ñ American Occupational Therapy Association (AOTA); 2015.
35. American Psychiatric Association. Practice guidelines for the psychiatric evaluation of adults, third edition; 2015.
36. American Society of Anesthesiologists. Practice guidelines for perioperative blood management: an updated report by the American Society of Anesthesiologists Task Force on Perioperative Blood Management; 2015.
37. American Society of Anesthesiologists. Practice advisory on anesthetic care for magnetic resonance imaging: an updated report by the American Society of Anesthesiologists Task Force on Anesthetic Care for Magnetic Resonance Imaging; 2015.
38. American Society of Clinical Oncology. Recommendations for the use of WBC growth factors: American Society of Clinical Oncology clinical practice guideline update; 2015.
39. American Society of Clinical Oncology. Systemic therapy for stage IV non-small-cell lung cancer: American Society of Clinical Oncology clinical practice guideline update; 2015.
40. American Society of Clinical Oncology. Use of biomarkers to guide decisions on systemic therapy for women with metastatic breast cancer: American Society of Clinical Oncology clinical practice guideline; 2015.
41. Raghu G, Rochwerg B, Zhang Y, Garcia CA, AzumaA BJ, et al. An official ATS/ERS/JRS/ALAT clinical practice guideline: treatment of idiopathic pulmonary fibrosis. An update of the 2011 clinical practice guideline. Am J Respir Crit Care Med. 2015;192(2):e3–19.
42. American Urological Association. Castration-resistant prostate cancer: AUA guideline; 2015.
43. Atkins DL, Berger S, Duff JP, Gonzales JC, Hunt EA, Joyner BL, et al. Part 11: pediatric basic life support and cardiopulmonary resuscitation quality: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care (reprint). Circulation. 2015; 132(18 Suppl 2):S519–25.
44. Black A, Guilbert E, Co Authors, Costescu D, Dunn S, Fisher W, et al. Canadian contraception consensus (part 2 of 4). J Obstet Gyneacol Can 2015;37(11):1033-1039.
45. Black A, Guilbert E, Co Authors, Costescu D, Dunn S, Fisher W, et al. Canadian contraception consensus (part 1 of 4). J Obstet Gyneacol Can. 2015;37(10):936-942.
46. British HIV Association. British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2015; 2015.
47. Canadian Task Force on Preventive Health Care. Recommendations for prevention of weight gain and use of behavioural and pharmacological interventions to manage overweight and obesity in adults in primary care; 2015.
48. CancerControl Alberta. Systemic therapy for well and moderately differentiated unresectable pancreatic neuroendocrine tumours; 2015.
49. CancerControl Alberta. Systemic therapy for unresectable stage III or metastatic cutaneous melanoma; 2015.
50. CancerControl Alberta. Merkel cell carcinoma; 2015.
51. Belgian Health Care Knowledge Centre. What are the recommended clinical assessment and screening tests during pregnancy?; 2015.
52. Department of Defense/Department of Veterans Affairs. VA/DoD clinical practice guideline for the management of substance use disorders; 2015.
53. Eastern Association for the Surgery of Trauma. Evaluation and management of blunt traumatic aortic injury: a practice management guideline from the Eastern Association for the Surgery of Trauma; 2015.
54. National Institute for Health and Care Excellence. Tuberculosis (CG117); 2015.
55. National Institute for Health and Care Excellence. Diabetes in pregnancy (CG63); 2015.
56. National Institute for Health and Care Excellence. Diabetes in children and young people (CG); 2015.
57. National Institute for Health and Care Excellence. Type 1 diabetes (CG15); 2015.
58. National Institute for Health and Care Excellence. Referral for suspected cancer (CG27); 2015.
59. National Institute for Health and Care Excellence. Coeliac disease (CG86); 2015.
60. National Institute for Health and Care Excellence. Type 2 diabetes: the management of type 2 diabetes (update) (CG66); 2015.
61. National Institute for Health and Care Excellence. Type 2 diabetes—footcare (CG10); 2015.
62. National Institute for Health and Care Excellence. Violence (CG25); 2015.
63. Siu AL, U.S. Preventive Services Task Force. Screening for iron deficiency anemia and iron supplementation in pregnant women to improve maternal health and birth outcomes: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2015;163(7):529–36.
64. Siu AL, U.S. Preventive Services Task Force. Screening for iron deficiency anemia in young children: USPSTF recommendation statement. Pediatrics. 2015;136(4):746–52.

65. Siu AL, U.S. Preventive Services Task Force. Screening for abnormal blood glucose and type 2 diabetes mellitus: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2015;163(11):861–8.
66. Siu AL, U.S. Preventive Services Task Force. Screening for high blood pressure in adults: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2015;163(10):778–86.
67. Siu AL, U.S. Preventive Services Task Force. Behavioral and pharmacotherapy interventions for tobacco smoking cessation in adults, including pregnant women: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2015;163(8):622–34.
68. Malaysia Health Technology Assessment Section (MaHTAS). Management of cervical cancer (second edition); 2015.
69. Malaysia Health Technology Assessment Section (MaHTAS). Management of type 2 diabetes mellitus (5th edition); 2015.
70. Malaysia Health Technology Assessment Section (MaHTAS). Management of dengue infection in adults (third edition); 2015.
71. Miyazaki M, Yoshitomi H, Miyakawa S, Uesaka K, Unno M, Endo I, et al. Clinical practice guidelines for the management of biliary tract cancers 2015: the 2nd English edition. J Hepatobiliary Pancreat Sci. 2015;22(4):249–73.
72. National Clinical Guideline Centre. Anaemia management in people with chronic kidney disease; 2015.
73. National Collaborating Centre for Nursing and Supportive Care. Irritable bowel syndrome in adults: diagnosis and management of irritable bowel syndrome in primary care; 2015.
74. National Collaborating Centre for Women. Diabetes (type 1 and type 2) in children and young people: diagnosis and management; 2015.
75. National Institute for Health and Care Excellence. Prophylaxis against infective endocarditis: antimicrobial prophylaxis against infective endocarditis in adults and children undergoing interventional procedures; 2015.
76. National Kidney Foundation. KDOQI clinical practice guideline for hemodialysis adequacy: 2015 update; 2015.
77. Scottisch Intercollegiate Guidelines Network (SIGN). Diagnosis and management of epilepsy in adults (SIGN CPG 143); 2015.
78. Lyman GH, Bohlke K, Khorana AA, Kuderer NM, Lee AY, Arcelus JI, et al. Venous thromboembolism prophylaxis and treatment in patients with cancer: American Society of Clinical Oncology practice guideline update 2014. J Clin Oncol. 2015;33(6):654–6.
79. Hesketh PJ, Bohlke K, Lyman GH, Basch E, Chesney M, Clark-Snow RA, et al. Antiemetics: American Society of Clinical Oncology focused guideline update. J Clin Oncol. 2016;34(4):381–6.
80. Hwang JP, Somerfield MR, De A-J, Cryer DR, Feld JJ, Kramer BS, et al. Hepatitis B virus screening for patients with cancer before therapy: American Society of Clinical Oncology provisional clinical opinion update. J Clin Oncol. 2015;33(19):2212–20.
81. Allegra CJ, Rumble RB, Hamilton SR, Mangu PB, Roach N, Hantel A, et al. Extended RAS gene mutation testing in metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy: American Society of Clinical Oncology provisional clinical opinion update 2015. J Clin Oncol. 2016;34(2):179–85.
82. Park IU, Introcaso C, Dunne EF. Human papillomavirus and genital warts: a review of the evidence for the 2015 Centers for Disease Control and Prevention sexually transmitted diseases treatment guidelines. Clin Infect Dis. 2015;61(Suppl 8):S849–55.
83. Prica A, Baldassarre F, Hicks LK, Imrie K, Kouroukis TC, Cheung M, et al. Rituximab in lymphoma and chronic lymphocytic leukemia: a clinical practice guideline, version 3; 2015.
84. Registered Nurses' Association of Ontario. Person- and family-centred care; 2015.
85. Shea B, Boers M, Grimshaw JM, Hamel C, Bouter LM. Does updating improve the methodological and reporting quality of systematic reviews? BMC Med Res Methodol. 2006;6:27.
86. Hasenfeld R, Shekelle PG. Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR) guidelines with those published subsequently. Qual Saf Health Care. 2003;12(6):428–34.
87. Brandt L, Vandvik PO, Alonso-Coello P, Akl EA, Thornton J, Rigua D, et al. Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial. BMJ Open. 2017;7(2):e011569.
88. Kastner M, Bhattacharyya O, Hayden L, Makarski J, Estey E, Durocher L, et al. Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review, 498. J Clin Epidemiol. 68(5):–509.
89. Newberry SJ, Shekelle PG, Vaiana M, Motala A. Reporting the findings of updated systematic reviews of comparative effectiveness: how do users want to view new information? Rockville: Agency for Healthcare Research and Quality; 2013.
90. Schunemann HJ, Al-Ansary LA, Forland F, Kersten S, Komulainen J, Kopp IB, et al. Guidelines international network: principles for disclosure of interests and management of conflicts in guidelines. Ann Intern Med. 2015;163(7):548–53.
91. Armstrong JJ, Goldfarb AM, Instrum RS, MacDermid JC. Improvement evident but still necessary in clinical practice guideline quality: a systematic review. J Clin Epidemiol. 2017;81:13–21.
92. Alonso-Coello P, Irfan A, Sola I, Gich I, Delgado-Noguera M, Rigau D, 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):e58.
93. Murad MH. Clinical practice guidelines: a primer on development and dissemination. Mayo Clin Proc. 2017;92(3):423–33.
94. Alonso-Coello P, Martinez Garcia L, Carrasco JM, Solà I, Qureshi S, Burgers JS, et al. The updating of clinical practice guidelines: insights from an international survey. Implement Sci. 2011;6:107.
95. Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol. 2014;67(1):33–45.
96. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [internet]. London: National Institute for Health and Care Excellence (NICE; 2015.
97. Chan WV, Pearson TA, Bennett GC, Cushman WC, Gaziano TA, Gorman PN, et al. ACC/AHA special report: clinical practice guideline implementation strategies: a summary of systematic reviews by the NHLBI implementation science work group: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. J Am Coll Cardiol. 2017;69(8):1076–92.
98. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: clinical practice guidelines. BMJ. 2016;i2089:353.
99. Martínez García L, Pardo-Hernandez H, Sanabria A, Alonso-Coello P, Pregnancy Clinical Guideline Updating Working Group. Continuous surveillance of a pregnancy clinical guideline: an early experience. Systematic Reviews. 2017;6:143.

6. DISCUSSION

6.1 Main results

This thesis is a compendium of three articles that aim to extend the knowledge and provide more guidance for updating CGs, including a SR about the guidance for the updating process described in methodological handbooks [116], the development process of CheckUp [117], and an evaluation of the reporting in updated CGs with CheckUp [118].

In the first study, we identified scarce guidance for updating CGs in the included methodological handbooks. For the most part, the handbooks restricted the guidance to *de novo* CG development. Methodological steps that are considered crucial for the updating process, including the identification of new scientific evidence, the assessment for the need of an update, and updating the recommendation are generally lacking or included solely a reference to *de novo* development [116]. This study has been published in the scientific journal "*Implementation Science*" (Impact factor 2016: 3:35, ranked 17 of 90 in category Health Care Sciences & Services, cited 54 times as in December 2017).

For the second study, we have constructed CheckUp through a rigorous development process. In these different development steps, an agreement and consensus was reached on what information regarding the updating process ought to be reported in an updated CG. CheckUp is the first reporting instrument about the updating process in CGs and covers an important field of research where more guidance was needed [117]. CheckUp is included in the EQUATOR library and is available in four languages: English, Spanish, Chinese, and Dutch (**Appendices VII to IX**) [97]. This study has been published in the scientific journal "*PLOS Medicine*" (Impact factor 2016: 11:86, ranked 7 of 155 in category General & Internal Medicine, cited 10 times in December 2017).

In our third study, we found suboptimal reporting of the updating process in updated CGs using CheckUp. The reporting of the presentation and justification of the updating process at recommendations level was often lacking. Additionally, the updating methods for evidence synthesis, external review, implementation of the updated recommendations into practice, and strategies for future updating were poorly reported. On the other hand, most CGs reported information about the editorial independence of the panel responsible for updating the CG. We found that, on average, CGs developed by European or international institutions had an overall higher score in CheckUp compared to North American or Asian institutions and agreement between the reviewers was considered excellent [118]. This study has been published in the scientific journal "*Implementation Science*" (Impact factor 2016: 3:35, ranked 17 of 90 in category Health Care Sciences & Services, not yet cited in December 2017).

6.2 Our results in the context of previous research

The work we present in this thesis expands the knowledge about updating CGs and will provide more guidance to the CG panels in the updating process. So far, several components of the updating process had been investigated in previous studies, mostly focussing on the identification of new evidence and the assessment for the need of an update [63,67,71,73,78,79]. However, studies like ours were greatly needed in this field given the scarce amount of research and guidance available, reflected in an international survey of 39 CG institutions conducted by our team [64].

6.2.1 Methodological approaches to clinical guideline updating

The survey aforementioned showed that the updating practices were not standardised and could be more rigorous. Almost half of the participating organisations reported that they were planning to improve the updating process, yet many called for better methodological evidence before modifying their current approach [64]. Nevertheless, although multiple practice-

changing articles were published between 2011 (the year of publishing the survey) and 2014 (publication of our SR of methodological handbooks) [63,67,73], there seemed to be little improvements over time [64,116].

For instance, the reported time-frame for updating CGs was similar between the survey and our SR. In the former, almost two-third of the agencies updated their CGs every five years compared with, in the latter, three-quarters of the CG institutions that pursue a five-year time-interval between updating rounds [64,116]. Albeit it might not be feasible to determine one fixed time-interval that can be applied to every individual CG, given that some clinical areas are subject to more fast-pace changes than others, there seems to be consensus that the CGs recommendations commonly lose validity earlier than five years [67,70].

Another example of stagnation in methodological advances for CG updating is the process for the identification of new evidence. Whereas 40% of the survey respondents modified their original search strategy to improve the specificity of the search strategy, 31% of the methodological handbooks indicated to do the same [64,116]. Consequently, this might result in an inefficient use of resources since pragmatic search strategies are considered more efficient and feasible for identifying new evidence to assess the need for an update [63,78,79].

6.2.2 Methodological guidance for clinical guideline updating

As far as we know, solely one other study investigated the updating process described in methodological CG handbooks [119]. Becker *et al.* (2014) assessed 47 methodological handbooks, using a similar approach to our SR. The authors concluded, similar to our findings, that most CG handbooks did not provide a comprehensible and systematic approach in CG updating [119]. Combining the results of Becker *et al.* (2014) and the results of our SR, there seems to be poor guidance on updating CGs in methodological handbooks of CG institutions and there is a need of including more explicit and rigorous guidance in this field.

This could, consequently, lead to a more optimal updating of CGs, and ultimately, to more valid trustworthy CGs. [116,119].

Although there is a paucity of studies on the methodological guidance of CG handbooks, there are some additional studies published that can inform about other aspects related to CG updating. This mostly relates to the development process of *de novo* CGs. For example, Schünemann *et al.* (2006) examined the different components of developing *de novo* CGs, and developed 19 components (including priority setting, CG panel, CoI, group processes, defining important outcomes, methodology process, among others) that a CG institutions handbook should include to adequately guide a CG development panel [120]. Similarly, in the SR of Ansari *et al.* (2012), including 19 CG methodological handbooks, the CG development process was examined and the handbooks were ranked on the reporting of the methodological steps of the CG development process. In general, guidance for the *de novo* development of CGs was lacking in most examined handbooks [121]. In these evaluations, the development and implementation process of *de novo CGs* received most attention in the methodological handbooks, whereas the updating process was generally overlooked [122]. Reflecting on this overview of SRs of CG methodological handbooks, including the results of our SR, it is clear that the handbooks do not provide adequate guidance for neither the *de novo* development nor the updating process.

6.2.3 Reporting checklists for clinical guidelines

While the AGREE II, the G-I-N McMaster Guideline Development checklist, and the RIGHT statement focus on the development process of *de novo CGs*, they neglect the specific methodological aspects related to CG updating to a great extent (**Table 5**) [34,105,106]. The AGREE II and G-I-N McMaster checklist include some questions about the updating process of CGs, mainly related to the procedure and timeline for future CG updating [34,105], while the RIGHT statement includes none [106]. This is understandable given that although there

are some methodological similarities between developing *de novo* CGs and updating CGs (e.g. evidence selection, risk of bias assessment of the included studies), the updating process should be considered a different step in the life-cycle of CGs with distinct reporting requirements.

Table 5: Differences and similarities of the clinical guidelines reporting tools

Item	Objectives	Domains	Number of items	Items about updating
AGREE II	The purpose of the AGREE II, is to provide a framework to: 1. assess the quality of guidelines. 2. provide a methodological strategy for the development of guidelines. 3. inform what information and how information ought to be reported in guidelines.	1. Scope and Purpose. 2. Stakeholder Involvement. 3. Rigour of Development. 4. Clarity of Presentation. 5. Applicability. 6. Editorial Independence.	23 items.	• A procedure for updating the guideline is provided.
G-I-N McMaster Checklist	The checklist is intended for use by guideline developers to plan and track the process of guideline development and to help the developers ensure that no key	1. Organization, Budget, Planning and Training. 2. Priority Setting. 3. Guideline Group Membership. 4. Establishing Guideline Group Processes. 5. Identifying Target	146 items.	• Set a policy, procedure and timeline for routinely monitoring and reviewing whether the guideline needs to be updated. • Decide who will be responsible for routinely

steps are missed.	Audience and topic Selection. 6. Consumer and Stakeholder Involvement. 7. Conflicts of interest Consideration. 8. (PICO) Question Generation. 9. Considering Importance of Outcome and Interventions, Values, Preferences, and Utilities. 10. Deciding what Evidence to Include and Searching for Evidence. 11. Summarizing Evidence and Considering Additional Information. 12. Judging Quality, Strength or Certainty of a Body of Evidence. 13. Developing Recommendations and Determining their Strength. 14. Wording of Recommendations and of Considerations of Implementation, Feasibility, and Equity. 15. Reporting and Peer review. 16. Dissemination and Implementation. 17. Evaluation and Use.	monitoring the literature and assessing whether new significant evidence is available. • Set the conditions that will determine when a partial or a full update is required. • Make arrangement for guideline group membership and participation after completion of the guideline. • Plan the funding and logistics for updating the guideline in the future. Document the plan and proposed methods for updating the guideline to ensure they are followed.	

		18. Updating.		
RIGHT	The RIGHT statement includes a checklist of 22 items and aims to ensure that guidelines contain clear statements of why and how they were developed, and who was involved. The ultimate goal is that healthcare practitioners can better understand and implement the recommendations contained in the guideline	1. Basic information. 2. Background. 3. Evidence. 4. Recommendations. 5. Review and quality assurance. 6. Funding and declaration and management of interests. 7. Other information.	22 items.	None.

6.2.4. Reporting the updating process

The three domains included in CheckUp are the presentation of an updated CG, editorial independence, and the methodology of the updating process.

6.2.4.1 Presentation of an updated CG

In the field of *de novo* CG development, two studies investigated presentation formats of recommendations [123,124]. The study of Brandt *et al.* (2017), which is result of a collaboration of GRADE and MAGIC (MAking Grade the Irresistible Choice), in the context

of the DECIDE project (www.decide-collaboration.eu), tested a new CG presentation format with clinicians. The study found that approximately three-quarters of the physicians preferred a multi-layered presentation format rather than the standard narrative presentation format. This new format displays the CGs recommendations upfront with supporting information in additional layers [123]. Similarly, the SR of Kastner *et al.* (2015) investigated the factors associated with the successful implementation of CGs and concluded that the formulation of recommendations in a simple, clear, and persuasive language could help to ensure successful implementation [124].

Although the aforementioned studies concern the presentation format of *de novo* developed CGs and no studies to date have examined the ideal presentation format of recommendations in an updated CGs, the fifth item of CheckUp concerns the presentation and labelling of recommendations as new, modified, or not changed. Solely one study investigated the presentation format of recommendations in a sample of updated CGs [125]. Martínez García *et al.* (2014) evaluated the reporting of recommendations in a sample of updated CGs produced by NICE [125]. The presentation format for updated recommendations was fairly heterogeneous. In the field of SRs, previous research identified that different interest groups have different information needs regarding the updating process. For example, clinicians indicated that they prefer a clear overview of what exactly had changed [126]. The EtD frameworks might help the CG panel to use the evidence in a structured and transparent way to inform decisions in the context of clinical recommendations. This framework might help the CG panel to ensure consideration of key criteria that determine the strength and direction of the recommendation. Consequently, the rationale of the recommendations will become justified [61,62].

6.2.4.2 Editorial independence

The editorial independence of CG updating should be described with the same rigour as the *de novo* development [42]. Two previous SRs collected individual studies with AGREE II evaluations and found both low scores in the *de novo* developed CGs for the editorial independence domain [127,128]. These findings are slightly different from the results of our systematic assessment with CheckUp in updated CGs, where adequate reporting of the CG panel composition and their CoI was identified [118]. However, the funding disclosure was similarly poor reported in the *de novo* and updated CGs [118,127,128]. Although the reporting of the editorial independence is recognised to be crucial for the CGs transparency in both *de novo* development and updating of CGs, our study shows that there is room for improvement, especially with relation to the funding sources.

6.2.4.3 Methodology of the updating process

As reported by the first study of this thesis [116], and as confirmed by previous similar studies [62,63,67], there is no consensus on the common methodological approach of CG updating. However, at present there is a consensus that the CG updating process involves: a prioritisation process for deciding which CG, clinical question, or recommendation requires attention; a surveillance process, including the identification of new scientific evidence and the assessment for the need of an update; and the updating of the recommendations [63,65,67,71,73]. All these methodological steps are included in CheckUp, except for the prioritisation process. This stage has only recently become progressively more important [114,118], given the important resources involved in updating, this theme probably would merit to be included as an independent item and/or part of the CheckUp explanation materials.

6.2.5. Assessing the reporting of CGs

Up to now there have been no studies that assessed the reporting of a big sample of updated CGs. As such, the third study of this thesis fulfils this research gap. There are, however, several studies that assessed the reporting of *de novo* developed CGs using the AGREE II instrument. As previously mentioned, there is some overlap between the development process and updating process in CGs and, therefore, we can compare the methodological items of previous evaluations of *de novo* developed CGs with the AGREE II instrument, and the results of our systematic assessment of updated CGs with CheckUp. The SR of Armstrong *et al.* (2017) and Alonso-Coello *et al.* (2010) collected, appraised, and synthesised the results of individual studies with AGREE II assessments [127,128]. These SRs show that although the quality of reporting has substantially improved over time [128], there is suboptimal reporting for the stakeholder involvement, editorial independence, and applicability domains of the AGREE instrument [127]. As in the case of editorial independence, stakeholder involvement, and clarity of presentation there is room for improvement in both the development and updating process of CGs [118,127,128]. Previous research identified that updated CGs were in general of worse quality compared with the previous version of the CG [129].

Similarly, in the field of SRs, no improvement was identified between the updated version of SRs and their previous version [130]. These findings are consistent with the results of our systematic assessment, including 60 updated CGs, which concluded, that for most items, more than half of the included CGs reported inadequately the required information [118]. Therefore, we might conclude with caution that the reporting might be similar, if not worse, between the updated CG and its predecessor. However, this should be further researched by a study that compares the methodological items (e.g. reporting of search strategy, evidence selection, or quality assessment) that are similar between AGREE II and CheckUp between a sample of updated CG and their respective predecessor (i.e. the *de novo* developed CG).

6.2.6. Assessment of CGs using CheckUp

So far, one study, outside our research group, has used CheckUp to assess the reporting of the updating process in CGs regarding acne treatment [131]. The authors found similar results to our systematic assessment of updated CGs [118], which consisted of suboptimal reporting in the items regarding the rationale for updating, the labelling of recommendations, and justification of changes in recommendations. The reporting was assessed in two updated CGs in this study: one of the included CGs reported solely one of the items of CheckUp clearly and the other CG reported five items of CheckUp adequately [131].

Furthermore, two conference abstracts with a systematic assessment of the reporting in updated CGs with CheckUp were presented at the Global Evidence Summit of 2017 (i.e. G-I-N conference). Vasconcelos *et al.* (2017) presented the assessment of the reporting of two updated CGs assessed with CheckUp, showing adequate reporting of the updating process in 5 of the 16 items of CheckUp for both CGs [132]. Ram *et al.* (2017) conducted a systematic assessment of ten updated CG with CheckUp and observed that the items regarding the justification of changes in the recommendations, the reviewed sections, the labelling of recommendations, the plan and methods for updating in the future, and the rationale for updating the CG, were poorly reported. Over the full sample of included CGs, approximately half of the items were suboptimal reported (i.e. answered with no or unclear) [133]. Given that the average score of CheckUp ranges from approximately 6% to 50% adequate reporting, we can conclude that, together with the third study of this thesis, room for improvement in the reporting of the updating process has been identified [118,131-133].

6.3 Strengths and limitations

The strengths and limitations of the studies included in this thesis are listed in Table 5 and summarised narratively below (**Table 6**). The main strength of the studies included in the thesis concern a rigorous methodology, including the use of protocols and a combination of a

consensus survey and qualitative research. However, the main limitations of our body of research are that we might not have included all handbooks or updated CGs and that there is a potential overlap between CheckUp and the other CG instruments.

Table 6: Strengths and limitations of the included studies.

Article	Strengths	Limitations
Vernooij *et al.* (2014)	• Systematic and exhaustive search. • Duplicate independent eligibility and data extraction. • Case-report form developed *de novo*, reviewed and piloted iteratively.	• Potential risk of not capturing all methodological handbooks available. • Indirect information about the real updating practice of organisations.
Vernooij *et al.* (2017)	• Use of EQUATOR and Moher's criteria for the development. • Use of a consensus Delphi survey, and qualitative research to explore the perceptions of both CG users and CG methodologists. • Fairly strong overall consensus regarding the inclusion or exclusion of items	• Use of consensus methods and convenience samples of CG stakeholders • Potential overlap of items between CheckUp and other instruments or checklists (e.g. AGREE II). • Lack of CoI evaluation of the CheckUp panel.
Vernooij *et al.* (2017)	• Use of a protocol. • Rigorous and transparent methods. • Excellent agreement between appraisers.	• Potential risk of not capturing all relevant CGs. • One eligibility criterion was an item from the checklist, which might have led to an overestimation of the reporting completeness.

If we compare the development methodology of CheckUp with other reporting instruments regarding CGs, such as the AGREE II, G-I-N McMaster Guideline Development checklist, and the RIGHT statement, we followed for the development of CheckUp a similar or more

elaborated methodology. While the G-I-N McMaster Guideline Development checklist conducted a SR of research regarding the quality of CGs, described these in quality criteria, and consequently reviewed these criteria with the panel [105], we elaborated the development methodology for CheckUp with a qualitative part (semi-structured interviews) and a formal Delphi consensus survey. Similarly, the RIGHT statement, developed their instrument by a SR of CG quality criteria, a systematic assessment with an initial version of the checklist to CGs, and conducted a Delphi consensus survey to improve their instrument [106], which is similar to the development methodology of CheckUp, except that we conducted a qualitative assessment of the completeness of the checklist with semi-structured interviews and an external review with both CG users and methodologists.

6.4 Implications for clinical guideline institutions

Taking into account the results of our systematic assessment [116], we can conclude that the CG methodological handbooks do not provide adequate methodological guidance regarding the updating process at this current moment. Handbooks could benefit from a critical revision that includes providing sufficient guidance. This guidance should include the use of CheckUp as a blueprint for the reporting of updated CGs [117]. We advise CG institutions to use CheckUp as a guidance in their updating process, especially since we identified room for improvement with potentially a more explicit and rigorous guidance and standardisation of the terminology in the actual updating process of CG institutions [63,64,116]. Adherence to CheckUp could likely improve the comprehensiveness and transparency of the updating process, which in turn potentially could ultimately benefit patients, healthcare professionals, and other stakeholders [116].

Although CheckUp solely assesses the reporting of the updating process, it also includes methodology items for the updating process. Therefore, CG panels responsible for updating can use the methodology domain to assess whether all methodological aspects are being

considered [117]. However, this guidance should be interpreted with caution, since there is no such a thing as a gold standard for the complete updating. It is likely that the complete CG enterprise, including the state-of-the-art methodology for updating CG, will change in the future, therefore, CheckUp should be considered a dynamic tool that will be modified as needed.

CheckUp can be used by different stakeholders to assess the completeness of the reporting of the updating process. For example, researchers can assess the reporting of the updating process in a sample of updated CGs (e.g. CGs covering one clinical area). The full evaluation of an updated CG might be considered in conjunction with another CG reporting instrument, with a more specific focus on the development process. For example, an evaluation with the AGREE II instrument only allows seeing if a procedure for updating CGs is in place, however, CheckUp allows the user to assess if this methodology of the updating process is properly reported. Ideally, researchers that are interested in assessing the quality of reporting of a sample of CGs, should use a reporting instrument for the assessment of the development process (e.g. AGREE II), as well as CheckUp, to assess both updated CGs and newly developed CGs [34,117].

6.5 Implications for research

The results of our systematic assessment of the current reporting of the updating process can be considered the baseline level before the implementation of CheckUp [118]. There is some evidence, for example a before-and-after study of the CONSORT instrument, that shows that the publication of a reporting tool improves the reporting [102,134,135]. Therefore, considering the suboptimal reporting of the updating process identified in our systematic assessment [118], a second evaluation of the updating process in a few years might be conducted to evaluate the implementation process of CheckUp and examine whether the

reporting of the updating process improved over the last years. In addition, agreement between reviewers in the CheckUp assessment should be evaluated in other contexts [118].

We have found room for improvement in the actual updating practices described in methodological handbooks [116,119]. Additionally, in the survey of Alonso-Coello *et al.* (2011), several CG institutions indicated that they were waiting for more evidence to modify their updating process [64]. Therefore, further research is needed to conduct an update of the current updating practices to evaluate whether the publication of the new landmark studies regarding the updating process, including CheckUp, have introduced changes in the updating practices of CG institutions.

7. CONCLUSIONS

- Guidance for updating CGs is overall poorly described in CG methodological handbooks. The CG updating process could potentially benefit from a more explicit and rigorous guidance and standardisation of the terminology (article 1).
- CheckUp, a 16 item checklist that addresses: 1) the presentation of an updated CG; 2) editorial independence; and 3) the methodology of the updating process, might be used for the evaluation of the reporting in updated CGs, and for informing CG developers about reporting requirements of updated CGs (article 2).
- Adherence to the reporting items of CheckUp is likely to improve the comprehensiveness and transparency of the updating process, potentially benefiting patients, healthcare professionals, and other relevant stakeholders (article 2).
- The reporting of the updating process varies considerably in contemporary CGs and there is significant room for improvement. CG developers are advised to use CheckUp as a blueprint for the reporting of CG updating (article 3).

8. REFERENCES

1. McColl A, Smith H, White P, Field J. General practitioner's perceptions of the route to evidence based medicine: a questionnaire survey. BMJ 1998;316(7128):361-5.
2. O'Donnell CA. Attitudes and knowledge of primary care professionals towards evidence-based practice: a postal survey. Journal of evaluation in clinical practice 2004;10(2):197-205.
3. Heiwe S, Kajermo KN, Tyni-Lenne R, Guidetti S, Samuelsson M, Andersson IL, et al. Evidence-based practice: attitudes, knowledge and behaviour among allied health care professionals. International journal for quality in health care 2011;23(2):198-209.
4. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS medicine 2010;7(9):e1000326.
5. Tsay MY, Yang YH. Bibliometric analysis of the literature of randomized controlled trials. Journal of the Medical Library Association 2005;93(4):450-8.
6. Haynes B, Haines A. Barriers and bridges to evidence based clinical practice. BMJ 1998;317(7153):273-6.
7. Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. Journal of clinical epidemiology 2006;59(8):770-8.
8. Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. Journal of clinical epidemiology 2010;63(6):620-6.
9. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. The New England journal of medicine 2003;348(26):2635-45.
10. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org.
11. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. Annals of internal medicine 2011;154(11):774-5.
12. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P, et al. Guidelines International Network: toward international standards for clinical practice guidelines. Annals of internal medicine 2012;156(7):525-31.
13. Hoffmann T, Erueti C, Thorning S, Glasziou P. The scatter of research: cross sectional comparison of randomised trials and systematic reviews across specialties. BMJ 2012;344:e3223.
14. World Health Organization: Guidelines for WHO Guidelines. Geneva: World Health Organization; 2003.
15. National Institute for Health and Clinical Excellence: The guidelines manual. London: National Institute for Health and Clinical Excellence; 2017.
16. Baumann MH, Lewis SZ, Gutterman D, American College of Chest P. ACCP evidence-based guideline development: a successful and transparent approach addressing conflict of interest, funding, and patient-centered recommendations. Chest 2007;132(3):1015-24.
17. American Urological Association Education and Research: Overview: Standard Operating Procedures. Linthicum: American Urological Association; 2011.

18. New Zealand Guidelines Group: Handbook for the Preparation of Explicit Evidence-base Clinical Practice Guidelines. Wellington: New Zealand Guidelines Group; 2001.
19. Harrison MB, Legare F, Graham ID, Fervers B. Adapting clinical practice guidelines to local context and assessing barriers to their use. Canadian Medical Association journal 2010;182(2):E78-84.
20. Dahm P, Yeung LL, Gallucci M, et al. How to use a clinical practice guideline. The Journal of urology 2009;181(2):472-9.
21. Browman GP. Improving clinical practice guidelines for the 21st century. Attitudinal barriers and not technology are the main challenges. International journal of technology assessment in health care 2000;16(4):959-68.
22. Browman GP, Makarski J, Robinson P, Brouwers M. Practitioners as experts: the influence of practicing oncologists "in-the-field" on evidence-based guideline development. Journal of clinical oncology 2005;23(1):113-9.
23. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. The British journal of general practice 2007;57(545):971-8.
24. Carlsen B, Norheim OF. "What lies beneath it all?"--an interview study of GPs' attitudes to the use of guidelines. BMC health services research 2008;8:218.
25. Woolf SH. Practice guidelines, a new reality in medicine. II. Methods of developing guidelines. Archives of internal medicine 1992;152(5):946-52.
26. Wollersheim H, Burgers J, Grol R. Clinical guidelines to improve patient care. Netherlands journal of medicine 2005;63(6):188-92.
27. Rosenfeld RM, Shiffman RN, Robertson P; Department of Otolaryngology State University of New York Downstate. Clinical Practice Guideline Development Manual, Third Edition: a quality-driven approach for translating evidence into action. Otolaryngology and head and neck surgery 2013;148(1 Suppl):S1-55.
28. Scottish Intercollegiate Guidelines Network: SIGN 50: A guideline developer's handbook. Edinburgh: Scottish Intercollegiate Guidelines Network; 2015.
29. Grupo de trabajo para la actualización del Manual de Elaboración de GPC. Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud. Actualización del Manual Metodológico. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad - GuíaSalud; 2016.
30. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 2. Priority setting. Health research policy and systems 2006;4:14.
31. Reveiz L, Tellez DR, Castillo JS, Mosquera PA, Torres M, Cuervo LG, et al. Prioritization strategies in clinical practice guidelines development: a pilot study. Health research policy and systems 2010;8:7.
32. Schünemann HJ, Hill SR, Kakad M, Bellamy R, Uyeki TM, Hayden FG, et al. WHO Rapid Advice Guidelines for pharmacological management of sporadic human infection with avian influenza A (H5N1) virus. The Lancet Infectious diseases 2007;7(1):21-31.
33. Schünemann HJ, Hill SR, Kakad M, Vist GE, Bellamy R, Stockman L, et al. Transparent development of the WHO rapid advice guidelines. PLoS medicine 2007;4(5):e119.

34. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. Canadian Medical Association journal 2010;182(18):E839-42.
35. van de Bovenkamp HM, Trappenburg MJ. Reconsidering patient participation in guideline development. Health care analysis 2009;17(3):198-216.
36. van de Bovenkamp HM, Zuiderent-Jerak T. An empirical study of patient participation in guideline development: exploring the potential for articulating patient knowledge in evidence-based epistemic settings. Health expectations 2015;18(5):942-55.
37. Norris SL, Holmer HK, Ogden LA, Burda BU. Conflict of interest in clinical practice guideline development: a systematic review. PloS one 2011;6(10):e25153.
38. Neuman J, Korenstein D, Ross JS, Keyhani S. Prevalence of financial conflicts of interest among panel members producing clinical practice guidelines in Canada and United States: cross sectional study. BMJ 2011;343:d5621.
39. Norris SL, Burda BU, Holmer HK, Ogden LA, Fu R, Bero L, et al. Author's specialty and conflicts of interest contribute to conflicting guidelines for screening mammography. Journal of clinical epidemiology 2012;65(7):725-33.
40. Ioannidis JP. Why most published research findings are false. PLoS medicine 2005;2(8):e124.
41. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294(2):218-28.
42. Schünemann HJ, Al-Ansary LA, Forland F, Kersten S, Komulainen J, Kopp IB, et al. Guidelines International Network: Principles for Disclosure of Interests and Management of Conflicts in Guidelines. Annals of internal medicine 2015;163(7):548-53.
43. Norris SL, Holmer HK, Burda BU, Ogden LA, Fu R. Conflict of interest policies for organizations producing a large number of clinical practice guidelines. PloS One 2012;7(5):e37413.
44. Norris SL, Holmer HK, Ogden LA, Selph SS, Fu R.. Conflict of interest disclosures for clinical practice guidelines in the national guideline clearinghouse. PloS One 2012;7(11):e47343.
45. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. Canadian Medical Association journal 1988;138(8):697-703.
46. Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. International journal of epidemiology 2002;31(1):150-3.
47. Jenuwine ES, Floyd JA. Comparison of Medical Subject Headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. Journal of the Medical Library Association 2004;92(3):349-53.
48. Stevinson C, Lawlor DA. Searching multiple databases for systematic reviews: added value or diminishing returns? Complementary therapies in medicine 2004;12(4):228-32.
49. Savoie I, Helmer D, Green CJ, Kazanjian A. Beyond Medline: reducing bias through extended systematic review search. International journal of technology assessment in health care 2003;19(1):168-78.

50. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924-6.
51. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of clinical epidemiology 2011;64(4):383-94.
52. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ, et al. What is "quality of evidence" and why is it important to clinicians? BMJ 2008;336(7651):995-8.
53. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). Journal of clinical epidemiology 2011;64(4):407-15.
54. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. Journal of clinical epidemiology 2011;64(12):1294-302.
55. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. Journal of clinical epidemiology 2011;64(12):1303-10.
56. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. Journal of clinical epidemiology 2011;64(12):1283-93.
57. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. Journal of clinical epidemiology 2011;64(12):1277-82.
58. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. Journal of clinical epidemiology 2011;64(12):1311-6.
59. Hussain T, Michel G, Shiffman RN. The Yale Guideline Recommendation Corpus: a representative sample of the knowledge content of guidelines. International journal of medical informatics 2009;78(5):354-63.
60. Schünemann HJ, Oxman AD, Akl EA, Brozek JL, Montori VM, Heffner J, et al. Moving from evidence to developing recommendations in guidelines: article 11 in Integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. Proceedings of the American Thoracic Society 2012;9(5):282-92.
61. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 2016;353:i2016.
62. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ 2016;353:i2089.
63. Martínez García L, Arévalo-Rodríguez I, Solà I, Haynes RB, Vandvik PO, Alonso-Coello P, et al. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implementation science 2012;7:109.

64. Alonso-Coello P, Martínez García L, Carrasco JM, Solà I, Qureshi S, Burgers JS, et al. The updating of clinical practice guidelines: insights from an international survey. Implementation science 2011;6:107.
65. Martínez García L, Pardo-Hernández H, Sanabria AJ, Alonso-Coello P, Penman K, McFarlane E, et al. Guideline on terminology and definitions of updating clinical guidelines: the Updating Glossary. Journal of clinical epidemiology 2017 pii: S0895-4356(17):31021-1
66. Shekelle P, Woolf S, Grimshaw JM, Schünemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implementation science 2012;7:62.
67. Martínez García L, Sanabria AJ, García Alvarez E, Trujillo-Martín MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. Canadian Medical Association journal 2014;186(16):1211-9.
68. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. Journal of clinical epidemiology 2014;67(1):52-5.
69. Neuman MD, Goldstein JN, Cirullo MA, Schwartz JS. Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations. JAMA 2014;311(20):2092-100.
70. Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B, et al. Updating clinical practice recommendations: is it worthwhile and when? International journal of technology assessment in health care 2012;28(1):29-35.
71. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001;286(12):1461-7.
72. Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? Lancet 2006;367(9514):881-3.
73. Agbassi C, Messersmith H, McNair S, Brouwers M. Priority-based initiative for updating existing evidence-based clinical practice guidelines: the results of two iterations. Journal of clinical epidemiology 2014;67(12):1335-42.
74. Martínez García L, Pardo-Hernandez H, Superchi C, Niño de Guzman E, Ballesteros M, Ibargoyen Roteta N, et al. Methodological systematic review identifies major limitations in prioritization processes for updating. Journal of clinical epidemiology 2017;86:11-24.
75. Alonso-Coello P, Martinez Garcia L. [Clinical guidelines: old and new challenges]. Medicina clinica 2014;143(7):306-8.
76. Amos Q, Chan W, Tom G. Maximising efficiency in updating guidelines through prioritisation of clinical questions. 10th G-I-N Conference. San Francisco, USA 2013.
77. Grupo de actualización de la Guía Metodológica: [Guía Metodológica para la elaboración de Guías de Práctica Clínica con Evaluación Económica en el Sistema General de Seguridad Social en Salud Colombiano]. In: Ministerio Salud Colombia; 2014.

78. Martínez García L, Sanabria AJ, Araya I, Lawson J, Solà I, Vernooij RW, et al. Efficiency of pragmatic search strategies to update clinical guidelines recommendations. BMC medical research methodology 2015;15:57.
79. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Assessing the need to update prevention guidelines: a comparison of two methods. International journal for quality in health care 2004;16(5):399-406.
80. Eccles M, Rousseau N, Freemantle N. Updating evidence-based clinical guidelines. Journal of health services research & policy 2002;7(2):98-103.
81. Parmelli E, Papini D, Moja L, Bandieri E, Belfiglio M, Ciccone G, et al. Updating clinical recommendations for breast, colorectal and lung cancer treatments: an opportunity to improve methodology and clinical relevance. Annals of oncology 2011;22(1):188-94.
82. Kaiser K, Miksch S. Versioning computer-interpretable guidelines: semi-automatic modeling of 'Living Guidelines' using an information extraction method. Artificial intelligence in medicine 2009;46(1):55-66.
83. Akl EA, Meerpohl JJ, Elliott J, Kahale LA, Schünemann HJ, Living Systematic Review Network. Living systematic reviews: 4. Living guideline recommendations. Journal of clinical epidemiology 2017;91:47-53.
84. Johnston ME, Brouwers MC, Browman GP. Keeping cancer guidelines current: results of a comprehensive prospective literature monitoring strategy for twenty clinical practice guidelines. International journal of technology assessment in health care 2003;19(4):646-55.
85. Uhlig K, Berns JS, Carville S, Chan W, Cheung M, Guyatt GH, et al. Recommendations for kidney disease guideline updating: a report by the KDIGO Methods Committee. Kidney International 2016;89(4):753-60.
86. Martínez García L, Pardo-Hernández H, Sanabria AJ, Alonso-Coello P, Pregnancy Clinical Guideline Updating Working Group. Continuous surveillance of a pregnancy clinical guideline: an early experience. Systematic reviews 2017;6(1):143.
87. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. The Medical journal of Australia 2006;185(5):263-7.
88. Smidt N, Rutjes AW, van der Windt DA, Ostelo RW, Bossuyt PM, Reitsma JB, et al. The quality of diagnostic accuracy studies since the STARD statement: has it improved? Neurology 2006;67(5):792-7.
89. Prady SL, Richmond SJ, Morton VM, Macpherson H. A systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. PloS One 2008;3(2):e1577.
90. Wang X , Chen Y, Yang N, Deng W, Wang Q, Li N, et al. Methodology and reporting quality of reporting guidelines: systematic review. BMC medical research methodology 2015;15:74.
91. Simera I, Altman DG, Moher D, Schulz KF, Hoey J. Guidelines for reporting health research: the EQUATOR network's survey of guideline authors. PLoS medicine. 2008;5(6):e139.
92. Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. Lancet 2005;365(9465):1159-62.
93. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Annals of internal medicine 2006;144(6):427-37.

94. Lee CW, Chi KN. The standard of reporting of health-related quality of life in clinical cancer trials. Journal of clinical epidemiology 2000;53(5):451-8.
95. Mallett S, Deeks JJ, Halligan S, Hopewell S, Cornelius V, Altman DG. Systematic reviews of diagnostic tests in cancer: review of methods and reporting. BMJ 2006;333(7565):413.
96. Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. BMJ 2004;329(7471):883.
97. Equator network – Enhancing the QUAlity and Transparency Of health Research [Internet]. [Consulted on 1 December 2017]. Available on: http://www.equator-network.org.
98. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS medicine 2010;7(3):e1000251.
99. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596):1453-7.
100. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
101. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.
102. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. Systematic reviews 2012;1:60.
103. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. The Cochrane database of systematic reviews 2012;11:MR000030.
104. Brouwers MC, Kerkvliet K, Spithoff K; AGREE Next Steps Consortium. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. BMJ 2016;352:i1152.
105. Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. Canadian Medical Association journal 2014;186(3):E123-42.
106. Chen Y, Yang K, Marušic A, Qaseem A, Meerpohl JJ, Flottorp S, et al. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. Annals of internal medicine 2017;166(2):128-32.
107. Tsertsvadze A, Maglione M, Chou R, Garritty C, Coleman C, Lux L, et al. Updating comparative effectiveness reviews: current efforts in AHRQ's Effective Health Care Program. Journal of clinical epidemiology 2011;64(11):1208-15.
108. Clark E, Donovan EF, Schoettker P. From outdated to updated, keeping clinical guidelines valid. International journal for quality in health care 2006;18(3):165-6.

109. Shekelle PG. Updating practice guidelines. JAMA 2014;311(20):2072-3.
110. Burgers JS, Cluzeau FA, Hanna SE, Hunt C, Grol R. Characteristics of high-quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. International journal of technology assessment in health care 2003;19(1):148-57.
111. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet 2017;390(10092):415-23.
112. Pieper D, Mathes T. Survey of instructions for authors on how to report an update of a systematic review: guidance is needed. Evidence-based medicine 2017;22(2):45-48.
113. Vernooij RW, Martínez García L, Brouwers M, Alonso-Coello P; CheckUp Panel. CheckUp: Explanation and elaboration of a checklist for the reporting of updating clinical guidelines. PLoS Medicine 2017;14(1):e1002207.
114. Martínez García L, Pardo-Hernandez H, Niño de Guzman E, Superchi C, Ballesteros M, McFarlane E, et al. Development of a prioritisation tool for the updating of clinical guideline questions: the UpPriority Tool protocol. BMJ open. 2017;7(8):e017226.
115. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
116. Vernooij RW, Sanabria AJ, Solà I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implementation science 2014;9:3.
117. Vernooij RW, Alonso-Coello P, Brouwers M, Martínez García L, CheckUp Panel. Reporting Items for Updated Clinical Guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp). PLoS Medicine 2017;14:e1002207.
118. Vernooij RWM, Martínez García L, Florez ID, Hildago Armas L, Poorthuis MHF, Brouwers M, et al. Updated clinical guidelines experience major reporting limitations. Implementation science 2017;12(1):120.
119. Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. Journal of clinical epidemiology 2014;67(1):33-45.
120. Schünemann HJ1, Fretheim A, Oxman AD; WHO Advisory Committee on Health Research. Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. Health research policy and systems 2006;4:13.
121. Ansari S, Rashidian A. Guidelines for guidelines: are they up to the task? A comparative assessment of clinical practice guideline development handbooks. PloS One. 2012;7(11):e49864.
122. Kredo T, Bernhardsson S, Machingaidze S, Young T, Louw Q, Ochodo E, et al. Guide to clinical practice guidelines: the current state of play. International journal for quality in health care 2016;28(1):122-8.
123. Brandt L, Vandvik PO, Alonso-Coello P, Akl EA, Thornton J, Rigau D, et al. Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial. BMJ Open 2017;7(2):e011569.
124. Kastner M, Bhattacharyya O, Hayden L, Makarski J, Estey E, Durocher L, et al. Guideline uptake is influenced by six implementability domains for creating and

communicating guidelines: a realist review. Journal of clinical epidemiology 2015;68(5):498-509.
125. Martínez García L, McFarlane E, Barnes S, Sanabria AJ, Alonso-Coello P, Alderson P. Updated recommendations: an assessment of NICE clinical guidelines. Implementation science 2014;9:72.
126. Newberry SJ, Shekelle PG, Vaiana M, Motala A. Reporting the Findings of Updated Systematic Reviews of Comparative Effectiveness: How Do Users Want To View New Information? Rockville (MD): 2013.
127. Alonso-Coello P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. Quality & safety in health care 2010;19(6):e58.
128. Armstrong JJ, Goldfarb AM, Instrum RS, MacDermid JC. Improvement evident but still necessary in clinical practice guideline quality: a systematic review. Journal of clinical epidemiology 2017;81:13-21.
129. Hasenfeld R, Shekelle PG. Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR) guidelines with those published subsequently. Quality & safety in health care 2003;12(6):428-34.
130. Shea B, Boers M, Grimshaw JM, Hamel C, Bouter LM. Does updating improve the methodological and reporting quality of systematic reviews? BMC medical research methodology 2006;6:27.
131. Eady EA, Layton AM, Sprakel J, Arents BW, Fedorowicz Z, van Zuuren EJ. AGREE II assessments of recent acne treatment guidelines: how well do they reveal trustworthiness as defined by the U.S. Institute of Medicine criteria? The British journal of dermatology 2017,10.1111/bjd.15777.
132. Vasconcelos L, Kalb Wainberg S, Molino CDGRC, Melo DOD. Quality of the reporting of Brazilian Ministry of Health updated clinical practice guidelines (CPGU) versions: Applying checkUP. Abstracts of the Global Evidence Summit, Cape Town, South Africa. Cochrane Database of Systematic Reviews 2017, Issue 9 (Suppl 1).
133. Silva RAM, Molino CDGRC, Santos NCL, Melo DOD. Assessment of the reporting quality of clinical practice guidelines using CheckUP list/tool. Abstracts of the Global Evidence Summit, Cape Town, South Africa. Cochrane Database of Systematic Reviews 2017, Issue 9 (Suppl 1).
134. Moher D1, Jones A, Lepage L; CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA 2001;285(15):1992-5.
135. Areia M, Soares M, Dinis-Ribeiro M. Quality reporting of endoscopic diagnostic studies in gastrointestinal journals: where do we stand on the use of the STARD and CONSORT statements? Endoscopy 2010;42(2):138-47.

APPENDIX I

Article: "CheckUp: Explanation and elaboration of a checklist for the reporting of updating clinical guidelines."

CheckUp: Explanation and elaboration of a checklist for the reporting of updating clinical guidelines

Vernooij RWM[1], Martínez García L[1]*, Brouwers M[2], Alonso-Coello P[1,2]; CheckUp Panel.

1. Iberoamerican Cochrane Centre, and Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain.

2. CIBER of Epidemiology and Public Health (CIBERESP), Spain.

3. Program in Evidence-based Care, Cancer Care Ontario, Canada.

Membership of the UpCheck Panel is provided in the Acknowledgements.

* laura.martinez.garcia@cochrane.es

Abstract

Background

Due to the continuous emergence of new evidence, clinical guidelines have a limited life-time. However, evidence based reporting standards for the updating of clinical guidelines are still not available. In this article we present CheckUp (CHECKlist for the reporting of UPdating clinical guidelines) and provide explanations, and examples for each item. In a previous article we present the methodology to develop CheckUp.

Methods and findings

We developed an initial list of items – including explanation and examples – based on an overview of research on clinical guideline updating, the AGREE (Appraisal of Guidelines for Research and Evaluation) II Instrument, and the advice of the CheckUp Panel. A multi-step process was used to refine this list including an assessment of existing guidelines, a Delphi consensus survey, key informant interviews, and a formal external review by both clinical guideline methodologists and users. CheckUp includes sixteen items that address the: 1) presentation (version, rationale for updating, changes in scope and purpose, reviewed sections, recommendation labels, changes in the recommendations); 2) editorial independence (panel participants and disclosures of interest); and 3) methodology of the updating process (methods for search strategy, evidence selection, quality assessment, evidence synthesis, external review, implementation, and future updates).

Conclusions

CheckUp can be used to evaluate the completeness of reporting in updated guidelines, and as a tool to inform guideline developers about reporting requirements. This explanation and elaboration article aims to facilitate the use of CheckUp both, for clinical guidelines users and developers, by explaining the rationale of the items, and providing relevant examples of adequate reporting.

Keywords

Evidence-Based Medicine/Standards, Practice Guidelines as Topic, Information Dissemination, Publishing/Standards, Quality Control, Terminology as topic, Program Development, Program Evaluation

Word count

Abstract: 255 words

Manuscript: 5186 words.

Background

Trustworthy clinical guidelines aim to assist decision making by providing recommendations that are informed by the best available evidence, and include an assessment of the benefits and harms of alternative care options [1,2]. Due to the continuous emergence of new research evidence (i.e. changes in available interventions, effects, or costs) [3], appropriate updating to maintain the trustworthiness of clinical guidelines is challenging since it requires regular reviewing [4,5]. To date, little attention has been paid to the methodology for updating clinical guidelines [5-8].

Very little guidance in reporting standards regarding the updating process of clinical guidelines has been identified [9,10]. To address this need, in a partnership of the Iberoamerican Cochrane Center [www.cochrane.org], the AGREE Collaboration [www.agreetrust.org] and the Guidelines International Network (G-I- N) Updating Working Group [www.g-i-n.net/working-groups/updating-guidelines], we have developed CheckUp (CHECKlist for the reporting of UPdated guidelines). The aim of CheckUp is to evaluate the completeness of reporting the updating process in clinical guidelines, and to inform clinical guideline developers about reporting requirements.

This series of two articles about CheckUp is targeted at guideline developers and users of guidelines. In the first article we presented the methodology of the development process [11]. In this second article we explain and elaborate on all checklist items including the explanation and examples with the goal to facilitate the use of CheckUp.

Methods

The development of CheckUp consisted of four phases: 1) panel selection; 2) generation of the initial checklist; 3) optimisation of the checklist; and 4) approval of the final checklist. We reported the detailed methodology of the development process of CheckUp in a previous article [11].

To advise on the development of CheckUp, a panel was convened comprised of individuals with expertise in clinical guideline development, updating and/or research methodology. The purpose of the panel was to provide expert advice on the checklist and to participate in a Delphi survey. A core group of authors (RWMV, LMG, MB, and PAC) was established to provide time-sensitive and operational advice.

We developed an initial list through discussion and brainstorming, taking into account the following aspects: 1) key research literature, including two systematic reviews about updating [7,8,10]; 2) the AGREE II instrument [12]; and 3) the CheckUp Panel experience. In this process the initial list of items, examples, and explanations were generated. We refined the checklist subsequently by a multi-step process that included: 1) an assessment of a sample of updated clinical guidelines; 2) semi-structured interviews; 3) a Delphi consensus survey; and 4) an external review with clinical guideline methodologists and users. Finally, all CheckUp Panel members reviewed and approved the final version of the checklist (Table 1).

85 ***Table 1. CheckUp (Checklist for reporting the Updating process) items and assessment***

Item	Assessment	Reported on page number.	Notes
1. The updated version can be distinguished from the previous version of the clinical guideline	Yes No Unclear Not applicable		
2. The rationale for updating the clinical guideline is reported	Yes No Unclear Not applicable		
3. Changes in the scope and purpose between the updated and previous version are described and justified	Yes No Unclear Not applicable		
4. The sections reviewed in the updating process are described	Yes No Unclear Not applicable		
5. Recommendations are clearly presented and labelled as new, modified, or not changed. Deleted recommendations are clearly noted	Yes No Unclear Not applicable		
6. Changes in recommendations are reported and justified	Yes No Unclear Not applicable		
7. The panel participants in the updated version are described	Yes No Unclear Not applicable		
8. Disclosures of interests of the group responsible for the updated version are recorded	Yes No Unclear Not applicable		

9.	The role of the funding body for the updated version is identified and described	Yes No Unclear Not applicable		
10.	The methods used for searching and identifying new evidence in the updating process are described	Yes No Unclear Not applicable		
11.	The methods used for evidence selection in the updating process are described	Yes No Unclear Not applicable		
12.	The methods used to assess the quality of the included evidence in the updating process are described	Yes No Unclear Not applicable		
13.	The methods used for the evidence synthesis in the updating process are described	Yes No Unclear Not applicable		
14.	The methods used for externally reviewing the updated version are described	Yes No Unclear Not applicable		
15.	The methods and plan for implementing the changes of the updated version in practice are described	Yes No Unclear Not applicable		
16.	The plan and methods for updating the new version in the future are reported	Yes No Unclear Not applicable		

86

Explanation and Elaboration of the items and examples of CheckUp

The answer possibilities for all items are: 1) Yes, 2) No, 3) Unclear, and 4) Not applicable. Whenever the users of CheckUp believe that certain concepts are not applicable in the updated guideline subject to assessment, the answer "Not applicable" can be used. Consequently, the denominator of total items should be adjusted. Similarly, the answer possibility "Unclear" can be used whenever the user is doubting whether the items is adequately described in the updated clinical guideline.

Item 1. The updated version can be distinguished from the previous version of the clinical guideline

Explanation

The clinical guideline users must be able to easily identify that they are consulting an updated version of the clinical guideline. If possible, a reference to the previous version should be provided and the clinical guideline has to be identified as an update.

Examples

- Infection prevention and control guideline for cystic fibrosis: 2013 update [13].
- This is an updated version of the clinical guideline previously published in 2014 (illustrative example).

In these examples, branding the updated version has been achieved by using the term "update" and the publication date (year) in the title. Other options include new version number/code in the title.

Item 2. The rationale for updating the clinical guideline is reported

Explanation

There are different situations that might trigger the updating process of a clinical guideline (e.g. elapsed time frame or identification of new evidence). Clinical guideline users should be able to identify these reasons. Whenever the rationale for updating differs among various sections in the clinical guideline, these differences should be reported.

Examples

- After publishing the clinical guideline in 2006, a time period of five years for starting the updating process was considered adequate. The updating process was therefore started in 2011. After review and analysis of the new evidence since the previous version of the clinical guideline, the Update Panel concluded that the new evidence was compelling enough to warrant substantive changes in various recommendations (illustrative example).
- Recently, three new randomised controlled trials have been published.Given the importance of these studies and their relevance to this clinical guideline, an update of this clinical guideline was undertaken by the Update Panel. This new evidence prompted a new systematic review of the literature (illustrative example).
- In the yearly revision of all the clinical guidelines with the clinical experts included in our working group, two members indicated the existence of two new randomised controlled trials that are not included in the previous version of this clinical guideline. Therefore, we decided to update this clinical guideline and conduct a new systematic review of the literature (illustrative example).

In these examples different rationales for updating a clinical guideline have been stated and other reasons are imaginable: 1) a certain pre-specified time frame elapsed; 2) new evidence is expected or known to exist; or, 3) alerts / clinical experts indicate a possible obsolescence of the recommendations in the current clinical guideline.

Item 3. Changes in the scope and purpose between the updated and previous version are described and justified

Explanation

If in the updating process the scope and purpose of the clinical guideline change, the clinical guideline users have to be able to identify the differences in the scope and purpose of the updated version in contrast with the previous version, including the justification for these changes. A statement should be included if there are no differences in the scope and purpose between the updated and the previous version of the clinical guideline.

Examples

- The scope of the clinical guideline changed since the working group decided to expand due to new identified evidence on the prevention of Urinary Tract Infections (UTIs) in patients treated with catheters, which was not included in the previous version of the clinical guideline. Consequently, the scope of the clinical guideline is extended to the prevention, diagnosis, and treatment of UTIs in patients treated with catheters (illustrative example).
- Given the availability of randomised controlled trials included in the updated version of this clinical guideline, the scope of the inclusion criteria was modified to only include higher level evidence. Consequently, the scope of the clinical guideline changed to give insights in the prevention, diagnosis, and treatment of schizophrenia in primary care setting, based only on the highest level evidence available. Case controlled studies included in the previous version were moved to appendix X and serve now as historical evidence (illustrative example).

In these examples, reasons for changes in the scope and purpose of the clinical guideline have been included, by providing modified or new clinical questions. Other plausible reasons for

changes in the scope and purpose are modifications to the target population, interventions, comparisons or outcomes, or a new clinical setting.

Item 4. The sections reviewed in the updating process are described

Explanation

An updated clinical guideline may be modified in all, some, or no sections (chapters, questions, etc.) compared to the previous version of the clinical guideline. Which sections have changed, which sections were reviewed but remained unchanged, and, if applicable, which sections have not changed should be transparently reported.

Examples

- Summary of updates, by section of the clinical guideline (Table 2) (illustrative example).

In this example, an overview has been provided with information about which sections have been, for example, reviewed and modified, reviewed but not modified, or not been reviewed. The terminology is not mandatory, it might be adapted to the needs of each organisation.

Table 2. Example for Item 4

Clinical Guideline Sections	**Actions**
1. Key recommendations	New section
2. Diagnosis of bacterial UTI in patients with catheters	Not reviewed
3. Management of bacterial UTI in adult women	Reviewed and modified
4. Management of bacterial UTI in pregnant women	Reviewed and modified
5. Management of bacterial UTI in adult men	Reviewed but not modified
6. Management of bacterial UTI in patients with catheters	Reviewed and modified

7. Management of bacterial UTI in new-borns	Deleted
8. Prevention of bacterial UTI in patients with catheters	New section
9. Provision of information	Reviewed and modified
10. Implementing the clinical guideline	Reviewed but not modified.
11. External review	Not repeated

Abbreviation: UTI: Urinary Tract Infection.

Item 5. Recommendations are clearly presented and labelled as new, modified, or not changed. Deleted recommendations are clearly noted

Explanation

If some recommendations change due to newly identified evidence, the clinical guideline users should be able to identify these changes at the recommendation level. Recommendations can be labelled as new, modified, or not changed. In case recommendations are deleted, clinical guideline users should be alerted.

Examples

- Discuss with people who have or who are at risk of breast-cancer related lymphoedema that there is no indication that exercise prevents, causes or worsens lymphoedema [new 2014] [14].
- Explore whether the patient experienced in the past allergies to any antibiotics that can be used in the treatment of UTIs [reviewed and not changed 2014] (illustrative example).
- Due to the fact that a new and more effective antibiotic Y for the treatment of UTIs is available since 2013, we decided to delete three recommendations that previously advocated the use of antibiotic X. These deleted recommendations have been included in appendix X (illustrative example).
- Overview of recommendations included in section X (Table 3) (illustrative example).

197 ***Table 3. Example for Item 5***

Recommendations	Status
1. Explore whether the patient experienced in the past allergies to any antibiotics that can be used in the treatment of UTIs.	Reviewed and not changed 2014
2. Explore whether the patients is using at this current moment any other antibiotics for the treatment of any other infection.	Not reviewed

198

199 In these examples, labels for the status of recommendations have been used, including [new
200 2014], whenever a recommendation is newly implemented in the updated version and was not
201 included in the previous version. Other possible terminology for labelling recommendations
202 could be: "[modified 2014]" if the recommendation has been reviewed and modified in 2014;
203 "[2009]" if the recommendation has not been reviewed since the previous version of the clinical
204 guideline; or in the case of reviewed but unchanged recommendations, a label like "[reviewed
205 and not changed 2014]" might be used. An overview with all the recommendations and their
206 corresponding labels can also be provided. If applicable, an alert about the deleted
207 recommendations should be included.

208

209 ***Item 6. Changes in recommendations are reported and justified***

210 **Explanation**

211 In the updated version, clinical guideline users should be able to identify the changes made in
212 the recommendations. Information regarding which specific parts of the recommendations have
213 been changed, including a justification of these changes, should be provided. To avoid possible
214 confusion between updated and previous recommendations this information should be included
215 in an appendix. A statement should be included if there are no differences between the
216 recommendations of the updated and previous version of the clinical guideline.

Examples

- Appendix: Overview of the new recommendation, previous recommendations, and the rationale for changes (Table 4) [15].

In this example an examination has been given to compare the previous recommendation and the recommendation after the update. To illustrate the rationale for updating the recommendations, a reflection about the evidence inducing the changes has been given.

Table 4. Example for Item 6

New recommendation	Previous recommendation	Reason for change/deletion
For people with unexplained infertility, mild endometriosis or 'mild male factor infertility', who are having regular unprotected sexual intercourse: • do not routinely offer intrauterine insemination, either with or without ovarian stimulation (exceptional circumstances include, for example, when people have social, cultural or religious objections to IVF) • advise them to try to conceive for a total of 2 years (this can include up to 1 year before their fertility investigations) before IVF will be considered. [new 2013]	Where intrauterine insemination is used to manage male factor fertility problems, ovarian stimulation should not be offered because it is no more clinically effective than unstimulated intrauterine insemination and it carries a risk of multiple pregnancy. [A]	New evidence has shown that IUI with or without stimulation is no more effective than no treatment. Therefore, it is no longer recommended.

Item 7. The panel participants in the updated version are described

Explanation

This item refers to the panellists who were involved in the updating process. It is plausible that in the time period between the previous and the updated version the composition of the clinical guideline panel changes, especially when the scope and purpose of the clinical guidelines vary. Additionally, the panel participants of the previous version of the clinical guideline should be

reported, especially in the case of recommendations that were not changed in the updating process. Patients, public, and other stakeholders' involvement in the panel responsible for updating the clinical guideline should be clearly reported.

Examples

- The panel members involved in *the 2014* update (including their names, disciplines, organisation, and the role of each panellist) are listed in appendix 1. The composition of the group responsible for the previous version of the clinical guideline is included in this appendix (illustrative example).

In this example, the name, expertise, affiliated organisations, geographical location, the role of each member of the panel responsible for updating the clinical guideline, and of those responsible for the previous version of clinical guideline are clearly described.

Item 8. Disclosures of interests of the group responsible for the updated version are recorded

Explanation

All group members who participated in the updating process should disclosure their conflicts of interest (CoI) before starting the updating process. The CoI of those involved in the updating process are important for the recommendations that have been updated or newly implemented. Additionally, the CoI of the participants who were involved in the previous version and not in the updating process should still be reported, since they remain relevant for the recommendations that are not reviewed or those that have not been changed.

Examples

- Members of the Update Panel completed a CoI form, which requires disclosure of financial and other interests that are relevant to the subject matter of the clinical

guideline, including relationships with commercial entities that are reasonably likely to experience direct regulatory or commercial impact with the implementation of the clinical guideline. These disclosure forms are included on the website [link] (illustrative example).

- In accordance with the procedures, the majority of members of the Update Panel did not disclose any such relationship. For the recommendations regarding medicine X, one participant reported substantial CoI, due to previous research funding and the involvement in publications regarding this medicine. Consequently, this participant has been excluded from the discussions about medicine X (illustrative example).
- In addition, the CoI of the panel responsible for the development process of the previous version of the clinical guideline are included in appendix VI. These CoI are still valid for the recommendations that are not reviewed (illustrative example).
- Measures for handling intellectual CoI, such as including published work of those responsible for constructing the recommendation, have been reported in appendix VI (illustrative example).

In these examples, a description of the CoI, type and in what form, how these were sought, and how these were addressed, has been included. The CoI of the people involved solely in the updating process, or in both the updating process and the development process of the previous version, or only in the development process of the previous version, have been described.

Item 9. The role of the funding body for the updated version is identified and described

Explanation

Clinical guidelines are frequently developed with the help of an external funding body, in the form of financial contribution used for realising the updating process. A description of the role of the funding body should be reported in the updated version, including an indication of the

amount of funds provided. If a clinical guideline is self-funded (with no external support) this should be explicitly reported. If the funding for the updating process differs from the funding of the previous version, a description of both funding bodies should be given.

Examples

- Support for this clinical guideline was provided by the Society for Healthcare Epidemiology of America (SHEA) and the Infectious Diseases Society of America (IDSA). The financial support of the previous version of this clinical guideline solely consists of the Society for Healthcare Epidemiology of America. Both organisations responsible for the external funding did not have any influence in the content of the clinical guideline [16].

In this example, an explicit statement is included in the updated version, providing details of the funding body. Further details on the amount of funds provided, the role of the funding body in the updating process, and its influence on the content of the clinical guideline are reported.

Item 10. The methods used for searching and identifying new evidence in the updating process are described

Explanation

A complete documentation of the search strategy should be included in the updated version to allow clinical guideline users to reiterate the search strategy, to be informed about the searched sources, and to evaluate the quality of the search strategy. Differences in the search strategy between the updated and the previous version should be justified. In addition, whenever different search strategies have been used for different recommendations or clinical questions, this should be stated.

Examples

- Electronic databases (MEDLINE and EMBASE) were searched from 2000 to 2008. The single exception was for question 7, where it was stipulated in the research question to exclude studies published before 2003. The full search strategy is described in appendix II [17].
- With the exception of the searched interval dates, the search strategy followed the previous strategy used in the preceding clinical guideline (appendix II): MEDLINE, EMBASE, and CINAHL: September 2010 to November 2014; Cochrane library: 2014 (illustrative example).

In these examples, details of the used search strategy for identifying the new evidence in the updated version are reported, including the search terms used for all recommendations and clinical questions (appendix), consulted databases (e.g. MEDLINE, EMBASE), and the time period considered. The complete search strategy is included in an appendix. Finally, the differences in methodology of the search strategy between the updated and the previous version are reported.

Item 11. The methods used for evidence selection in the updating process are described

Explanation

Criteria for including or excluding the new identified evidence in the updating process should be explicitly described. If the inclusion and exclusion criteria in the updated version differ from those used in the previous version, these changes should be reported and justified. A statement should be included if there are no differences in methodology for evidence selection between the updated and the previous version.

Examples

- The inclusion and exclusion criteria differ from those used in the 2009 clinical guideline. For the updated version, we included prevention of UTI for patients treated with catheters. Therefore, the inclusion criteria are now: systematic reviews or randomised controlled trials published in English; addressing the prevention, diagnosis, or treatment of UTIs for patients treated with catheters (illustrative example).
- We applied the same eligibility criteria (inclusion and exclusion criteria) as in the previous clinical guideline. See appendix III (illustrative example).

In these examples, the inclusion and exclusion criteria used are clearly reported. These criteria refer to study design, languages, and study objectives. In addition, in the inclusion criteria reference has been made to the clinical question of the clinical guideline.

Item 12. The methods used to assess the quality of the included evidence in the updating process are described

Explanation

A complete documentation of the evidence assessment allows clinical guideline users to examine the rating of the quality of the evidence. The differences in methodology used to assess the evidence between the updated and the previous version should be reported, including the rationale behind those differences. Additionally, whenever the methodology (e.g. instrument or tool) is used in the updated version differs from that used in the previous version it should be described whether the evidence included in the previous version was also assessed with the new methodology. A statement should be included if there are no differences in the methodology for evidence assessment between the updated and previous version.

Examples

- In the updating process of the clinical guideline, the quality assessment was conducted with the GRADE (Grading of Recommendations Assessment, Development and Evaluation) system. In the GRADE approach, the quality of the evidence is obtained by taking into consideration several factors. An assessment of the risk of bias, inconsistency, indirectness, imprecision, and other considerations (including publication bias) has been conducted. We have used GRADE to evaluate the whole evidence base, not just the new studies identified (illustrative example).
- The methods used to assess the quality of the studies found during the updating were similar as those used in the previous version of the clinical guideline (illustrative example).

In these examples, the methodology used to assess the risk of bias and/or the quality of the evidence is described (e.g. GRADE methodology or other instruments), as well as the differences with the previous versions.

Item 13. The methods used for the evidence synthesis in the updating process are described

Explanation

The methodology used to combine multiple sources of quantitative and/or qualitative evidence should be reported in the updated version. Furthermore, the methodology used to translate the evidence into recommendations should be stated in the updated version. A statement should be included if there are no differences in methodology for evidence synthesis between the updated and the previous version.

Examples

- Generated summaries of the evidence by outcome: 1) randomised studies: meta-analysed, where appropriate and reported in GRADE profiles for clinical studies. 2) Observational studies: data presented as a range of values in GRADE profiles. 3) Qualitative studies: each study summarised in a table (available in an appendix) where possible, and the quality of included studies assessed against the NICE quality checklists for qualitative studies. Key common themes between studies which were relevant to the review question were summarised and presented with a comment of the quality of studies contributing to the themes in the main clinical guideline document. GRADE does not have a system for rating the quality of evidence for qualitative studies or surveys, and therefore there are no GRADE quality ratings for the themes identified [18].
- The methods for evidence synthesis during the updating process followed those used in the development of the previous version of the clinical guideline (illustrative example).
- Studies in the updated and previous version were not meta-analysed due to the paucity of evidence reporting individual outcomes. GRADE evidence profiles have, therefore, been presented in the appendix XI, with the new evidence clearly identified (illustrative example).

In these examples, the methodology used for synthesising the newly identified evidence (quantitative and/or qualitative) with the evidence included in the previous version is reported.

Item 14. The methods used for externally reviewing the updated version are described

Explanation

An updated version should be externally reviewed by experts in the clinical area, methodological experts, and/or public consultation. The reviewers should not have been

involved in the updating process of the clinical guideline. On the other hand, if no changes have been made, an external review of the updated version might seem less appropriate. Where external review was deemed appropriate, the methods, results and impact should be described, including information of the reviewers.

Examples

- The updated version was reviewed in draft form by the following expert referees, who were members of the previous clinical guideline development group. All expert referees made declarations of interest and further details of these are available on request from the SIGN Executive [18].
- Given that the recommendations were not changed, an external review of the updated version was not undertaken (illustrative example).
- Given that solely minor changes were made to the recommendations, a modified strategy of the external review process was undertaken. Specifically, we only asked the group participating in the external review to provide comments on the recommendations that have been changed (illustrative example).

In these examples, documentation of the methodology used to conduct the external review has been presented. If an external review of the updated version has been conducted, it is stated whether specific instructions to appraise the updated version were provided to the reviewers.

Item 15. The methods and plan for implementing the changes of the updated version in practice are described

Explanation

If applicable, specific accompanying materials produced to support the implementation of the updated version should be provided. The implementation plan should lay emphasis on the new recommendations of the updated version, or on the recommendations with significant changes.

In addition, the implementation plan from the previous version might have been evaluated, and consequently improved for the updated version. A statement should be included if there are no differences in methodology for the implementation plans or strategies between the updated and the previous version.

Examples

- The implementation plan of the updated version includes the following aspects: 1) Regular audit, with feedback of non-adherence to local clinical guidelines (including specific clinician feedback). This should be actively discussed and acted upon on a regular basis. 2) Active involvement and support from local senior staff or respected opinion leaders for the implementation strategy programme. 3) Indicate the changes introduced in recommendations in the updated version with respect to the previous version, by directing clinical guideline users to appendix X (illustrative example).
- The methods for the implementation plan during the updating procedure followed those used in the development of the previous version (illustrative example).

In these examples the methodology and the specific aspects of the implementation plan of the updated clinical guideline have been given. If this consists of different aspects, they have all been illustrated.

Item 16. The plan and methods for updating the new version in the future are reported

Explanation

Updating is a crucial process for maintaining the validity of recommendations. A clear statement about the methodology and the plan of the forthcoming updating procedures should be provided. It should be clearly noted whenever the updating process will differ among clinical questions or recommendations. Additionally, if specific cases exist that trigger an update before

the time frame, these should be reported. If deemed appropriate, a specific time period for the next update, including a rationale for this period, must be provided.

Examples

- This clinical guideline was issued in 2014 and will be considered for review within two years to assure the validity of this clinical guideline because of the relatively quick advances in this clinical area. The panel responsible for updating the clinical guideline in 2014 will be contacted again for participation in the updating process in 2016. Any updates to the clinical guideline in the interim period will be noted on the website. In principle, the same methodology for this current update will be used for the update in 2016 (illustrative example).

In this example, the procedure, including the time interval, for updating the clinical guideline in the future is provided in the updated version, including an established panel that will be responsible for updating the clinical guideline.

Discussion

We developed CheckUp using a rigorous development process, including the use of systematic reviews, assessment of updated clinical guidelines, and engagement of a large international clinical guideline community through semi-structured interviews, a Delphi consensus survey, and a wide external review process. CheckUp includes 16 items regarding the presentation, the editorial independence, and the methodology used in the clinical guideline updating process.

Like similar explanation and elaboration documents from other instruments [19-25] the primary aim of this manuscript is to outline a framework for adequate use of the CheckUp checklist. With this article we aim to improve its usability by explaining the rationale of the items, and by providing relevant examples of optimal reporting.

Our results in the context of previous research

The methodology for updating clinical guidelines is not standardised and should be more rigorous [7]. This lack of rigour, compared to de novo development of clinical guidelines, might be due to different reasons including: a) the scarcity of methodological research on the updating process [8]; b) lack of guidance in handbooks of clinical guideline institutions [9,10]; c) a clear emphasis on developing rather than updating in most organisations over the last two decades [5].

In CheckUp, the format of updated clinical guidelines, recommendations, and updated sections is covered in items 1 to 6. Nowadays, there is a gap in knowledge regarding the presentation formats of updated clinical guidelines. An analysis of clinical guidelines updated by the National Institute of Care Excellence (NICE), a key organisation in the clinical guideline area, observed concerning variability in the presentation formats of the recommendations' changes, and a lack of justification for those changes [26]. The DECIDE (Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence; http://www.decide-collaboration.eu/) project, an initiative by the GRADE working group, has developed several presentation formats for recommendations of clinical guidelines [27,28]. However, the project did not address the presentation of updated recommendations.

A description of the updating panel, the disclosures of interest, and the role of the funding body are covered in items 7 to 9. Both, the role in the working group and disciplines of the various members of the panel, responsible for updating a clinical guideline need to be carefully taken into account [29]. Similarly, the conflicts of interest of the members of the original clinical guideline should be regularly reviewed throughout the process of monitoring and updating. [30].

In CheckUp, the methodology of the updating process is covered in items 10 to 16. Updating of a clinical guideline should be performed with a similar degree of rigour and explicitness as in the development of a de novo clinical guideline. However, in practice updating clinical

guidelines seems to be methodologically poorer conducted than developing de novo clinical guidelines [31].

Although the checklist provides methodological principles on updating and reporting, it does not advocate a single operation strategy for updating a clinical guideline. There is no evidence for a "gold standard", there are is little research evidence in the field, and often the original de novo methodology is used during the updating process [8]. In response to this gap, CheckUp is a first step, determining what content has to be reported in updated clinical guidelines.

Strengths and limitations

We believe that the strength of our study concerns the elaborate development process of the CheckUp. For the development process we followed the criteria of the EQUATOR (Enhancing the QUAlity and Transparency Of health Research) network and included the recommendations stated by Moher et al (2011) [32,33]. In the development process of CheckUp we used several methods to ensure inclusion of developers´ and users´ input internationally. By applying a formal consensus method (Delphi survey) and collecting experts' opinions using diverse methods (semi-structured interviews and external reviews), we reached a fair understanding of clinical guideline methodologists' and users' perceptions of the updating of clinical guidelines. Complemented by empirical evidence from the literature and an assessment of updated clinical guidelines, we believe that CheckUp includes all relevant aspects of updating clinical guidelines.

CheckUp has some limitations. One is the use of consensus methods with convenience samples of clinical guideline stakeholders. There is, hence, a potential bias with our sampling frames. However, across the different processes, an alignment and consensus of opinion emerged on what clinical guideline developers, users, and researchers expect to see reported in an updated clinical guideline. Additionally, for all items we tried to present optimal presentation formats examples from published updated clinical guidelines. However, we were not able to find for some items real examples of adequate presentation formats. Therefore, illustrative examples of

what we believe to be realistic and adequate reporting have been included. However, our use of a particular example does not imply that the terminology used is mandatory.

Implications for practice and research

CheckUp is a checklist for users and clinical guideline appraisers to assess the reporting of the updating process in updated clinical guidelines. Besides, CheckUp can be used by clinical guideline developers as a helping checklist for planning and conducting the reporting of clinical guidelines that need to be updated. The checklist can, hence, provide guidance to developers who are updating clinical guidelines through methodological principles that should be incorporated into the clinical guideline updating process, and strategies for reporting the clinical guideline.

CheckUp will be updated in the near future whenever real examples of adequate reporting will be identified. We advise the updating panels that are responsible for updating clinical guidelines to take the items of CheckUp into account. Additionally, there is a need to evaluate the reporting of updated guidelines with CheckUp. It would be desirable to also assess the impact of this checklist in the clinical guideline field over the next few years. The Guidelines International Network (G-I-N) Updating Working Group [www.g-i-n.net/working-groups/updating-guidelines] will be playing a crucial role in this work and in moving forward the updating agenda in the clinical guideline enterprise.

In addition, like every resource designed to improve evidence-based methodology, we recognize that CheckUp is the first iteration of the checklist. As specific methods for updating continue to best tested and evolve, and standard practices emerge, these innovations will be reflected in future revisions of CheckUp.

Acknowledgements

Vernooij RWM, Alonso-Coello P, Brouwers M, Martínez García L, Florez ID, Iorio A, James R, Sanabria AJ, Selva A, Shekelle PG, and Vandvik PO are members of the G-I-N Updating Guidelines Working Group. The members of the CheckUp panel are: Vernooij RWM, Alonso-Coello P, Brouwers M, Martínez García L, Ada L, Alemán A, Arévalo-Rodriguez I, Becker M, Burgers JS, Chan W, Delvaux N, Duggan G, Enciso Olivera CO, Etxeandia-Ikobaltzeta I, Florez ID, Follmann M, Gartlehner G, Iorio A, James R, Jones SL, Kotzeva A, Lloyd M, López Gallegos D, Louro-González A, Marin Leon I, Martí-Carvajal A, Meerpohl JJ, Pardo R, Rojas-Reyes MX, Rotaeche R, Sanabria AJ, Selva A, Shekelle PG, Sierra Matamoros FA, van de Velde S, Vandvik PO, Willett S.

Robin W.M. Vernooij is a doctoral candidate at the Paediatrics, Obstetrics and Gynaecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain. The authors thank Sandra Pequeño Saco (Iberoamerican Cochrane Centre) for her support in assessing updated clinical guidelines. The authors also thank María Victoria Leo for her help editing the manuscript.

References

1. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. Ann Intern Med. 2011;154(11): 774-775.
2. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P, et al. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med. 2012;156(7):525-31.
3. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, Woolf SH. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001;286(12):1461-7.
4. Grupo de trabajo sobre actualización de GPC. Actualización de Guías de Práctica Clínica en el Sistema Nacional de Salud. Manual Metodológico. Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad y Política Social. Instituto Aragonés dológico. de Ciencias de la Salud-I+CS; 2009. Guías de Práctica Clínica en el SNS: I+CS Nº 2007/02-01
5. Shekelle PG. Updating practice guidelines. JAMA. 2014;311(20):2072-3.
6. Burgers JS, Cluzeau FA, Hanna SE, Hunt C, Grol R. Characteristics of high-quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. Int J Technol Assess Health Care. 2003;19(1):148-57.
7. Alonso-Coello P, Martínez García L, Carrasco Gimeno JM, Solà I, Qureshi S, Burgers JS; Updating Guidelines Working Group. The updating of clinical practice guidelines: insights from an international survey. Implement Sci 2011, 6:107.
8. Martínez García L, Arévalo-Rodríguez I, Solà I, Haynes RB, Vandvik PO, Alonso-Coello P; Updating Guidelines Working Group. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci 2012, 7(1):109.
9. Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol. 2014;67(1):33-45.

10. Vernooij RW, Sanabria AJ, Solà I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci. 2014;9:3.
11. Vernooij RWM, Alonso-Coello P, Brouwer MC, Martínez García L; CheckUp Panel. CheckUp: reporting items for updated clinical guidelines. Submitted.
12. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L; AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.
13. Saiman L, Siegel JD, LiPuma JJ, Brown RF, Bryson EA, Chambers MJ, Downer VS, Fliege J, Hazle LA, Jain M, Marshall BC, O'Malley C, Pattee SR, Potter-Bynoe G, Reid S, Robinson KA, Sabadosa KA, Schmidt HJ, Tullis E, Webber J, Weber DJ. Infection prevention and control guideline for cystic fibrosis: 2013 update. Infect Control Hosp Epidemiol. 2014;35:Suppl 1:S1-S67.
14. NICE (2014) Advanced breast cancer (update. Diagnosis and treatment NICE clinical guideline 81). Available at https://www.nice.org.uk/Guidance/CG81 [NICE guideline].
15. [NICE (2013) Fertility: Assessment and treatment for people with fertility problems. Available at https://www.nice.org.uk/guidance/CG156 [NICE guideline]
16. Cohen SH, Gerding DN, Johnson S, Kelly CP, Loo VG, McDonald LC, Pepin J, Wilcox MH; Society for Healthcare Epidemiology of America; Infectious Diseases Society of America. Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). Infect Control Hosp Epidemiol. 2010;31(5):431-55.
17. Scottish Intercollegiate Guidelines Network (SIGN). Antibiotic prophylaxis in surgery. Edinburgh: SIGN; 2008. (SIGN publication no.104). [July 2008]. Available from URL: http://www.sign.ac.uk

18. NICE (2012) Infection: prevention and control of healthcare-associated infections in primary and community care Available at https://www.nice.org.uk/guidance/CG139 [NICE guideline].
19. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Annals Internal Medicine. 2001;134(8):663-94.
20. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Plos Med. 2009;6(7):e1000100.
21. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med. 2008;148(4):295-309.
22. Welch V, Petticrew M, Petkovic J, Moher D, Waters E, White H, Tugwell P; PRISMA-Equity Bellagio group. Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E 2012): explanation and elaboration. J Clin Epidemiol. 2015 Sep 5. pii: S0895-4356(15)00420-5.
23. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349:g7647.
24. Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleza-Jeric K, Laupacis A, Moher D. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013 Jan 8;346:e7586.
25. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015 Jan 6;162(1):W1-73.

644 26. Martínez García L, Sanabria AJ, García Alvarez E, Trujillo-Martín MM, Etxeandia-
645 Ikobaltzeta I, Kotzeva A, Rigau D, Louro-González A, Barajas-Nava L, Díaz Del Campo P,
646 Estrada MD, Solà I, Gracia J, Salcedo-Fernandez F, Lawson J, Haynes RB, Alonso-Coello
647 P; Updating Guidelines Working Group. The validity of recommendations from clinical
648 guidelines: a survival analysis. CMAJ. 2014;186(16):1211-9.
649 27. Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brożek J, Davoli M, Flottorp
650 S, Harbour R, Hill S, Liberati A, Liira H, Schünemann HJ, Rosenbaum S, Thornton J,
651 Vandvik PO, Alonso-Coello P; DECIDE Consortium. Developing and Evaluating
652 Communication Strategies to Support Informed Decisions and Practice Based on Evidence
653 (DECIDE): protocol and preliminary results. Implement Sci. 2013;8:6.
654 28. Kristiansen A, Brandt L, Alonso-Coello P, et al. Development of a novel electronic
655 multilayered presentation format for clinical practice guidelines. CHEST 2014, early online
656 September 2014
657 29. Schünemann HJ, Al-Ansary LA, Forland F, Kersten S, Komulainen J, Kopp IB, Macbeth F,
658 Phillips SM, Robbins C, van der Wees P, Qaseem A; Board of Trustees of the Guidelines
659 International Network. Guidelines International Network: Principles for Disclosure of
660 Interests and Management of Conflicts in Guidelines. Ann Intern Med. 2015;163(7):548-53.
661 30. Grupo de trabajo sobre actualización de GPC. Elaboración de Guías de Práctica Clínica en
662 el Sistema Nacional de Salud. Actualización del Manual Metodológico. March 2016.
663 Available: http://portal.guiasalud.es/emanuales/elaboracion_2/
664 31. Hasenfeld R, Shekelle PG. Is the methodological quality of guidelines declining in the US?
665 Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR)
666 guidelines with those published subsequently. Qual Saf Health Care. 2003;12(6):428-34.
667 32. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research
668 reporting guidelines. Plos Med. 2010;7(2):e1000217.
669 33. Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG, Schulz KF, Miller D,
670 Simera I, Grimshaw J, Hoey J. Describing reporting guidelines for health research: a
671 systematic review. J Clin Epidemiol. 2011;64(7):718-74

APPENDIX II

Article: "Efficiency of pragmatic search strategies to update clinical guidelines recommendations"

Martínez García et al. BMC Medical Research Methodology (2015) 15:57
DOI 10.1186/s12874-015-0058-2

BMC Medical Research Methodology

RESEARCH ARTICLE Open Access

Efficiency of pragmatic search strategies to update clinical guidelines recommendations

L. Martínez García[1*], AJ. Sanabria[1], I. Araya[2], J. Lawson[3], I. Solà[1], RWM. Vernooij[1], D. López[4], E. García Álvarez[5], MM. Trujillo-Martín[6], I. Etxeandia-Ikobaltzeta[7], A. Kotzeva[8,9], D. Rigau[1], A. Louro-González[10], L. Barajas-Nava[1], P. Díaz del Campo[11], MD. Estrada[8,9], J. Gracia[12], F. Salcedo-Fernandez[13], RB. Haynes[3], P. Alonso-Coello[1,9] and Updating Guidelines Working Group

Abstract

Background: A major challenge in updating clinical guidelines is to efficiently identify new, relevant evidence. We evaluated the efficiency and feasibility of two new approaches: the development of restrictive search strategies using PubMed Clinical Queries for MEDLINE and the use of the PLUS (McMaster Premium Literature Service) database.

Methods: We evaluated a random sample of recommendations from a national guideline development program and identified the references that would potentially trigger an update (key references) using an exhaustive approach.
We designed restrictive search strategies using the minimum number of Medical Subject Headings (MeSH) terms and text words required from the original exhaustive search strategies and applying broad and narrow filters. We developed PLUS search strategies, matching Medical Subject Headings (MeSH) and Systematized Nomenclature of Medicine (SNOMED) terms with guideline topics. We compared the number of key references retrieved by these approaches with those retrieved by the exhaustive approach.

Results: The restrictive approach retrieved 68.1 % fewer references than the exhaustive approach (12,486 versus 39,136), and identified 89.9 % (62/69) of key references and 88 % (22/25) of recommendation updates. The use of PLUS retrieved 88.5 % fewer references than the exhaustive approach (4,486 versus 39,136) and identified substantially fewer key references (18/69, 26.1 %) and fewer recommendation updates (10/25, 40 %).

Conclusions: The proposed restrictive approach is a highly efficient and feasible method to identify new evidence that triggers a recommendation update. Searching only in the PLUS database proved to be a suboptimal approach and suggests the need for topic-specific tailoring.

Keywords: Clinical guidelines, Diffusion of innovation, Dissemination and implementation, Evidence-based medicine, Information storage and retrieval, Knowledge translation, Methods, Updating

Background

Clinical guidelines, like systematic reviews and other evidence summaries, require periodic reassessment of research evidence to remain valid [1–4]. Current guidance usually recommends revision and update within two to three years of their publication [5, 6]. New evidence to update clinical guidelines is generally identified using the original exhaustive search strategies [7].

A major challenge for guideline developers is to efficiently screen for new, relevant evidence that justifies a clinical guideline update. Unfortunately, little empirical work has been conducted to date to test the effectiveness and efficiency of searching processes [7]. More than a decade ago, Shekelle et al. developed a strategy based on retrieving reviews, editorials, and commentaries in high impact general journals and specialised journals, complemented with a survey by clinical experts [8].

* Correspondence: laura.martinez.garcia@cochrane.es
[1]Iberoamerican Cochrane Centre - Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain
Full list of author information is available at the end of the article

Gartlehner et al. compared a modified version of this strategy versus an exhaustive search strategy [9]. The results so far have shown that restrictive approaches are promising, but more information is needed about the timing and type of search [7].

Similarly, researchers are testing alternative search strategies to update systematic reviews [10–13]. Haynes et al. developed the McMaster Premium Literature Service (PLUS) database, from the McMaster Health Knowledge Refinery [14, 15]. PLUS contains a searchable subset of pre-appraised primary studies and systematic reviews from more than 120 journals and it can identify key articles needed to update systematic reviews [14, 15]. Clinical Queries search filters in MEDLINE and EMBASE have also shown a high sensitivity to detect key articles [11].

We designed a study to evaluate the efficiency and feasibility of two approaches to identify the need to update clinical guidelines recommendations: 1) restrictive search strategies using PubMed Clinical Queries search filters for MEDLINE and 2) the use of PLUS database.

Methods

Design

We conducted a descriptive study of search strategies to identify the references that update recommendations from clinical guidelines. We developed three search strategies to identify the need to update the recommendations: an exhaustive approach, a restrictive approach, and a PLUS approach.

The sample was obtained from a previous study and included a stratified random sample of recommendations from the Spanish National Health System Clinical Guidelines Program [1, 16]. The selection process involved two phases: 1) we stratified guidelines by topic and by year of publication; when multiple guidelines per strata were available, we randomly selected one; 2) we performed a stratified random sampling of recommendations by guideline topic and by turnover (number of pertinent references linked per recommendation in the updating process).

1) Exhaustive approach

Guideline methodologists with experience designing search strategies developed exhaustive literature search strategies for each clinical question: 1) based on the original searches; and 2) applying the filters of the original study. An example of the exhaustive search strategy is available in Additional file 1. We also contacted clinical experts to identify new studies. We obtained a reference database of clinical questions. We screened the references and assessed them qualitatively as: 1) *Pertinent references:* Randomised controlled trials or systematic reviews related to the topic of the clinical guideline; 2) *Relevant references:* pertinent references that could be used when considering an update to a recommendation, but that would not necessarily trigger a potential update; and 3) *Key references:* relevant references that would potentially trigger an update because of their impact on the population, the intervention, the comparison, the outcome, the quality of the evidence, the direction and/or the strength of the recommendation. Using the results of the reference screening we classified recommendations as: 1) need for updating*:* with one or more key references linked; or 2) still valid: without key references linked.

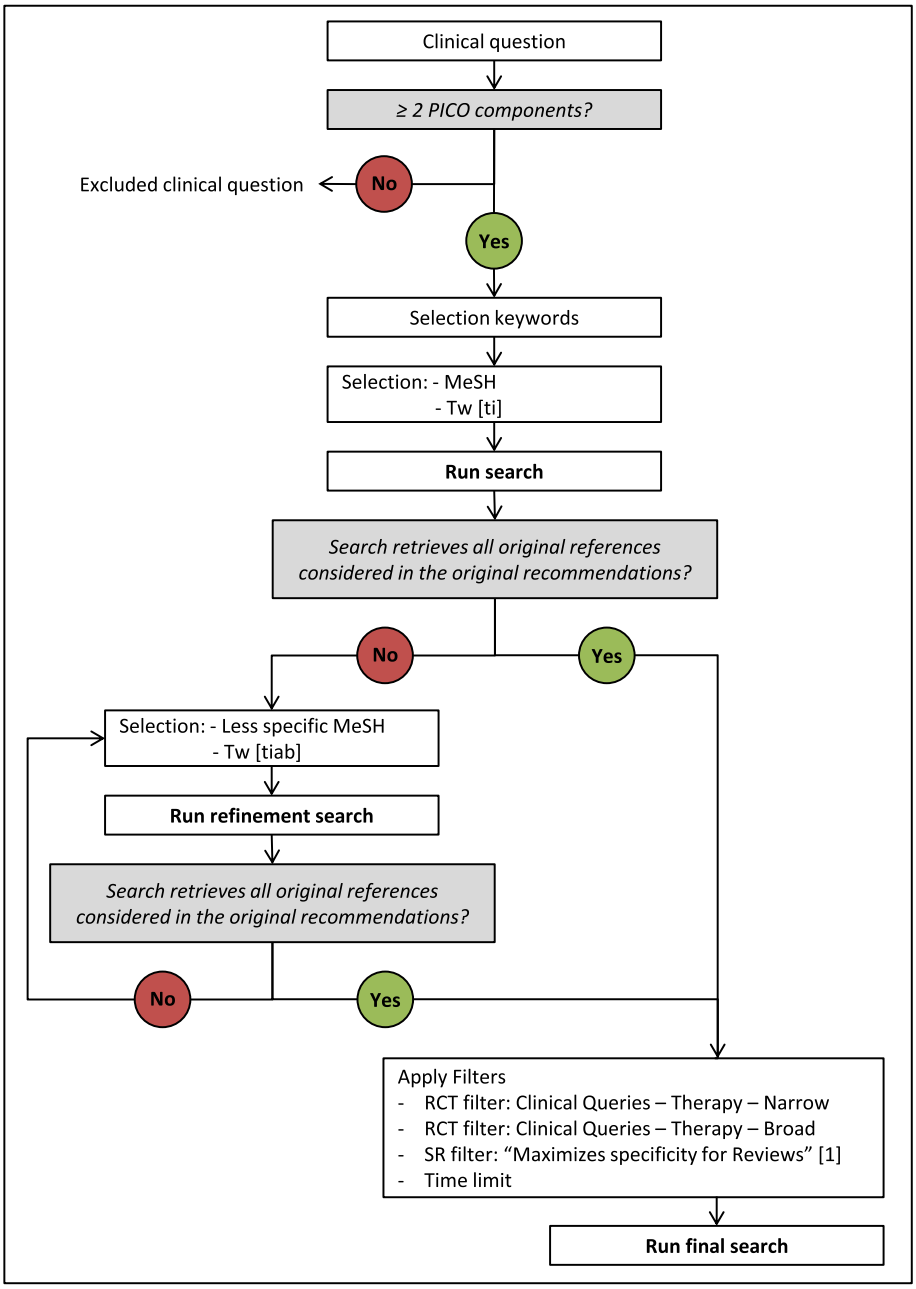
A more complete description of this approach is available in the previously published protocol and survival analysis results [1, 16].

2) Restrictive approach

Guideline methodologists, trained by researchers with experience designing search strategies, developed restrictive search strategies for each clinical question using the PubMed Clinical Queries search filters for the MEDLINE database. We considered clinical questions that had at least two PICO (population, intervention, comparator or outcome) components. We developed the restrictive search strategies considering the minimum number of Medical Subject Headings (MeSH) terms and text words required from the original exhaustive searches strategies. The search strategies were designed in four stages [Fig. 1]: 1) Development: we selected keywords from the clinical questions and identified Medical Subject Headings (MeSH) terms and text words in titles; 2) Validation: we evaluated whether each search retrieved all the original references for its corresponding recommendation; 3) Refinement: If a search did not retrieve all the original references, we selected and searched less specific Medical Subject Headings (MeSH) and/or text words in the title or abstract; and 4) Application of each of a broad and a narrow treatment Clinical Queries filter (www.ncbi.nlm.nih.gov/pubmed/clinical), and a systematic review filter [17]. We used the same date limits as with the exhaustive approach (from the complete year in which the original exhaustive searches was completed onwards). An example of a restrictive search strategy is available in Additional file 1.

3) PLUS approach

An information specialist from the Health Information Research Unit developed a PLUS search strategy for each guideline topic. We matched Medical Subject Headings (MeSH) and Systematized Nomenclature of Medicine (SNOMED) indexing terms in the PLUS database with clinical guideline topics. Both primary and review papers were included. To take into account the time delay associated with the critical appraisal process (CAP) the articles go through, we ran the PLUS searches strategies from the beginning of the year in which



Abbreviations: MeSH: Medical Subject Headings; RCT: Randomized controlled trial; SR: Systematic review; Ti: Title; Tiab: Title or abstract; Tw: Text words

1. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ. 2005;330(7482):68.

Fig. 1 Restrictive approach algorithm. Figure01_RestrictiveApproachAlgorithm. Microsoft PowerPoint presentation

the original exhaustive searches were run, until approximately three months beyond the latest date of the exhaustive searches. An example of a PLUS search strategy is available in Additional file 1.

Outcome

Our primary outcome was the number of key references identified by each alternative approach.

Statistical methods

We performed a descriptive analysis of the data. We calculated absolute and relative frequencies or median and range, as appropriate.

Two investigators independently retrieved the key references (identified in the exhaustive approach) in each of the alternative approach results. We analysed the number of key references in: 1) the results of restrictive

search strategies per clinical question; 2) restrictive search strategies results per clinical guideline (clustering all references identified by clinical question) [Fig. 2]; and 3) results of PLUS strategies per clinical guideline. We did not identify additional pertinent, relevant or key references from the alternative approaches. We did not develop restrictive search strategies for clinical questions with less than two of the four PICO components, prognosis or diagnostic clinical questions. In these instances we used the updated exhaustive search strategies.

We identified the recommendations that needed an update (with one or more key references) retrieved by each alternative approach. We compared the recommendations identified with those that were not identified according to clinical guideline topic (cancer, cardiovascular disease, mental health or metabolic disease), strength of recommendation (A, B, C, D or good practice point [18]), clinical purpose (prevention, screening, diagnosis, treatment or other), and turnover. Each recommendation was classified according to the number of linked pertinent references: none, ≤ median number (low turnover), or > median number (high turnover). We used Pearson's chi-square test or Fisher's exact test, as appropriate.

We recorded the number of hours spent on designing each approach and the number of researchers involved.

We accepted p values of less than 0.05 as significant in all calculations. We performed the analyses using SPSS 21.0 (SPSS Inc., Chicago, Illinois).

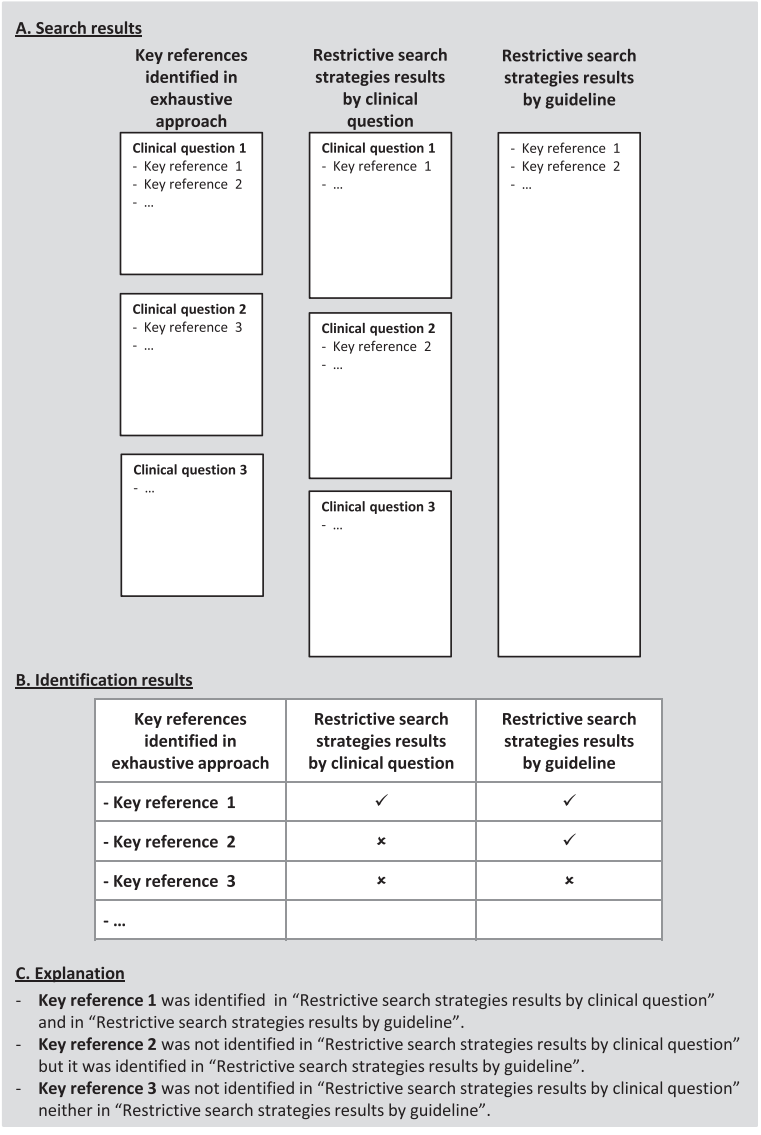


Fig. 2 References analysis. Figure02_ReferencesAnalysis. Microsoft PowerPoint presentation

Results

We included a cohort of four clinical guidelines from the Spanish National Health System Clinical Guidelines Programme, corresponding to 87 clinical questions and 249 recommendations [19–22]. After the random selection process, the final recommendation sample included 43 clinical questions and 113 recommendations.

Exhaustive approach results

This approach retrieved a total of 39,136 references from the four clinical guidelines included. From the recommendations sample, we identified a total of 69 key references and 25 recommendations that potentially needed an update [Table 1].

Restrictive approach results

We applied the restrictive approach to 88.5 % (77/87) clinical questions from the included clinical guidelines, corresponding to 85 % (96/113) of the recommendations from our recommendation sample. We excluded eight questions that did not present a minimum of two PICO components (population, intervention, comparator or outcome) and one diagnostic question.

The restrictive searches covered a mean of 4.6 years (range 3.9 – 5.1 years) from 2008–2009 to 2011 – 2012 [Table 2].

For the clinical guidelines included, we retrieved a total of 40,021 references using the broad filter and 9,958 references using the narrow filter [Table 2]. We retrieved more key references when we clustered results of references per guideline rather than per question (40 [87 %] and 39 [84.8 %] compared with 26 [56.5 %] and 25 [54.3 %] using the broad and narrow filters, respectively) [Table 2, Additional file 2]. Similarly, clustered results of references per guideline identified a higher number of recommendations that were considered to potentially need an update (18 [90.0 %] and 17 [85 %] compared with 15 [75 %] and 14 [70 %] respectively [Table 2].

When we used exhaustive search strategies for the clinical questions not developed by the restrictive approach (narrow filter and clustering by all questions), we retrieved a total of 12,486 references, and we identified a total of 62 (89.9 %) key references and 22 (88.0 %) recommendations that potentially needed an update [Table 4].

The restrictive approach (narrow filter and clustering by all questions) failed to identify seven key references (15.2 %): four (57.1 %) references were systematic reviews and three references (42.9 %) were congress abstracts (not indexed in MEDLINE) [Fig. 3].

The recommendations that potentially needed an update not identified by the restrictive approach were similar to those that were identified in terms of topic, strength of the recommendations, clinical purpose, and turnover [Additional file 3].

PLUS approach results

The PLUS searches covered a median of 5.0 years (range 4.1 – 5.3 years) from 2008–2009 to 2011 – 2012 [Table 3].

For the clinical guidelines included, we retrieved a total of 4,486 references (range 137 – 3,059) [Table 3]. For the recommendation sample, we retrieved 18 (26.1 %) key references; these references potentially update 10 (40 %) recommendations [Table 3, Additional file 2].

The PLUS approach failed to identify 51 key references (73.9 %); most (41 references, 80.4 %) were from journals not included in PLUS database [Fig. 4].

Recommendations with a high turnover were more likely to be identified by the PLUS approach. The remaining factors (clinical guideline topic, strength of the recommendations, and clinical purpose) were not significantly associated with the need to update [Additional file 3].

Resource use

Three guideline methodologists spent a total of 174 h in designing and running the restrictive search strategies [Table 4]. The PLUS search strategies were developed by an information specialist who designed and ran the searches in 28 h [Table 4].

Discussion

We evaluated two search strategies to identify signals for updating recommendations and compared them to an

Table 1 Exhaustive approach results

	Major depression in adults 2008 [19]	Obesity in childhood and adolescence 2009 [20]	Prostate cancer treatment 2008 [21]	Secondary prevention of stroke 2009 [22]	Total
Search period (years)	4.8	3.9	4.5	5.1	
References retrieved in search for clinical guidelines, n	11243	9763	3343	14787	39136
Key references identified from recommendation sample, n	13	32	11	13	69
Potential update recommendations identified from recommendation sample, n (%)	3	8	7	7	25

Table 2 Restrictive approach results

	Major depression in adults 2008 [19]		Obesity in childhood and adolescence 2009 [20]		Prostate cancer treatment 2008 [21]		Secondary prevention of stroke 2009 [22]		Total	
Search period (years)	4.8		3.9		4.5		5.1			
References retrieved in search for clinical guidelines, n										
- Broad filter	9223		10561		6939		13294		40017	
- Narrow filter	2814		3976		976		2187		9953	
Key references identified from recommendation sample, n (%)[a]										
- Exhaustive approach[b]	13		16		4		13		46	
- Broad filter										
by individual clinical questions	5	38.5	11	68.8	4	100.0	6	46.2	26	56.5
by clustering all clinical questions	11	84.6	16	100.0	4	100.0	9	69.2	40	87.0
- Narrow filter										
by individual clinical questions	4	30.8	11	68.8	4	100.0	6	46.2	25	54.3
by clustering all clinical questions	10	76.9	16	100.0	4	100.0	9	69.2	39	84.8
Potential update recommendations identified from recommendation sample, n (%)[a]										
- Exhaustive approach[b]	3		6		4		7		20	
- Broad filter										
by individual clinical questions	3	100.0	4	66.7	4	100.0	4	57.1	15	75.0
by clustering all clinical questions	3	100.0	6	100.0	4	100.0	5	71.4	18	90.0
- Narrow filter										
by individual clinical questions	2	66.7	4	66.7	4	100.0	4	57.1	14	70.0
by clustering all clinical questions	2	66.7	6	100.0	4	100.0	5	71.4	17	85.0

[a]Percentage of references and recommendations identified regarding the exhaustive strategy
[b]Exhaustive strategy results without clinical questions and recommendations not included in ReSe strategy

exhaustive search strategy using a random sample of recommendations from a cohort of clinical guidelines from a national guideline development program.

The restrictive approach (using a narrow PubMed Clinical Queries filter, clustering results per clinical guideline and imputing exhaustive search results for clinical questions not developed) retrieved 68.1 % fewer references than the exhaustive approach, and identified most of the key references (62/69, 89.9 %) and recommendations updates (22/25, 88.0 %). We developed search strategies for each clinical question but obtained better results by considering the results across all questions included in a clinical guideline. The restrictive approach proved to be relatively simple to develop, not needing the expertise of information retrieval specialists. Over half of the very few missing key references with this approach were systematic reviews. Three references were missed due to a mistake in the design of restrictive searches, and one was missed by the filter used [17], reflecting the need to pay more attention to the design and quality check of search strategies. Additional searches for systematic reviews in specific databases, like Epistemonikos, could prove useful [www.epistemonikos.org/]).

Our results show that PLUS approach retrieved 88.5 % fewer references than the exhaustive approach but identified a substantially lower number of key references (18/69, 26.1 %) and potential updates (10/25, 40 %) than the restrictive approach. These results were similar independently of the searches being performed by a PLUS information specialist (using search strategies) or directly using the PLUS interface using topic

Why Restrictive Approach (narrow filter and clustering all questions) did not identify 7 key references?

4 (57.1%) references were SRs

3 (42.9%) references were congress conferences (without PMID)

Abbreviations: PMID: PubMed Unique Identifier; ReSe: Restrictive Search; SR: Systematic review.

Fig. 3 Key references not identified by restrictive approach. Figure03_RefNotIdentifiedRestrictive. Microsoft PowerPoint presentation

Table 3 PLUS approach results

	Major depression in Adults 2008 [19]		Obesity in childhood and adolescence 2009 [20]		Prostate cancer treatment 2008 [21]		Secondary prevention of stroke 2009 [22]		Total	
Search period (years)	5.3		4.1		4.8		5.3			
References retrieved in search for clinical guidelines, n	973		317		137		3059		4486	
Key references identified from recommendation sample, n (%)[a]										
- Exhaustive strategy	13		32		11		13		69	
- PLUS strategy	4	(30.8)	9	(28.1)	1	(9.1)	4	(30.8)	18	(26.1)
Potential update recommendations identified from recommendation sample, n (%)[a]										
- Exhaustive strategy	3		8		7		7		25	
- PLUS strategy	2	(66.7)	4	(50.0)	1	(14.3)	3	(42.9)	10	(40.0)

[a]Percentage of references and recommendations identified regarding the exhaustive strategy

synonyms (*post-hoc analysis*). This poor performance was mainly due to most of these key references (80.4 %) being from journals not included in PLUS database.

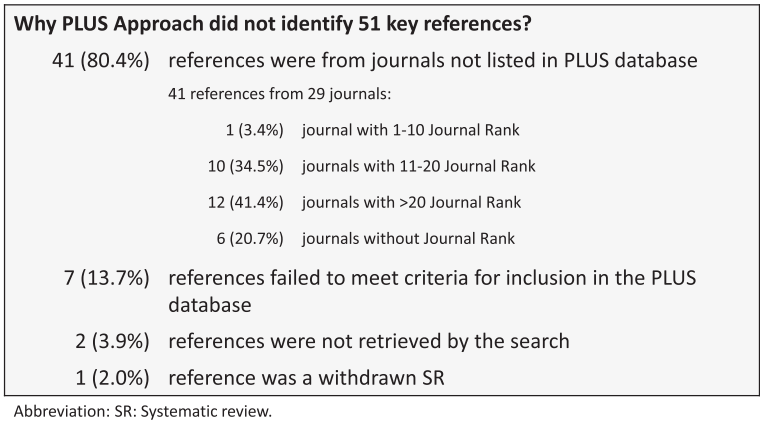
The PLUS approach performed differently across topics with major depression performing best (66.7 % of key references retrieved) and prostate cancer worst (14.3 %). This poor performance in the prostate cancer guideline is explained by the fact that the PLUS database does not include a large number of urology journals. This resource includes a limited number of journals with a stronger focus on a limited number of specialties and health topics. Given these findings and building on previous research in the systematic reviews and clinical guidelines fields, post-hoc we explored a potential approach of tailoring the PLUS approach by adding a limited number of journals for each specialty (e.g. those with a higher impact factor) [8, 9, 12, 13]. However, missing key references were published in a highly heterogeneous sample of journals, with only 3.4 % being in the first decile [Fig. 4].

The two search strategies we tested were far less time consuming than the exhaustive search strategy. The restrictive approach needs initial tailoring and takes each original guideline, question, search and references into account. In contrast, the PLUS approach could be potentially executed directly in its interface simply using topic synonyms from clinical guidelines.

Our results in the context of previous research

Only one previous study of clinical guidelines compared a different type of restrictive approach versus an exhaustive approach [9]. However, this study considered prevention topics as the unit of analysis rather than the individual recommendations. Furthermore, the authors restricted the search to MEDLINE, using publication types (review articles, editorials, guidelines and commentaries) and limiting the search to core and specialty clinical journals [9].

A recent evaluation of NICE clinical guidelines for interventional procedures also showed that updated recommendations that required a modification generally had a



Abbreviation: SR: Systematic review.

Fig. 4 Key references not identified by PLUS approach. Figure04_RefNotIdentifiedPLUS. Microsoft PowerPoint presentation

Table. 4 Summary results by approach

	Exhaustive approach		Restrictive approach[a]		PLUS approach	
	n	%	n	%[b]	n	%[b]
References identification						
References retrieved in search for clinical guidelines	39136		12486	31.9	4486	11.5
Key references identified from recommendation sample	69		62	89.9	18	26.1
Recommendation identification						
Potential update recommendations identified from recommendation sample	25		22	88.0	10	40.0
Resource use						
Guidelines methodologists	4		3	75.0	-	
Information specialist	-		-		1	25.0
Time to perform the search (hours)	279		174.3	62.5	28	10.0

[a]Narrow filter, clustered by all questions, and imputed exhaustive search results for the clinical questions not included in the restrictive approach
[b]Percentage regarding the exhaustive approach

greater increase in their evidence base (number of patients included in observational studies published) than non-updated recommendations [23]. Our results are consistent with this finding, showing a higher efficiency of the PLUS approach in recommendations with a higher turnover.

There is indirect evidence about the performance of PLUS for clinical guidelines from a previous study that evaluated the updating of systematic reviews [11]. Only 13 out of 87 systematic reviews (14.9 %) included all the new studies in PLUS. In 39 (44.8 %) reviews there was no statistically significant difference between PLUS and non-PLUS new studies (ROR: 0.99; 95 % confidence interval: 0.87-1.14). Thirty-five updated reviews (40.2 %) had no new studies indexed in PLUS (although conclusions were seldom altered by addition of new studies) [11]. Despite these results in systematic reviews, the PLUS database did not perform similarly in the context of clinical guidelines. However, we did not routinely determine the change in effect sizes with key references, so we could not assess their quantitative relationship. Neither did we assess whether references identified in the PLUS database could have reliably signalled the need to update for topics that were in the journals that are included.

The same study by Hemens et al. confirmed the high sensitivity of Clinical Queries filters for MEDLINE and EMBASE in detecting randomized controlled trials [11]. This is consistent with our results showing that incorporating Clinical Queries filters (to identify randomized controlled trials) and Montori's et al. filter (to identify systematic reviews) significantly reduces the citation screening burden [17].

Strengths and limitations

We used a rigorous and explicit methodology building on previous research in this area, improving its deficiencies, and implementing an innovative solution. We also used the exhaustive approach as a standard, improving the validity of the results and, hence, the strength of our inferences. We independently screened and extracted the data in pairs and included methodologists and panel members from the original guidelines as far as possible. Finally, we laid out a structured framework (e.g., outcome definitions) that could prove useful in the future for other researchers in the field.

Our study has some limitations. We did not assess all references retrieved by each alternative approach, so we were not able to evaluate whether other key references were identified by any of these approaches. Our sample is limited to recommendations from four guidelines topics. However, this potential limitation is mitigated because our sample covers broad areas such as cancer, cardiovascular diseases, mental health and lifestyle and behavioural issues. Additionally, we based our exhaustive search strategies on searches specifically designed during the original guidelines development. A post-hoc analysis revealed several mistakes and inconsistencies in search strategies that could have been avoided through peer review process [24]. However, the validation of the accuracy of the original search strategies was beyond the scope of our study. We are unable to estimate how this issue could affect the recall of the exhaustive search strategies, although we think that these deficiencies are minor and that they do not alter our conclusions. We included only randomised controlled trials and systematic reviews and did not incorporate observational studies, diagnostic questions or evidence about values and preferences or resource use considerations. Finally, some authors had conflicts of interest due to their involvement in the PLUS database and Clinical Queries filter development. However, they did not participate in the identification of key references.

Conclusions

Our results have important implications both for the updating of guidelines and for future research in this field. The proposed method of developing restrictive search strategies, using PubMed Clinical Queries filters in the MEDLINE database, provides a feasible and efficient method for guideline developers to identify significant new studies that are likely to trigger a recommendation update. Searching only in the PLUS database was a suboptimal approach that needs topic specific tailoring.

Our results highlight the need for additional methodological research in this field. For this future work, investigators are likely to find our framework helpful.

Additional files

Additional file 1: Search strategies examples. We reported an example of exhaustive strategy, restrictive strategy, PLUS strategy. (PDF 82 kb)

Additional file 2: Key references by approach. We reported key references by strategy, according to their clinical guidelines and linked to recommendation. (PDF 62 kb)

Additional file 3: Additional tables. We reported complementary results (PDF 40 kb)

Competing interests

Laura Martínez García, Andrea Juliana Sanabria, David Rigau, Leticia Barajas-Nava, Ivan Solà and Pablo Alonso-Coello have received research grants from Instituto de Salud Carlos III (FIS PI10/00346). R. Brian Haynes and Jennifer Lawson have received research grants from the Canadian Institutes of Health Research. R. Brian Haynes and Jennifer Lawson have been involved in developing the McMaster PLUS Project and RBH originated the development of Clinical Queries, so did not participate in identifying key references or potential update recommendations. No other competing interests were declared.

Authors' contributions

LMG, AJS, IA, PAC, DR and IS conceived the idea of the study. IA, JL, IS, RWMV and DL designed and/or ran searches strategies. EGA, MMTM, IEI, AK, DR, ALG, LBN, PDC and MDE screened the references. LMG and AJS performed the statistical analysis. LMG and PAC drafted the manuscript. All of the authors revised the manuscript critically for important intellectual content and approved the final version submitted for publication.

Acknowledgements

The authors thank Mercedes Cabañas (Técnicas Avanzadas de Investigación en Servicios de Salud [TAISS], Madrid, Spain) and Joanna Kelly (Healthcare Improvement Scotland, Glasgow, UK) for their support in developing and running the exhaustive literature searches, and Carolyn Newey for her help editing the manuscript.
Laura Martínez García is a doctoral candidate at the Pediatrics, Obstetrics and Gynecology, and Preventive Medicine Department, Universitat Aunònoma de Barcelona, Barcelona, Spain.

Data sharing

The dataset is available from the corresponding author at laura.martinez.garcia@cochrane.es.

Funding

This project is funded by research grants from the Instituto de Salud Carlos III (FIS PI10/00346 to Pablo Alonso-Coello) and from the Canadian Institutes of Health Research (R. Brian Haynes). Laura Martínez García and Andrea Juliana Sanabria are funded by Río Hortega research contracts from the Instituto de Salud Carlos III (CM11/00035 and CM12/00168 respectively). Pablo Alonso-Coello is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (CP09/00137).

Members of the updating guidelines working group

Martínez García L; Sanabria AJ; Araya I; Lawson J; Solà I; Vernooij RWM; López D; García Álvarez E; Trujillo-Martín MM; Etxeandia-Ikobaltzeta I; Kotzeva A; Rigau D; Louro-González A; Barajas L; Díaz del Campo P; Estrada MD; Gracia J; Salcedo-Fernandez F; Haynes RB; Alonso-Coello P; Álvarez Ariza M; Argente Oliver J; Armario García P; Arrieta Antón E; Balaguer Santamaría A; Borque Fernando A; Buñuel Álvarez JC; Cervera Álvarez A; De Las Heras Liñero ME; Dopico L; Eddy Ives LS; Escó Barón R; Ferreira González I; Ferrer F; Gil Sanz MJ; Lago Deibe FI; Martí Canales JC; Martínez Rubio A; Martos Moreno GA; Mejuto Marti T; Morales Ortiz A; Rioja Sanz LA; Rioja Zuazu J; Rodríguez González A; Rodríguez-Arias Palomo JL; Sáenz Cusí A.

Author details

[1]Iberoamerican Cochrane Centre - Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain. [2]Evidence Based Dentistry Unit, Faculty of Dentistry, Universidad de Chile, Santiago, Chile. [3]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada. [4]Department of Epidemiology, Sub Secretariat of Public Health, Ministry of Health, Santiago, Chile. [5]NHS Ayrshire and Arran, Ayr, UK. [6]Fundación Canaria de Investigación y Salud (FUNCIS), Red de Investigación en Servicios de Salud en Enfermedades Crónicas (REDISSEC), Tenerife, Spain. [7]Osteba, Basque Office for Health Technology Assessment, Vitoria, Spain. [8]Agency for Health Quality and Assessment of Catalonia (AQuAS), Barcelona, Spain. [9]CIBER of Epidemiology and Public Health (CIBERESP), Barcelona, Spain. [10]Centro de Saúde de Cambre, Xerencia de Xestión Integrada de A Coruña SERGAS, A Coruña, Spain. [11]Health Technology Assessment Unit (UETS), Subdirección General de Tecnología e Innovación Sanitaria, Consejería de Sanidad, Madrid, Spain. [12]National Clinical Practice Guideline Programme of the NHS, Madrid, Spain. [13]GuíaSalud-Aragon Institute of Health Sciences, Zaragoza, Spain.

Received: 22 May 2015 Accepted: 22 July 2015
Published online: 31 July 2015

References

1. Martínez García L, Sanabria AJ, García Alvarez E, Trujillo-Martín MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ. 2014;186(16):1211–9.
2. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of national institute for health and care excellence clinical guidelines was about 60 months. J Clin Epidemiol. 2014;67(1):52–5.
3. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007;147(4):224–33.
4. Banzi R, Cinquini M, Liberati A, Moschetti I, Pecoraro V, Tagliabue L, et al. Speed of updating online evidence based point of care summaries: prospective cohort analysis. BMJ. 2011;343:d5856.
5. Vernooij RW, Sanabria AJ, Solà I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci. 2014;9:3.
6. Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol. 2014;67(1):33–45.
7. Martínez García L, Arévalo-Rodríguez I, Solà I, Haynes RB, Vandvik PO, Alonso-Coello P, et al. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci. 2012;7:109.
8. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001;286(12):1461–7.
9. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Assessing the need to update prevention guidelines: a comparison of two methods. Int J Qual Health Care. 2004;16(5):399–406.
10. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Surveillance search techniques identified the need to update systematic reviews. J Clin Epidemiol. 2008;61(8):755–62.
11. Hemens BJ, Haynes RB. McMaster Premium Literature Service (PLUS) performed well for identifying new studies for updated cochrane reviews. J Clin Epidemiol. 2012;65(1):62–72. e1.
12. Sagliocca L, De Masi S, Ferrigno L, Mele A, Traversa G. A pragmatic strategy for the review of clinical evidence. J Eval Clin Pract. 2013;19(4):689–96.
13. Shekelle PG, Motala A, Johnsen B, Newberry SJ. Assessment of a method to detect signals for updating systematic reviews. Syst Rev. 2014;3:13.
14. Haynes RB, Cotoi C, Holland J, Walters L, Wilczynski N, Jedraszewski D, et al. Second-order peer review of the medical literature for clinical practitioners. JAMA. 2006;295(15):1801–8.
15. Holland J, Haynes RB, McMaster PLUS, Team Health Information Research Unit. McMaster Premium Literature Service (PLUS): an evidence-based medicine information service delivered on the Web. AMIA Annu Symp Proc. 2005;340:4.
16. Martínez García L, Sanabria AJ, Araya I, Lawson J, Haynes RB, Rigau D, et al. Strategies to assess the validity of recommendations: a study protocol. Implement Sci. 2013;8(1):94.
17. Montori VM, Wilczynski NL, Morgan D, Haynes RB, Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ. 2005;330(7482):68.

18. Guidelines by topic. Edinburgh (Scotland): Scottish Intercollegiate Guidelines Network — Healthcare Improvement Scotland; 2013. Available: www.sign.ac.uk/guidelines/published/index.html (accessed 2014 Oct 1).
19. Working Group on the Management of Major Depression in Adults. Clinical practice guideline on the management of major depression in adults, National plan for the SHN of the MHCA. Axencia de avaliación de tecnoloxías sanitarias de Galicia (avalia-t). Madrid: Clinical Practice Guidelines in the Spanish SHN: avalia-t No 2006/06; 2008.
20. Working Group of the Guideline for the Prevention and Treatment of Childhood and Juvenile Obesity. Iberoamerican Cochrane Centre, coordinator; Clinical Practice Guideline for the Prevention and Treatment of Childhood and Juvenile Obesity; CPGs: Quality Plan for the Spanish National Healthcare System of the Spanish Ministry for Health and Social Policy; Catalan Agency for Health Technology Assessment; 2009. Clinical Practice Guideline: CAHTA no. 2007/25
21. Working group of the Clinical Practice Guideline on Prostate Cancer Treatment. Clinical practice guidelines on prostate cancer treatment, National plan for the NHS of the MSC. Aragon institute of health sciences (I + CS). Madrid: Clinical Practice Guidelines in the NHS I + CS No 2006/02; 2008.
22. Development group of the stroke prevention Guideline. Quality plan for the national health system of the ministry of health and consumer affairs; Catalan agency for health technology assessment and research. Madrid: Clinical Practice Guideline: AATRM Number 2006/15; 2008.
23. Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. Updating clinical practice recommendations: is it worthwhile and when? Int J Technol Assess Health Care. 2012;28(1):29–35.
24. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol. 2009;62(9):944–52.

APPENDIX III

Article: "Methodological systematic review identifies major limitations in prioritization processes for updating."

Journal of Clinical Epidemiology 86 (2017) 11−24

Methodological systematic review identifies major limitations in prioritization processes for updating

Laura Martínez García[a,*], Hector Pardo-Hernandez[a], Cecilia Superchi[a], Ena Niño de Guzman[a], Monica Ballesteros[a], Nora Ibargoyen Roteta[b], Emma McFarlane[c], Margarita Posso[d], Marta Roqué i Figuls[a], Rafael Rotaeche del Campo[e], Andrea Juliana Sanabria[a,f], Anna Selva[a], Ivan Solà[a,g], Robin W.M. Vernooij[a], Pablo Alonso-Coello[a,g]

[a]*Iberoamerican Cochrane Centre, Hospital de la Santa Creu i Sant Pau, C/ Sant Antoni Maria Claret 167, Pavelló 18 - planta 0, Barcelona 08025, Spain*
[b]*Department of Health and Consumer Affairs, Basque Office for Health Technology Assessment (Osteba), Basque Government, Calle Donostia 1, Vitoria-Gasteiz ES-01010, Spain*
[c]*National Institute for Health and Care Excellence, Level 1A City Tower, Piccadilly Plaza, Manchester M1 4BT, UK*
[d]*Service of Clinical Epidemiology and Public Health, Hospital de la Santa Creu i Sant Pau, C/ Sant Antoni Maria Claret 167, Pavelló 18 - planta 0, Barcelona 08025, Spain*
[e]*Alza Health Centre, Osakidetza-Basque Health Service, Avda Larratxo s/n, Donostia-San Sebastián 20013, Spain*
[f]*Clinical Epidemiology and Cancer Screening, Corporació Sanitaria Parc Taulí de Sabadell, Parc Taulí s/n, Sabadell 08208, Spain*
[g]*CIBER de Epidemiología y Salud Pública (CIBERESP), Spain*

Accepted 10 May 2017; Published online 24 May 2017

Abstract

Objectives: The aim of the study was to identify and describe strategies to prioritize the updating of systematic reviews (SRs), health technology assessments (HTAs), or clinical guidelines (CGs).

Study Design and Setting: We conducted an SR of studies describing one or more methods to prioritize SRs, HTAs, or CGs for updating. We searched MEDLINE (PubMed, from 1966 to August 2016) and The Cochrane Methodology Register (The Cochrane Library, Issue 8 2016). We hand searched abstract books, reviewed reference lists, and contacted experts. Two reviewers independently screened the references and extracted data.

Results: We included 14 studies. Six studies were classified as descriptive (6 of 14, 42.9%) and eight as implementation studies (8 of 14, 57.1%). Six studies reported an updating strategy (6 of 14, 42.9%), six a prioritization process (6 of 14, 42.9%), and two a prioritization criterion (2 of 14, 14.2%). Eight studies focused on SRs (8 of 14, 57.1%), six studies focused on CGs (6 of 14, 42.9%), and none were about HTAs. We identified 76 prioritization criteria that can be applied when prioritizing documents for updating. The most frequently cited criteria were as follows: available evidence (19 of 76, 25.0%), clinical relevance (10 of 76; 13.2%), and users' interest (10 of 76; 13.2%).

Conclusion: There is wide variability and suboptimal reporting of the methods used to develop and implement processes to prioritize updating of SRs, HTAs, and CGs.

Keywords: Clinical guidelines; Methodology; Prioritization; Systematic review; Technology assessment; Updating

1. Introduction

Systematic reviews (SRs), health technology assessments (HTAs), and clinical guidelines (CGs) are closely related health decision-making tools that help patients, health care providers, and other stakeholders to make informed decisions. These documents share common steps in their development (i.e., assessment of the available evidence and systematic synthesis) [1−3].

As new evidence can change the conclusions of SRs, HTAs, and CGs, rigorous updating strategies are crucial in the maintenance of these documents [4−6]. We define updating as an iterative process with a systematic and

Funding: This study has been funded by the Instituto de Salud Carlos III through the project "Development and implementation of prioritization tool for updating clinical questions of clinical practice guidelines (PI15/00325)" (co-funded by European Regional Development Fund/European Social Fund, "Investing in your future"). L.M.G. is funded by a Juan Rodés contract from the Instituto de Salud Carlos III (JR15/00004). P.A-.C. is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (MSII15/00034).

* Corresponding author. Iberoamerican Cochrane Centre, Hospital de la Santa Creu i Sant Pau, C/ Sant Antoni Maria Claret 167, Pavelló 18 - planta 0, 08025 Barcelona, Spain. Tel./fax: +34-93-553-78-08.
E-mail address: laura.martinez.garcia@cochrane.es (L. Martínez García).

http://dx.doi.org/10.1016/j.jclinepi.2017.05.008
0895-4356/

What is new?

Key findings

- We identified 14 studies about prioritization process for updating (eight for SRs, six for clinical guidelines [CGs], and none for health technology assessments [HTAs]).
- We identified 76 prioritization criteria that can be applied when prioritizing SRs or CGs for updating.
- There is wide variability and suboptimal reporting of the methods used to develop and to implement prioritization processes in SRs, HTAs, and CGs.

What this adds to what was known?

- We provide an exhaustive description on methods to prioritize SRs, HTAs, and CGs for updating.
- We proposed hierarchical classification of the studies based on three levels: updating strategy, prioritization process, and prioritization criteria.
- We proposed different domains to standardize the prioritization criteria.

What is the implication and what should change now?

- Updating strategies may include a rigorous and transparent prioritization process to efficiently maintain SRs, HTAs, and CGs up to date.
- Further rigorous methodological research is required to optimize the prioritization process for updating.

explicit methodology that includes identification, review, and assessment of new evidence that is not included in the original document [7−9]. If the new evidence is relevant, the document needs to be reviewed and, if necessary, modified. Moreover, the updating strategies provide an opportunity to improve the overall methodology and edition of the document (e.g., correction of mistakes or enhancement to the writing).

Cochrane Handbook suggests updating of SRs every 2 years, although there might be exceptions to this rule (e.g., SRs need more frequent updates if relevant research is being published frequently or alternatively SRs are still current in some topics where new data emerge slowly or are unlikely to emerge) [2]. However, most SRs are updated less frequently than recommended [10−12]. The updating of SRs is resource intensive and time-consuming [10,11]; therefore, different stakeholders are starting to advocate for an approach based on the prioritization of SRs or topics for updating as opposed to predetermined time frames [13,14].

Currently, little is known about the updating of HTAs. Although some HTA developers or initiatives included this aspect in their reports, they do not provide specific guidance on how to implement it [1,15,16].

Several studies have assessed the validity of CGs and their recommendations, specifically the length of time they remain valid [4,5,17−19]. Based on this evidence, most CG developers adopt updating policies based on predetermined time frames [9]. However, the decision to update a CG is a complex process that needs to consider other factors like the volume of new research, available resources, or the balance between updating and developing CGs de novo [20]. In this context, there is a growing interest in approaches that help stakeholders determine which CGs or topics should be prioritized for updating [21]. As an example, Agbassi et al. [22] developed and implemented two questionnaires: one to classify CGs in order of priority for updating and another one to determine the effect of newly available evidence on CGs recommendations.

Methods to prioritize health decision-making tools for updating would ensure that resources are invested to update the documents that are most relevant to different stakeholders. Until now, the prioritization processes for updating reported in literature have not been systematically reviewed. We therefore undertook an SR to identify and describe processes to prioritize updating of SRs, HTAs, and CGs.

2. Methods

2.1. Information sources and search strategy

We searched in MEDLINE (accessed through PubMed, from 1966 onward) and The Cochrane Methodology Register (accessed through The Cochrane Library, Issue 8 2016) in August 2016. We did not establish limitations according to the language or publication status. The search strategy is available in the supplementary data (Supplementary Data 1 at www.jclinepi.com). Additionally, we hand searched the G-I-N Conferences abstract books (2011 onward), HTA International Meetings abstract books (2011 onward), and reference lists of the included studies. We also consulted experts and contacted authors of the included studies.

2.2. Eligibility criteria

Inclusion criteria: Studies that described or implemented one or more strategies to prioritize updating of SRs, HTAs, or CGs.

Exclusion criteria: SRs, HTAs, or CGs methodological handbooks; updated SRs, HTAs, or CGs; letters; comments; or editorials.

2.3. Study selection

Two reviewers (L.M.G. and H.P-.H.) independently screened titles and abstracts to identify potentially eligible

references for inclusion. They obtained the full-text copies of potentially eligible references for further assessment. Disagreements were initially solved by consensus; if necessary, a third reviewer (E.N.d.G.) was consulted. Study authors were contacted via email when additional information was needed.

2.4. Data extraction

We designed a data extraction form that included the following information: (1) study identification, (2) description of the study, (3) description of the methods used to develop the prioritization process for updating, (4) description of the methods used to implement the prioritization process for updating, and (5) strengths and weaknesses of the prioritization process as provided by authors (Supplementary Data 2 at www.jclinepi.com).

We classified the included studies by type of design (descriptive or implementation study), hierarchical level (updating strategy, prioritization process, or prioritization criteria), and type of updated document (SR, HTA, or CG) (Figs. 1 and 2).

Two reviewers (H.P-.H. and C.S.) pilot-tested and refined the data extraction form. Subsequently, two reviewers (H.P-.H. and C.S.) independently extracted data from all the studies that met the inclusion criteria. Disagreements were solved by consensus; if necessary, a third reviewer (L.M.G.) was consulted.

2.5. Data analysis

For the quantitative date, we calculated absolute frequencies and proportions.

For the qualitative data, we used content analysis to summarize and draw conclusions. We reviewed the data in detail and developed initial codes based on topics present in the narrative descriptions and organized sentences and paragraphs into these codes [23]. We constructed a conceptual framework that guided the organization and presentation of results based on the content analysis. We selected

Type of design	
• **Descriptive study**	Study that described one or more strategies to prioritize updating of SRs, HTAs, or CGs.
• **Implementation study**	Study that described but also pilot-tested and/or implemented one or more strategies to prioritize updating of SRs, HTAs, or CGs.
Hierarchical level	
• **Updating strategy**	Step-by-step method used to update documents. The essential stages in an updating strategy are: 1) identifying new evidence; 2) assessing the impact of this evidence on the document, and deciding the need to update the document; and 3) reviewing and, if needed, modifying the document.
• **Prioritization process**	Method used to determine which document should be prioritized for updating. The essential stages in a prioritization process are: 1) assessment of documents using one or more prioritization criteria, and 2) classification of documents in prioritization groups (e.g. high, medium or low relevance for updating). Prioritization process could be included in different stages of the updating strategy
• **Prioritization criteria**	Document-related factors that can be used to classify documents in order of priority for updating
Type of updated document	
• **Systematic review**	"A systematic review attempts to collate all empirical evidence that fits prespecified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made" [1].
• **Health technology assessment**	"Systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision making regarding health technologies" [2].
• **Clinical guideline**	"Clinical practice guidelines are 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" [3].

Fig. 1. Studies classification.

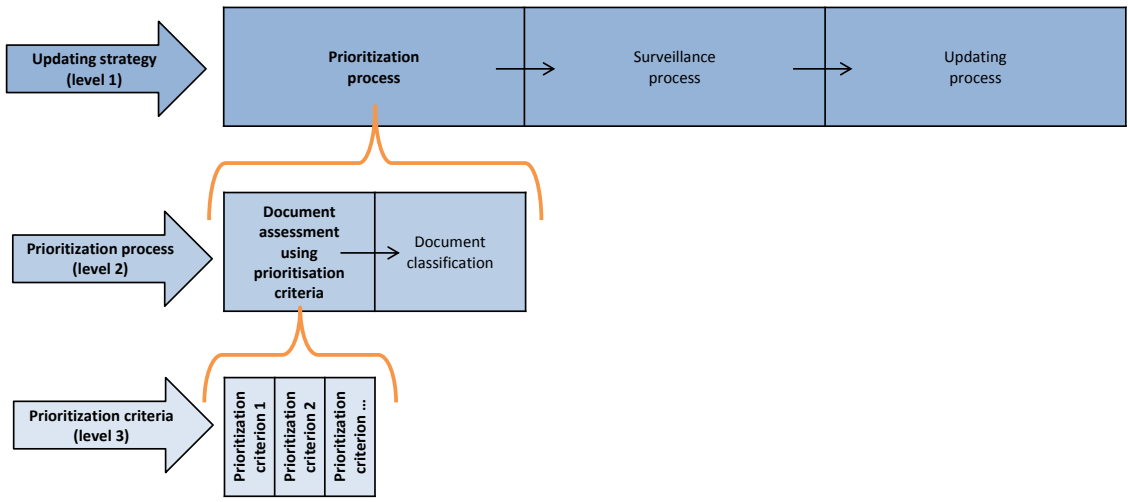


Fig. 2. Hierarchical classification diagram.

by consensus relevant quotes that were illustrative of the strengths and weaknesses of the identified prioritization process for updating.

We summarized narratively and in tables all the data extracted.

We adhered to Preferred Reporting Items for Systematic reviews and Meta-Analyses guideline for the reporting of this SR [24].

3. Results

3.1. Study selection

The screening process is summarized in a flow diagram (Fig. 3). We initially identified 4,724 references and excluded 4,677 references after examining their titles and abstracts. We reviewed 47 full texts and excluded 17 references (Supplementary Data 3 at www.jclinepi.com). We finally included 14 studies corresponding to 30 individual publications (Supplementary Data 4 at www.jclinepi.com) [22,25–53]. Six studies were only available as abstracts and/or in presentation format (6 of 14, 42.9%) [26,28,32,39,50,52]. We successfully contacted the authors of these studies and obtained additional information.

3.2. Study characteristics

The main characteristics of the 14 included studies are reported in Table 1. Most studies were conducted from Canada, the UK, or the United States (11 of 14, 78.6%) [22,26,28,32,35,36,39,42,43,48,53]. Most of the studies were published during the past 5 years (10 of 14, 71.4%) [22,26,28,32,35,36,42,48,52,53].

We classified six studies as descriptive (6 of 14, 42.9%) [28,32,35,43,50,52] and eight as implementation (8 of 14, 57.1%) [22,26,33,36,39,42,48,53]. Six studies reported an updating strategy (6 of 14, 42.9%) [28,33,36,42,43,52], six a prioritization process (6 of 14, 42.9%) [22,26,32,48,50,53], and two a prioritization criterion (2 of 14, 14.2%) [35,39]. Eight studies focused on SRs (8 of 14, 57.1%) [28,33,35,39,42,43,48,53], six studies focused on CGs (6 of 14, 42.9%) [22,26,32,36,50,52], and none were about HTAs. Most of the studies focusing on SRs were conducted by research groups affiliated with Cochrane (5 of 8, 62.5%) [33,39,43,48,53].

Studies that implemented a prioritization process provided the most detail (4 of 14, 28.6%), typically assessing and classifying documents according to the need for updating [22,26,48,53].

3.3. Methods used to develop the prioritization processes for updating

Five studies included a description of the development method for their proposed prioritization process for updating (5 of 14, 35.7%, Table 2) [22,32,36,50,52]. Most development methods followed a common pathway: conducting a literature review [22,32,36,52] and/or consultation with experts (survey or interview) [32,36,50,52].

Lord et al. [36] performed a study to compare the priorities for updating obtained from stakeholder surveys with those elicited from an economic model of diagnostic and treatment pathways. Two NICE CGs were used as case studies (prostate cancer and atrial fibrillation). A discrete event simulation model was used to model the recommended care pathway and estimate consequent costs and outcomes. A health economics research group invested approximately 24 months to develop the models for the two CGs.

3.4. Methods used to implement the prioritization processes for updating

The eight studies, that implemented their prioritization process for updating (8 of 14, 57.1%), usually adopted

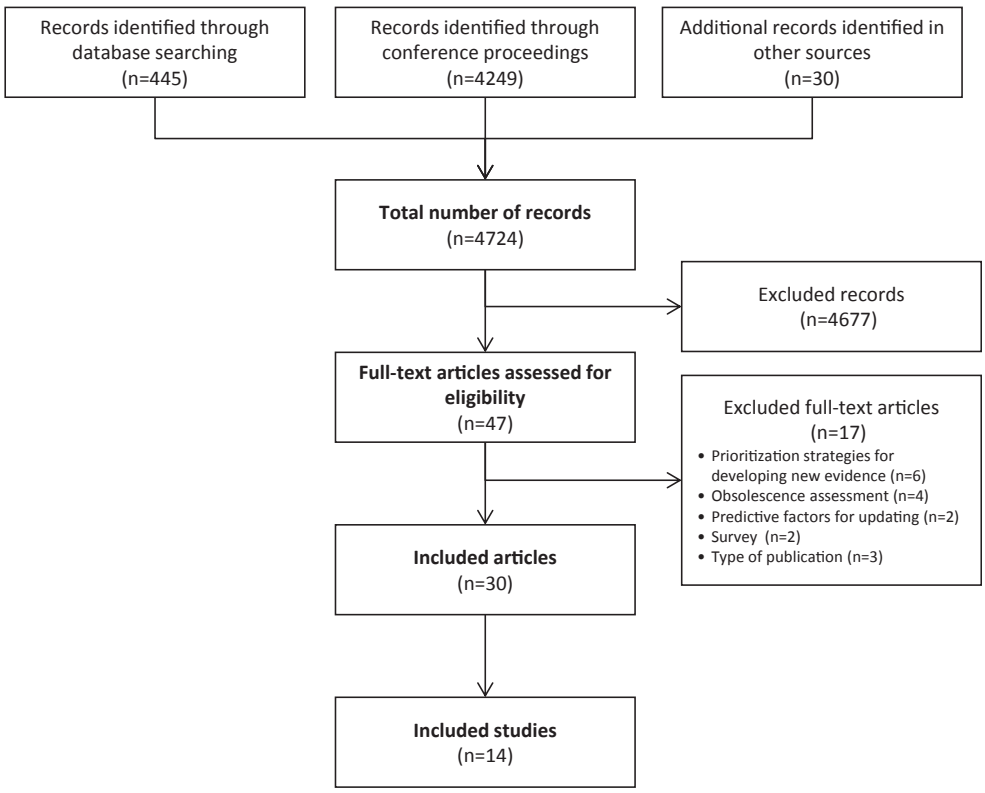


Fig. 3. Flow diagram of the screening process.

one or both of the following approaches: (1) assessment of the candidate documents at a given point in time using a questionnaire [26,33,36] or (2) assessment of candidate documents using a step-by-step algorithm (Table 3) [22,36,42,48,53]. The latter typically involved a pragmatic literature searches to identify new evidence (e.g., a streamlined SR or abbreviated literature searches [22,42]).

Five studies (5 of 8, 62.5%) incorporated categories to classify documents according to their priority for updating (range 1−5), mostly reflecting of the relative need to update (e.g., high, medium, or low) [22,26,42,53]. Experts in the development process (i.e., methodologists), researchers, and/or clinicians were the most common participants in the implementation of these prioritization processes [22,26,33,36,42,53]. A health economics research group collaborated in the study by Lord et al. [36], which involved economic modeling.

Six studies (6 of 8; 75.0%), which reported quantitative results, assessed a total of 660 documents (Table 3) [22,26,36,42,48,53]. Four studies reported high-priority documents for updating (4 of 8; 50%), identifying 11.2% documents in this category (71 of 623; Table 3) [22,26,42,53].

Half of the studies shared similar methodological tools (4 of 8; 50.0%), including questionnaires, algorithms, examples of pathways, or prediction equations (Supplementary Data 5 at www.jclinepi.com) [22,33,36,48].

Two studies reported the time taken to implement the proposed prioritization process for updating (2 of 8; 25%). Amos et al., in which eligible documents (clinical questions from CGs) were assessed using a questionnaire (with three prioritization criteria), required 2 hours for each document [26]. Agbassi et al., in which documents (CGs) were evaluated through a step-by-step process (with five prioritization criteria), required a median of 167 days for each document (range 18−358 days) [22]. Two studies reported the time between assessments (2 of 8; 25%), for an interval of 6 and 12 months, respectively [22,42].

3.5. Prioritization criteria for updating

We identified 76 prioritization criteria that can be applied when prioritizing documents for updating (Table 4); and 17.1% (13 of 76) of criteria were defined [22,35,36,39,42,48,53]. The studies included a mean of 3.5 criteria (range 1−18) [22,26,28,32,33,35,36,39,42,43,48,50,52,53]. We clustered criteria in 11 domains: (1) available evidence; (2) clinical relevance; (3) users' interest; (4) impact on resources use and costs; (5) impact on public health policies; (6) adequacy of the clinical question; (7) time frame (time between document development and assessment for update); (8) cluster criteria (including more than one criterion); (9) impact on stakeholders' views and experience; (10) complex criteria (which requires economics or statistics knowledge); and (11) others.

The most frequently reported criteria were related to the ''available evidence'' domain (19 of 76, 25.0%). Those

Table 1. Characteristics of included studies

Author and year[a]	Institution (country)	Design	Process level	Updated document
Agbassi et al. 2014 [22]	Program in Evidence-based Care, Cancer Care Ontario (Canada)	Implementation study	Prioritization process	Clinical guideline
Amos et al. 2013[b] [26]	Kaiser Permanente (USA)	Implementation study	Prioritization process	Clinical guideline (clinical questions)
Chang 2014[b] [28]	Agency for Health Research and Quality (USA)	Descriptive study	Updating strategy	Systematic review
Jamshidi et al. 2016[b] [32]	Université Laval (Canada)	Descriptive study	Prioritization process	Clinical guideline
Jordan et al.2008 [33]	NZ Cochrane Branch of the Australasian Cochrane Center, Cochrane Menstrual Disorders and Subfertility Group (New Zealand)	Implementation study	Updating strategy	Systematic review
Langan et al.2012 [35]	University of Leeds (UK)	Descriptive study	Prioritization criteria	Systematic review
Lord et al. 2013 [36]	National Institute for Health Research (UK)	Implementation study	Updating strategy	Clinical guideline
Salzwedel and Wright 2010[b] [39]	Cochrane Hypertension Review Group (Canada)	Implementation study	Prioritization criteria	Systematic review
Shekelle et al. 2014 [42]	Agency for Healthcare Research and Quality (USA)	Implementation study	Updating strategy	Systematic review
Soll 2008 [43]	Cochrane Neonatal Review Group (CNRG) (USA)	Descriptive study	Updating strategy	Systematic review
Takwoingi et al. 2013 [48]	Cochrane Editorial Unit (UK)	Implementation study	Prioritization process	Systematic review
Theobald et al. 1999[b] [50]	French Federation of Comprehensive Cancer Centers (France)	Descriptive study	Prioritization process	Clinical guideline
Venhorst et al. 2014[b] [52]	Knowledge Institute of Medical Specialists (the Netherlands)	Descriptive study	Updating strategy	Clinical guideline
Welsh et al. 2015 [53]	Cochrane Airways group (UK)	Implementation study	Prioritization process	Systematic review

[a] In alphabetical order.
[b] Only published abstract and/or presentation available.

Table 2. Methods used to develop the prioritization process

Author and year[a]	Description of the process	Participants
Agbassi et al. 2014 [22]	• Literature review.	• Research team.
Jamshidi et al. 2016[b] [32]	• Systematic review. • Survey. • Calculation of the criteria weights (analytical hierarchy process).	Survey • CGs development institutions. • Experts on the field.
Lord et al. 2013 [36]	Survey • Not reported Economic modeling • Individual level discrete event simulation.	Survey • Research team Economic modeling • Health economics research group
Shekelle et al. 2014 [42]	• Not reported.	• The RAND Corporation, Southern California Evidence-based Practice Center. • University of Ottawa Evidence-based Practice Center.
Soll 2008 [43]	• Not reported.	• Members of the Cochrane Neonatal Review Group.
Theobald et al. 1999[b] [50]	• Survey.	• Clinicians.
Vernhorst 2014[b] [52]	• Literature review • Interviews.	Interviews • CGs development group.

Abbreviation: CG, Clinical guideline.
[a] In alphabetical order.
[b] only published abstract and/or presentation available.

criteria were included in nine studies (9 of 14, 64.3%) [22,26,32,33,43,48,50,52,53], and six studies included more than one criterion related to this domain (6 of 14; 42.9%) [26,32,33,43,50,53]. Jamshidi et al. [32], the only study that ranked the criteria, reported higher weights for available evidence criteria: ''changes in the evidence on the benefits and harms of existing interventions'' (weight: 0.144), ''strength of the evidence'' (weight: 0.132), and ''quality of the evidence'' (weight: 0.131). Additionally, a criterion defined as ''new evidence'' was frequently used (5 of 76; 6.6%) [22,43,48,52,53], often requiring a pragmatic literature search to identify new evidence [22,48,53].

''Clinical relevance'' was rather an unspecific domain usually related with burden of disease or relevance of the document (i.e., topic, clinical question, or CG) (10 of 76; 13.2%) [26,33,48,50,52,53].

Criteria in the domain ''users' interest'' and related to ''citations, downloads, or Web site hits'' clustered similar items: ''frequency of downloads'' [28], ''citation in other scientific literature, including clinical practice guidelines'' [28], ''number of citations for SR'' [39], ''reviews that were most highly cited'' [53], or ''reviews that had the most Web site hits'' [53].

Three studies developed more complex criteria (3 of 76; 3.9%) [35,36,48]. Langan et al. [35] 2012 illustrated the potential impact of one new study on a given meta-analysis using funnel plots, which could be informative when prioritizing SRs for updating. Takwoingi et al. [48] developed a prediction equation for estimating the probability of conclusions changing after the addition of new studies to a meta-analysis. The prediction equation uses two signals: (1) ''ratio of the total weight of the new studies to the total weight of the old studies in an updated meta-analysis (weight ratio)'' and (2) number of new studies [48]. Lord et al. [36] modeled service pathways of two CGs to estimate the cost-effectiveness of possible changes to the pathways. The economic prioritization criterion was determined based on: (1) the probability that the currently recommended option is not the optimal strategy using a cost-effectiveness threshold of £20,000 per QALY gained and (2) the magnitude of the potential gain in net benefits (difference between the optimal and base-case strategies). These statistics were used by the authors to rank topics in order of importance for inclusion in an update of the two CGs.

3.6. Strengths and weaknesses of the prioritization processes for updating

Relevant quotes that illustrated the strengths and weaknesses reported by study authors are provided in Fig. 4. The strengths of the proposed strategies reflected the potential impact of the prioritization process for updating on the efficiency of the updating process (rigorous process that saves both time and resources) [22,26,32,35,43,48,52,53]. Conversely, the following limitations were highlighted: external validity, the resources required, or the need to conduct surveillance of the new evidence [22,48,52,53].

4. Discussion

4.1. Main findings

We identified 14 studies (reported in 30 publications) about the prioritization process for updating health decision-making tools (eight for SRs, six for CGs, and none for HTAs). The studies included either of the following approaches to prioritize documents for updating: a more pragmatic assessment (using a questionnaire that included different prioritization criteria) or more formal assessment (using a step-by-step algorithm that included literature searches to identify new evidence). We identified 76 prioritization criteria that can be applied when prioritizing documents for updating. The most frequently cited criteria were as follows: available evidence, clinical relevance, and users' interest.

There is wide variability of methods used to develop and to implement prioritization processes in methodological studies. Furthermore, the applied methods are poorly reported (i.e., few studies described the development method or the prioritization criteria).

4.2. Our results in the context of previous research

4.2.1. Aspects related with updated documents

4.2.1.1. Strategies to prioritize updating of SRs. In the field of SRs, the most recent consensus document on when and how to update SRs reported examples of how different organizations decide which SRs within their portfolio should be updated (i.e., AHRQ and Cochrane) [13]. However, and consistent with our findings, authors found that the studies that propose an overall updating strategies of SRs did not provide a detailed description of the prioritization criteria used [33,42,43].

4.2.1.2. Strategies to prioritize updating of HTAs. We did not identify any prioritization processes for updating HTAs. However, we did find studies that aimed to prioritize the evaluation of newly introduced health technologies or to identify potentially obsolete technologies, or those that may need additional evidence generation [1,54–58]. As such, Ruano Raviña et al. [56] developed a prioritization tool (PriTec) to identify, prioritize, and assess of potentially obsolete health technologies.

4.2.1.3. Strategies to prioritize updating of CGs. Previous SRs on CGs updating strategies did not identify or only partially described processes that could inform the decision of which CGs should be prioritized for updating [7,9,59]. There are, nevertheless, new studies published ever since that underscore the relevance of the prioritization process in CG updating [22,26,32,36,52], coinciding with a growing interest to shift from developing to updating CGs among developers [60].

4.2.2. Aspects related with prioritization process

4.2.2.1. When to prioritize documents for updating. Different prioritization processes could be implemented in

Table 3. Methods used to implement the prioritization process

Author and year[a]	Description of the process	Priority classification
Agbassi et al. 2014 [22]	Step 1. Document assessment: 1. Application of DAQ. 2. Classification of the documents (endorsed, deferred, reviewed, archived, and special cases). 3. Prioritization of the documents in review category (urgent, high, medium, or low priority). Step 2. Document review: 1. Application of DRQ. 2. Conduction of a streamlined SR (without conducting a full quality assessment). 3. Classification of the documents (endorsed, updated, and archived). 4. Consensus by larger expert panel.	• Urgent • High • Medium • Low
Amos et al. 2013[b] [26]	Ranking of clinical questions by importance for literature monitoring and clinical importance (9-point Likert scale) (1).	• High (rank 7–9) • Medium (rank 4–6) • Low (rank 1–3)
Jordan et al. 2008 [33]	1. Nomination of out-of-date reviews by Cochrane Review Groups. 2. Application of the Cochrane Infectious Diseases Group new priority setting approach.	Without classification.
Lord et al. 2013 [36]	Survey Survey to stakeholders. Economic modeling Application of a ''service pathway model' and a ''disease process model.''	Survey • Very important • Important • No opinion • Somewhat important • Not important Economic modeling Without classification.
Salzwedel and Wright 2010[b] [39]	Use of Google Scholar to determine the number of times an SR has been cited.	Without classification.
Shekelle et al. 2014 [42]	1. Conduction of an abbreviated literature search. 2. Obtaining expert opinion. 3. Obtaining safety alerts from different sources. 4. Assessment of the need to update each conclusion for each key question. 5. Assessment of priority status to updating the full report.	• High • Medium • Low
Takwoingi et al. 2013 [48]	1. Classification of the clinical question (''current question, no longer being updated'' or ''historical question, no longer being updated'') based on the new evidence. 2. Identification of new factors from the existing included studies, new methodology, response to feedback from users of the review, inclusion in policy decision-making, or clinical practice guidelines. 3. Application of the prediction tool if new studies are identified as relevant.	• Not reported
Welsh et al. 2015 [53]	1. Understanding of patient uncertainties about asthma. 2. Piloting of a prioritization tool to assess whether individual reviews require updating [Takwoingi2013]. 3. Survey of the Cochrane Airways Group Editorial Board. 4. Prioritization of new review titles by horizon scanning.	• High

Abbreviations: AF, atrial fibrillation; DAQ, document assessment questionnaire; DRQ, document review questionnaire; NB, net benefit; PC, prostate cancer; SR, systematic review.

Notes: (1) Question 1: Rate the importance of including the topic in an update of the guideline; (2) The priority strategy was not implemented due to the small response of Cochrane Review Groups; (3) ''The estimated economic priorities for update topics differed from those elicited from stakeholders''; (4) The total number of reviews assessed is not reported.

[a] In alphabetical order.

[b] Only published abstract and/or presentation available.

different time points of the updating strategy. For example, in CGs context, a prioritization process could be implemented to identify the CGs in greatest need of update; to identify the CGs in greatest need of update after a surveillance process; or to identify the clinical questions of a prioritized CG in greatest need of update. None of the included studies provided this degree of granularity.

4.2.2.2. Pragmatic vs. formal approaches to prioritize updating of documents. Among the studies that reported the

Participants	Piloted	Results (*n* prioritizations/ *n* assessments)	Implementation tools available
• Clinical leaders. • Experts in the development process.	Not reported.	2011 Assessment: 19/109 (17%) urgent; 16/109 (15%) high; 6/109 (6%) medium; 21/109 (19%) low. 2012 Assessment: 8/88 (9%) urgent; 2/88 (2%) high; 10/88 (11%) medium; 18/88 (20%) low.	Yes (two questionnaires, algorithm).
• Clinical experts. • Experts in the development process.	Not reported.	16/127 (13%) high; 22/127 (17%) medium; 42/127 (33%) low.	No.
• Cochrane Review Groups. • Updating officer.	Not reported.	Not reported (2).	Yes (nonformal questionnaire).
Survey • Research team • Stakeholders Economic modeling • Health economics research group	Survey Yes Economic modeling Not reported.	Survey PC 2/8 (25%) very important; AF 2/9 (22%) very important (3) Economic modeling PC 3/6 (50%) high potential for increased NB; AF 2/5 (40%) high potential for increased NB (3)	Survey No Economic modeling Yes (service pathway model).
Not reported.	Not reported.	1/17 (5.9%) highly cited.	Not reported.
• Research team. • Experts.	Yes.	Ahmadzai 201 2/24 (8%) high; 5/24 (21%) medium; 17/24 (71%) low. Newberry 2013 2/14 (14%) high; 3/14 (21%) medium; 9/14 (64%) low.	Yes (nonformal questionnaire)
• Review or editorial team.	Yes.	1/n high; 1/n low (4)	Yes (prediction equation, algorithm).
• Editorial Board (clinicians and researchers). • Information specialist.	Yes.	30/270 (11%) high.	No.

implementation of a prioritization process for updating, we identified pragmatic as well as more formal approaches. The pragmatic approaches (based mostly on surveys among stakeholders) could be considered less resource intensive and time-consuming compared with formal approaches (based on step-by-step algorithms that frequently including literature searches) [22,26,33,36,42,48,53]. Lord et al. [36] pointed out that there were substantial differences between the results of both approaches. The topics prioritized for updating obtained by surveying stakeholder largely differed from those obtained by modeling formal economic pathways [36]. Consequently, upon implementing a prioritization

Table 4. Prioritization criteria for updating

Domain	Priority criteria	No of studies (%)
1. Available evidence		19 (25.0)
	"New evidence" [22,43,48,52,53].	5 (6.6)
	"Fast-changing field" [32,33,53].	3 (3.9)
	"Differences between evidence and current practice" [32,50].	3 (3.9)
	Others	8 (10.5)
	• "Changes in the evidence on the benefits and harms of existing interventions" [32].	
	• "Continued uncertainty regarding the question at hand" [43].	
	• "Controversy about interpretation of the evidence" [50].	
	• "Existence of high-quality systematic reviews" [26].	
	• "Knowledge that current evidence was relatively unchanged" [26].	
	• "Number of trials and participants identified by new search in relation to the number of participants already included" [33].	
	• "The quality of the evidence" [32].	
	• "The strength of the evidence" [32].	
2. Clinical relevance		10 (13.2)
	"Burden of disease" [33,50,52].	3 (3.9)
	"Relevance, or not, of clinical question" [26,48,53].	3 (3.9)
	Others	4 (5.3)
	• "Frequency of disease" [50].	
	• "Relative importance of including each topic in a potential future update of the guideline" [36].	
	• "Relevance" [22].	
	• "The current relevance of the CG" [32].	
3. Users' interest		10 (13.2)
	"Citations/downloads/website hits" [28,39,53].	5 (6.6)
	Others	5 (6.6)
	• "Equipoise where there is current debate" [33].	
	• "Interest from stakeholder partners" [28].	
	• "Interest of the health professionals" [50].	
	• "Interest of the patient groups" [50].	
	• "Performance evaluations and feedback on guideline use" [32].	
4. Impact on resources use and costs		7 (9.2)
	"Impact on access to care" [22,32].	2 (2.6)
	Others:	5 (6.6)
	• "Changes in the resources available for health care" [32].	
	• "Economic dimension" [50].	
	• "Financial aspects" [52].	
	• "Interest of the payers" [50].	
	• "The required resources" [32].	
5. Impact on public health policies		5 (6.6)
	• "Impact in terms of health policy" [50].	
	• "Impact in terms of prevention" [50].	
	• "Policy based on the current results" [33].	
	• "Potential benefits of updating a guideline for public health" [32].	
	• "Strategic priority" [33].	
6. Adequacy of the clinical question		4 (5.3)
	• "Changes in the available interventions" [32].	
	• "Feasible" [22].	
	• "New clinical contexts that needed to be addressed" [43].	
	• "The appropriateness of the questions and search criteria" [32].	
7. Time frame criteria		4 (5.3)
	• "Age of module" [52].	
	• "Age of review" [33].	
	• "Deferred" [22].	
	• "The last review date of CG" [32].	
8. Cluster criteria		4 (5.3)
	• "Magnitude of out-of-date conclusions (e.g., consideration of magnitude/direction of changes in estimates, potential changes in practice or therapy preference, safety issue including withdrawn from the market drugs/black box warning, availability of a new treatment)" [42].	
	• "New relevant factors (information from existing included studies, new methodology, response to feedback from users of the review, and inclusion in policy decision or clinical practice guidelines)" [48,53].	
	• "Number of conclusions of the CER are up to date, possibly out of date, or certainly out of date" [42].	

(*Continued*)

Table 4. Continued

Domain	Priority criteria	No of studies (%)
9. Impact on patients' values and preferences		3 (3.9)
	• "Changes in outcomes that are considered important" [32].	
	• "Changes in the values placed on outcomes" [32].	
	• "Patient uncertainties" [53].	
10. Complex criteria		3 (3.9)
	• "Likelihood of a change in statistical significance using funnel plots" [35].	
	• "Prediction equation for estimating the probability of conclusions changing after the addition of new studies to a meta-analysis" [48].	
	• "Modeled service pathways" [36].	
11. Others criteria		7 (9.2)
	• "Ethical considerations" [50].	
	• "Impact on quality of care" [52].	
	• "Interventions relevant to developing countries or to Millennium Development Goals" [33].	
	• "Legal dimension" [50].	
	• "Size of patient population" [52].	
	• "The risk of leaving the outdated document publicly available" [32].	
	• "The scope of guideline" [32].	
Total		76 (100)

Abbreviation: CG, clinical guideline.

process for updating, developers should consider the balance between the advantages (i.e., less resource intensive and time-consuming) and disadvantages (i.e., literature searches or validation process) of more pragmatic approaches.

4.2.2.3. Prioritization processes in more dynamic SRs and CGs. The definition and methods to develop and implement dynamic or living SRs or CGs are still unclear [61–64]; furthermore, this approach is resource intensive and time-consuming [65]. A recent experience in continuous surveillance and updating of a pregnancy CG was stopped early because of lack of financial resources for maintaining the surveillance strategy [66]. Prioritization process in dynamic SRs and CGs is of even more importance, given the need to allocate the limited resources in those topics that required an update [21,67].

4.2.2.4. Involvement of stakeholders. There is a global movement to increase the involvement of different stakeholders (consumers [public, patient, and caregiver]; clinicians; policy makers; or researchers) in research development and in determining research priorities [68–70]. This engagement net benefit still needs to be further investigated in updating prioritization processes, as very few studies considered this aspect [28,50].

Strengths

- *"Institutes will have a formal and rigorous process for deciding when a guideline should be updated"* [1].
- *"Time savings realized by focusing only on those questions that are clinically relevant"* [2].
- *"The decision tool can promote channelling limited resources into updating systematic reviews that are most sensitive to change"* [3].

Limitations

- *"Test the universality of the procedures to determine steps that are generalizable and potential efficiencies that can be made in the process "* [4].
- *"It still takes a long time (median 167 days) to complete a Document Assessment and Review process".* [4].
- *"The use of the tool depends on monitoring the literature for new evidence"* [3].

Fig. 4. Examples of strengths and weaknesses.

4.2.3. Aspects related with methodological research

4.2.3.1. Other information sources to identify prioritization criteria for updating. During the review process, we identified studies that could be informative regarding the prioritization criteria for updating. There are, for instance, studies on processes for prioritizing topics for SRs, HTAs, and CGs development [71–75]. Some of the criteria for prioritizing documents for development and update may overlap; consequently, developers should consider that the criteria adopted for prioritizing development and updating are consistent.

A second example is the work conducted on identifying of potential factors associated with updating documents [4,6,18,19,76,77]. The identified predictive factors are consistent with our findings, especially regarding the importance of "available evidence" criteria (i.e., updated recommendations were associated with high volume of new publications or with being based on limited evidence) [4,19,77].

4.2.3.2. Reporting items for methodological research. To increase clarity and understanding, future methodological research should standardize the reporting of the methods used to develop and to implement the proposed prioritization processes. The data extraction form used (Supplementary Data 2 at www.jclinepi.com) and the outcomes assessed (time to implement and results of implementation) in this study may prove helpful as a tentative

guidance for reporting prioritization processes for updating. However, providing guidance in this area goes beyond the present study [78,79].

4.3. Strengths and limitations

Our SR has several strengths. We developed an SR protocol (available upon request from the authors) and performed an exhaustive SR. We included the most popular health decision-making tools (SRs, HTAs, and CGs) that could potentially share common steps in updating strategies. We proposed a hierarchical classification of the studies based on three levels: updating strategy, prioritization process, and prioritization criteria. Finally, we also proposed different domains to standardize the terminology for prioritization criteria.

Our SR is subject to some limitations. It is possible that we did not identify all potentially eligible studies due to publication bias or to omission of some more specialized information sources (e.g., methodological handbooks). Furthermore, synthesizing and comparing complex methodological studies without standardized reporting guidance made the analysis and interpretation of results challenging. Finally, we identified the most frequently cited prioritization criteria, although this did not imply their potential relevance for prioritization processes.

5. Conclusions

Updating strategies may include a rigorous and transparent prioritization process to efficiently ensure SRs, HTAs, and CGs remain up to date. However, the wide variability in the methods used to develop and implement prioritization processes and the suboptimal reporting of the methodological studies makes the choice of optimal prioritization criteria a challenge. Further rigorous methodological research is required to optimize the prioritization process for updating.

Acknowledgments

The authors would like to thank María Victoria Leo for her help editing the final version of the manuscript.

Supplementary Data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2017.05.008.

References

[1] Sampietro-Colom L, Lach K, Cicchetti A, Kidholm K, Pasternack I, Fure B, et al. The AdHopHTA handbook: a handbook of hospital-based Health Technology Assessment (HB-HTA); public deliverable; The AdHopHTA Project (FP7/2007-13 grant agreement nr 305018) 2015.

[2] Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. Available at. www.handbook.cochrane.org. Accessed November 4, 2016.

[3] National Institute for Clinical Excellence. Developing NICE guidelines: the manual: National Institute for Clinical Excellence 2015.

[4] Martínez García L, Sanabria AJ, Garcia Alvarez E, Trujillo-Martin MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ 2014;186(16):1211−9.

[5] Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001;286:1461−7.

[6] Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 2007;147:224−33.

[7] Martínez García L, Arevalo-Rodriguez I, Sola I, Haynes RB, Vandvik PO, Alonso-Coello P. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci 2012;7:109.

[8] Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? Lancet 2006;367:881−3.

[9] Vernooij RW, Sanabria AJ, Sola I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci 2014;9:3.

[10] French SD, McDonald S, McKenzie JE, Green SE. Investing in updating: how do conclusions change when Cochrane systematic reviews are updated? BMC Med Res Methodol 2005;5:33.

[11] Jaidee W, Moher D, Laopaiboon M. Time to update and quantitative changes in the results of cochrane pregnancy and childbirth reviews. PLoS One 2010;5(7):e11553.

[12] Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and reporting characteristics of systematic reviews of Biomedical research: a Cross-Sectional study. PLoS Med 2016;13(5):e1002028.

[13] Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ 2016;354:i3507.

[14] Nasser M, Welch V, Tugwell P, Ueffing E, Doyle J, Waters E. Ensuring relevance for Cochrane reviews: evaluating processes and methods for prioritizing topics for Cochrane reviews. J Clin Epidemiol 2013;66:474−82.

[15] Guidelines for the economic evaluation of health technologies: Canada Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

[16] Velasco Garrido M, Kristensen FB, Nielsen CP, Busse R, editors. Health Technology Assessment and Health PolicyMaking in Europe − Current Status, Challenges and Potential. Copenhagen: World Health Organisation; 2008. Available at. http://www.euro.who.int/Document/E91922.pdf.

[17] Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. J Clin Epidemiol 2014;67:52−5.

[18] Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. Updating clinical practice recommendations: is it worthwhile and when? Int J Technol Assess Health Care 2012;28:29−35.

[19] Neuman MD, Goldstein JN, Cirullo MA, Schwartz JS. Durability of class I American College of Cardiology/American Heart association clinical practice guideline recommendations. JAMA 2014;311:2092−100.

[20] Alonso-Coello P, Martínez García L. Clinical guidelines: old and new challenges. Med Clin (barc) 2014;143(7):306−8.

[21] Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic

development of a comprehensive checklist for a successful guideline enterprise. CMAJ 2014;186(3):E123−42.
[22] Agbassi C, Messersmith H, McNair S, Brouwers M. Priority-based initiative for updating existing evidence-based clinical practice guidelines: the results of two iterations. J Clin Epidemiol 2014; 67:1335−42.
[23] Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008; 8:45.
[24] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097.
[25] Ahmadzai N, Newberry SJ, Maglione MA, Tsertsvadze A, Ansari MT, Hempel S, et al. A surveillance system to assess the need for updating systematic reviews. Syst Rev 2013;2:104.
[26] Amos Q, Chan W, Tom G. Maximising efficiency in updating guidelines through prioritisation of clinical questions [abstract]. 10th G-I-N Conference; San Francisco, California, USA; 2013.
[27] Brown J, Lethaby A, Clarke J, Jordan V, Showell M, Farquhar C. The cochrane updating project. In: Evidence in the Era of Globalisation. Abstracts of the 16th Cochrane Colloquium; 2008 3−7 Oct; Freiburg, Germany. Available at http://abstracts.cochrane.org/2008-freiburg/cochrane-updating-project.
[28] Chang S. When to update, when to Archive? The AHRQ EPC Program. Updating Cochrane Reviews Workshop; Hamilton, Canada; 2014.
[29] Crossan C, Westby M, Lord J. Modelling of clinical pathways to assess cost-effectiveness in NICE guidelines: impact on stakeholder view of the importance of potential update topics [abstract]. 9th G-I-N Conference; Berlin, Germany; 2012.
[30] Hopewell S, Loudon K, Clarke M, Moher D, Scholten R, Eisinga A, et al. A decision tool for updating Cochrane reviews. Cochrane Editorial Unit; 2007.
[31] Hopewell S, Loudon K, Clarke M, Moher D, French S, Scholten R, et al. Development of a decision tool for updating Cochrane reviews. In: Evidence in the era of globalisation. Abstracts of the 16th Cochrane Colloquium; 2008 3−7 Oct; Freiburg, Germany.
[32] Jamshidi A, Lamontagne M, Ait-kadi D, Ruiz A, Routhier F. Developing a comprehensive priority-based framework for updating clinical practice guidelines (CPGs): a systematic review and international survey. Huitième Colloque Étudiant du CIRRIS; Québec, Canada; 2016.
[33] Jordan V, Farquhar C, Brown A, Clarke J, Lethaby A, Jeffery M. Centralised updating officer pilot project a project of the Cochrane Collaboration final report: August 2008. NZ Cochrane Branch of the Australasian Cochrane Centre & Cochrane, Menstrual Disorders and Subfertility Group; 2008.
[34] Langan D, Higgins J, Gregory W, Crowther M, Sutton A. Graphical augmentations to the funnel plot for assessing the impact of additional evidence on a meta-analysis. In: Abstracts of the 19th Cochrane Colloquium; 2011 19−22 Oct; Madrid, Spain. John Wiley & Sons; 2011. Available at http://abstracts.cochrane.org/2011-madrid/graphical-augmentations-funnel-plot-assessing-impact-additional-evidence-meta-analysis.
[35] Langan D, Higgins JP, Gregory W, Sutton AJ. Graphical augmentations to the funnel plot assess the impact of additional evidence on a meta-analysis. J Clin Epidemiol 2012;65:511−9.
[36] Lord J, Willis S, Eatock J, Tappenden P, Trapero-Bertran M, Miners A, et al. Economic modelling of diagnostic and treatment pathways in National Institute for Health and Care excellence clinical guidelines: the modelling algorithm pathways in guidelines (MAPGuide) project. Health Technol Assess 2013;17:v−vi. 1-192.
[37] Newberry SJ, Ahmadzai N, Motala A, Tsertsvadze A, Maglione M, Ansari MT, et al. Surveillance and identification of signals for updating systematic reviews: implementation and early experience. Rockville, MD: Agency for Healthcare Research and Quality US; 2013.
[38] Cancer Care Ontario [Homepage]. Document assessment and review protocol version 1.2. Canada: Program in Evidence-Based Care; 2015. Available at https://www.cancercare.on.ca/cms/one.aspx?objectId=122178&contextId=1377.
[39] Salzwedel D, Wright J. Citation rates of systematic reviews of interventions for hypertension and blood pressure. In: Bero L, Montgomery P, Robinson K, Pigott T, Krause K, editors. Bringing Evidence-Based Decision-Making to New Heights. Abstracts of the 2010 Joint Colloquium of the Cochrane and Campbell Collaborations; 2010 18−22 Oct; Keystone, USA. John Wiley & Sons; 2010. Available at http://abstracts.cochrane.org/2010-keystone/citation-rates-systematic-reviews-interventions-hypertension-and-blood-pressure.
[40] Shekelle PG, Newberry S, Maglione M, Shanman R, Johnsen B, Carter J, et al. Assessment of the need to update comparative effectiveness reviews: Report of an initial rapid program assessment (2005-2009). Rockville, MD: Agency for Healthcare Research and Quality US; 2009.
[41] Shekelle PG, Motala A, Johnsen B. Assessment of a method to detect signals for updating systematic reviews. Rockville MD: Agency for Healthcare Research and Quality US; 2014.
[42] Shekelle PG, Motala A, Johnsen B, Newberry SJ. Assessment of a method to detect signals for updating systematic reviews. Syst Rev 2014;3:13.
[43] Soll RF. Updating reviews: the experience of the Cochrane Neonatal Review Group. Paediatr Perinat Epidemiol 2008;22:29−32.
[44] Sutton A, Gamble C, Donegan S, Garner P. Evaluating two statistical methods to predict when a meta-analysis will gain sufficient power to change the conclusions of a Cochrane review. In: Come to the craic. Abstracts of the 14th Cochrane Colloquium; 2006 23-26 Oct; Dublin, UK.
[45] Sutton AJ, Donegan S, Takwoingi Y, Garner P, Gamble C, Donald A. An encouraging assessment of methods to inform priorities for updating systematic reviews. J Clin Epidemiol 2009;62: 241−51.
[46] Takwoingi Y, Sutton A, Donegan S, Garner P, Gamble C, Donald A. Prioritising the updating of systematic reviews. In: Evidence in the Era of Globalisation. Abstracts of the 16th Cochrane Colloquium; 2008 3−7 Oct; Freiburg, Germany.
[47] Takwoingi Y, Hopewell S, Sutton A, Marshall R, Tovey D. Predicting the need to update a systematic review. In: Abstracts of the 19th Cochrane Colloquium; 2011 19−22 Oct; Madrid, Spain. John Wiley & Sons; 2011.
[48] Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. BMJ 2013;347:f7191.
[49] Takwoingi Y, Marshall R, Hopewell S, Tovey D, Sutton AJ. The updating decision tool: guidance for cochrane review groups and authors. Cochrane Editorial Unit; 2014.
[50] Theobald S, Blanc-Vincent MP, Farsi F, Gory-Delabaere G, Bataillard A, Guillo S, et al. The identification of questions in the updating process of clinical practice guidelines for oncology [abstract]. 15th Annual Meeting of the International Society of Technology Assessment in Health Care; Edinburgh, UK; 1999.
[51] Tovey D, Marshall R, Bazian L, Hopewell S, Rader T. Fit for purpose: centralised updating support for high-priority Cochrane Reviews. Cochrane Editorial Unit; 2011.
[52] Venhorst K, Ploegmakers M, Holtus A, Van Barnelveld T. What is needed to realize a dynamic process of guideline updating [abstract]. 11th G-I-N Conference; Melbourne, Australia; 2014.
[53] Welsh E, Stovold E, Karner C, Cates C. Cochrane Airways Group reviews were prioritised for updating using a pragmatic approach. J Clin Epidemiol 2015;68:341−6.
[54] Maceira Rozas MC, Varela Lema L, Atienza Merino G. Tecnologías sanitarias nuevas y emergentes 2015. Lista priorizada. Red Española de Agencias de Evaluación de Tecnologías y Prestaciones del SNS. Axencia Galega para a Xestión do Coñecemento en Saúde, Unidad

de Asesoramiento Científico-técnico, avalia-t; 2016. Informes de evaluación de tecnologías sanitarias.

[55] Puñal-Riobóo J, Baños Álvarez E, Varela Lema L, Castillo Muñoz MA, Atienza Merino G, Ubago Pérez R, et al. en representación del Grupo de trabajo de la Guía para la elaboración y adaptación de informes rápidos de evaluación de tecnologías sanitarias. Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del SNS. Agencia Gallega para la Gestión del Conocimiento en Salud. Unidad de Asesoramiento Científico-técnico, avalia-t; 2016.

[56] Ruano Raviña A, Velasco González M, Varela Lema L, Cerdá Mota T, Ibargoyen Roteta N, Gutiérrez Ibarluzea I, et al. Identificación, priorización y evaluación de tecnologías obsoletas. Guía metodológica. Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad y Política Social. Axencia de Avaliación de Tecnoloxías Sanitarias de Galicia; 2007. Informes de Evaluación de Tecnologías Sanitarias: avalia-t N$^{\circ}$. 2007/01.

[57] Varela-Lema L, Atienza Merino G, López-García M. Selección de prioridades sanitarias: criterios, proceso y marco estratégico. Revisión sistemática. Red Española de Agencias de Evaluación de Tecnologías y Prestaciones del SNS. Axencia Galega para a Xestión do Coñecemento en Saúde. Unidade de Asesoramento Científico-técnico, avalia-t. Informes de evaluación de tecnologías sanitarias; 2016.

[58] European Network For Health Technology Assessment (EUnetHTA). Description of criteria to select and prioritize health technologies for additional evidence generation. Work Package 7 Lead Partner: HAS, French National Authority for Health, 2012.

[59] Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol 2014;67:33−45.

[60] Baker M. Maintaining the currency of a Large guidelines portfolio; Necessity is the Mother of Invention. In Abstracts directory of the 13th G-I-N Conference; 2016 27−30 September; Philadelphia, USA; 2016. Available at http://www.g-i-n.net/document-store/g-i-n-conferences/philadelphia-2016/abstract-directory-2016.

[61] Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JP, Mavergames C, et al. Living systematic reviews: an emerging opportunity to narrow the evidence practice gap. PLoS Med 2014;11(2):e1001603.

[62] Johnston ME, Brouwers MC, Browman GP. Keeping cancer guidelines current: results of a comprehensive prospective literature monitoring strategy for twenty clinical practice guidelines. Int J Technol Assess Health Care 2003;19(4):646−55.

[63] MAGIC (making GRADE the irresistible choice) [Internet]. [cited 1 October 2016]. Available at http://magicproject.org.

[64] McFarlane E, Sparrow K, Willett S, Alderson P, White C, Sanabria AJ, et al. Live guideline surveillance: the future of surveillance? In Abstracts directory of the 13th G-I-N Conference; 2016 27−30 September; Philadelphia, USA; 2016. Available at http://www.g-i-n.net/document-store/g-i-n-conferences/philadelphia-2016/abstract-directory-2016.

[65] Alonso-Coello P, Martínez García L, Carrasco JM, Sola I, Qureshi S, Burgers JS. The updating of clinical practice guidelines: insights from an international survey. Implement Sci 2011;6:107.

[66] Martínez García L, Pardo-Hernández H, Sanabria AJ, Alonso-Coello P. Pregnancy Clinical Guideline Updating Working Group. Continuous surveillance of a pregnancy clinical guideline: an early experience, *Syst Rev*, 2017 [accepted].

[67] Grupo de trabajo sobre GPC. Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud. Actualización del Manual Metodológico [Internet]. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad - GuíaSalud; 2016. Available at http://portal.guiasalud.es/emanuales/elaboracion/index-02.html.

[68] Barnieh L, Jun M, Laupacis A, Manns B, Hemmelgarn B. Determining research priorities through partnership with patients: an overview. Semin Dial 2015;28:141−6.

[69] Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014;14:89.

[70] Guise JM, O'Haire C, McPheeters M, Most C, Labrant L, Lee K, et al. A practice-based tool for engaging stakeholders in future research: a synthesis of current practices. J Clin Epidemiol 2013;66:666−74.

[71] Atkins D, Perez-Padilla R, Macnee W, Buist AS, Cruz AA. ATS/-ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD guideline development. Priority setting in guideline development: article 2 in Integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. Proc Am Thorac Soc 2012;9:225−8.

[72] Bero LA, Binder L. The Cochrane Collaboration review prioritization projects show that a variety of approaches successfully identify high-priority topics. J Clin Epidemiol 2013;66:472−3.

[73] Husereau D, Boucher M, Noorani H. Priority setting for health Technology assessment at CADTH. Int J Technol Assess Health Care 2010;26:341−7.

[74] Noorani HZ, Husereau DR, Boudreau R, Skidmore B. Priority setting for health technology assessments: a systematic review of current practical approaches. Int J Technol Assess Health Care 2007;23:310−5.

[75] Cochrane Community Blog [Internet]. Foxlee R Prioritization Work Within Cochrane Rev Groups: a Ser case Stud [cited 1 October 2016]. Available at http://community.cochrane.org/news/prioritization-work-within-cochrane-review-groups-series-case-studies.

[76] Martínez García L, McFarlane E, Barnes S, Sanabria AJ, Alonso-Coello P, Alderson P. Updated recommendations: an assessment of NICE clinical guidelines. Implement Sci 2014;9:72.

[77] Peterson K, McDonagh MS, Fu R. Decisions to update comparative drug effectiveness reviews vary based on type of new evidence. J Clin Epidemiol 2011;64:977−84.

[78] Vernooij RW, Alonso-Coello P, Brouwers M, Martínez García L, CheckUp Panel. Reporting items for updated clinical guidelines: checklist for the reporting of updated guidelines (CheckUp). PLoS Med 2017;14(1):e1002207.

[79] Gentles SJ, Charles C, Nicholas DB, Ploeg J, McKibbon KA. Reviewing the research methods literature: principles and strategies illustrated by a systematic overview of sampling in qualitative research. Syst Rev 2016;5:172.

APPENDIX IV

Article: "Development of a prioritisation tool for the updating of clinical guideline questions: the UpPriority Tool protocol."

BMJ Open

Development of a prioritisation tool for the updating of clinical guideline questions: the UpPriority Tool protocol

Laura Martínez García,[1] Hector Pardo-Hernandez,[1,2] Ena Niño de Guzman,[1] Cecilia Superchi,[1] Monica Ballesteros,[1] Emma McFarlane,[3] Katrina Penman,[3] Margarita Posso,[1] Marta Roqué i Figuls,[1] Andrea Juliana Sanabria,[1] Anna Selva,[4] Robin WM Vernooij,[1] Pablo Alonso-Coello[1,2,5]

To cite: Martínez García L, Pardo-Hernandez H, Niño de Guzman E, *et al*. Development of a prioritisation tool for the updating of clinical guideline questions: the UpPriority Tool protocol. *BMJ Open* 2017;**7**:e017226. doi:10.1136/bmjopen-2017-017226

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-017226).

Received 8 April 2017
Revised 7 July 2017
Accepted 7 July 2017

[1]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain
[2]CIBER of Epidemiology and Public Health (CIBERESP), Barcelona, Spain
[3]National Institute for Health and Care Excellence, Manchester, UK
[4]Clinical Epidemiology and Cancer Screening. Corporació Sanitaria Parc Taulí, Sabadell, Barcelona, Spain
[5]Clinical Epidemiology and Biostatistics Department, McMaster University, Hamilton, Canada

Correspondence to
Dr Laura Martínez García; laura.martinez.garcia@cochrane.es

ABSTRACT

Introduction Due to a continuous emergence of new evidence, clinical guidelines (CGs) require regular surveillance of evidence to maintain their trustworthiness. The updating of CGs is resource intensive and time consuming; therefore, updating may include a prioritisation process to efficiently ensure recommendations remain up to date. The objective of our project is to develop a pragmatic tool to prioritise clinical questions for updating within a CG.

Methods and analysis To develop the tool, we will use the results and conclusions of a systematic review of methodological research on prioritisation processes for updating and will adopt a methodological approach we have successfully implemented in a previous experience. We will perform a multistep process including (1) generation of an initial version of the tool, (2) optimisation of the tool (feasibility test of the tool, semistructured interviews, Delphi consensus survey, external review by CG methodologists and users and pilot test of the tool) and (3) approval of the final version of the tool. At each step of the process, we will (1) calculate absolute frequencies and proportions (quantitative data), (2) use content analysis to summarise and draw conclusions (qualitative data) and (3) draft a final report, discuss results and refine the previous versions of the tool. Finally, we will calculate intraclass coefficients with 95% CIs for each item and overall as indicators of agreement among reviewers.

Ethics and dissemination We have obtained a waiver of approval from the Clinical Research Ethics Committee at the Hospital de la Santa Creu i Sant Pau (Barcelona). The results of the study will be published in peer-reviewed journal and communicated to interested stakeholders. The tool could support the standardisation of prioritisation processes for updating CGs and therefore have important implications for a more efficient use of resources in the CG field.

Strengths and limitations of this study

- ► To develop the tool, we will use the results and conclusions of a systematic review of methodological research on prioritisation processes for updating.
- ► We will adopt a methodological approach we have successfully implemented in a previous experience.
- ► We will collect views from clinical guidelines (CG) developers (semistructured interviews and external reviews), CG methodological experts (Delphi consensus survey) and CG users (semistructured interviews); these will allow us to pool different stakeholders' opinions about CG updating prioritisation processes.
- ► The principal limitation of the study is that we will not perform a formal validation of the tool.

INTRODUCTION

Clinical guidelines (CGs) are 'statements that include recommendations intended to optimise patient care that are informed by systematic reviews (SRs) of evidence and an assessment of the benefits and harms of alternative care options'.[1] Due to a continuous emergence of new evidence,[2 3] CGs require regular surveillance of evidence to maintain their trustworthiness.[4 5]

Several studies have assessed length of time that CGs and their recommendations remain valid.[4–8] Based on this evidence, most CG developers have adopted updating strategies based on predetermined time frames.[9]

An updating strategy involves different processes including the identification of new evidence; the assessment of the impact of new evidence on the current CG recommendations and whether an update is required and the update of the CG if needed.[9 10] The updating of CGs is resource intensive and time consuming.[11] In the current context of restricted resources, there is a growing interest in approaches that support decision-making for updating CGs.[12]

We define the prioritisation process for updating of CGs as the methodology used to determine which CGs should be prioritised to ensure that resources are invested

in updating the topics that are most relevant to different stakeholders.[12] The prioritisation process includes two main stages: (1) assessment of CGs using prioritisation criteria (eg, availability of new evidence, clinical relevance or users' interest) and (2) classification of CGs in groups according to priority for updating (eg, high, medium or low relevance for updating).[12]

Different prioritisation processes could be implemented at different time points within an updating strategy. For example, a prioritisation process could be implemented to identify the CGs in greatest need of update (prioritisation across available CGs)[13 14] or to identify the clinical questions in greatest need of update within a prioritised CG (prioritisation within a CG).[15 16]

Until now, there is wide variability and suboptimal reporting of the methods used to develop and implement processes to prioritise updating of CGs.[12]

AIMS AND OBJECTIVES

Primary objective

To develop a pragmatic tool to prioritise clinical questions for updating within a CG.

Secondary objectives

- To identify the most important items required to prioritise clinical questions for updating within a CG.
- To describe each item, establish a rating scale of items and provide a guidance on how to rate them.
- To develop guidance on how to calculate and present priority scores to support decision-making for updating clinical questions within a CG.

METHODS AND ANALYSIS

To develop the UpPriority Tool, we will use the results and conclusions of a systematic review of methodological research on prioritisation processes for updating[12] and will adopt a methodological approach we have successfully implemented in a previous experience.[17] We will perform a multistep process including (1) generation of an initial version of the tool, (2) optimisation of the tool (feasibility test of the tool, semistructured interviews, Delphi consensus survey, external review by CG methodologists and users and pilot test of the tool) and (3) approval of the final version of the tool (table 1, figure 1).

Generation of the initial version of the tool

Objective

The objective is to develop the initial version of the tool (items, scoring calculation and summary report).

Method

The UpPriority Steering Group (UpSG) will participate in informal discussion and will approve the initial version of the tool.

Participants

UpSG.

OPTIMISATION OF THE TOOL

Feasibility test of the tool

Objective

The objective is to explore the feasibility and refine the initial version of the tool.

Study design

Methodological survey.

Participants

A CG developed within the Spanish National Health System Clinical Guideline Program, published within the last 2 years and with <50 clinical questions.

Main outcome

Time to apply the tool.

Other variables

Response rate, characteristics of participants and workplace, characteristics of clinical questions, priority scores (single item and overall items) and overall assessment of the tool (table 2).

Data collection

Two reviewers from the original Guideline Development Group (GDG) and two reviewers from the UpSG will apply the initial version of the tool. We will use online software to design the survey and collect responses (www.digestepiclin.com).

Bias

To minimise non-response bias, the survey will be available online for 1 month; weekly email reminders will be sent to reviewers. To minimise observer bias, two reviewers from outside the UpSG will apply the tool.

Study size

Convenience sample.[18]

Data analysis

For quantitative data, we will calculate absolute frequencies and proportions. For qualitative data, we will use content analysis to summarise and draw conclusions (atlasti.com).[19] Questionnaires with no response in over 20% of the items will be withdrawn. We will draft a final report, discuss results and refine the initial version of the tool with the UpSG.

Semistructured interviews

Objective

The objective is to identify current practices in prioritisation processes for updating CGs and to refine the initial version of the tool.

Study design

Semistructured interviews (face-to-face, telephone or internet).

Participants

CG developers that (1) have experience in CG development and/or updating (defined as having participated in GDG and/or Guideline Updating Group (GUG) at least once in the past year) and (2) are fluent in English or Spanish. We will identify participants with the help of

Table 1 Characteristics of the multistep development process

		Optimisation of the tool						
	Generation of the initial version	**Feasibility test**	**Semistructured interviews**	**Delphi consensus survey**	**External review with clinical guidelines developers**	**External review with clinical guidelines users**	**Pilot test**	**Approval of the final version**
Objective	To develop the initial version of the tool	To explore the feasibility of the tool	To identify current practices in prioritisation processes for updating CGs	To reach a consensus about the included items of the tool	To assess the usefulness* and understanding of each item of the tool	To assess the usefulness* and understanding of each item of the tool	To explore the interobserver reliability of the final version of the tool	To approve the final version of the tool
Study design	–	Methodological survey	Semistructured interviews	Delphi consensus survey	Survey	Semistructured interviews	Methodological survey	–
Participants	UpSG	CG	CG developers	CG methodological experts from G-I-N Updating Guidelines Working Group	CG developers from G-I-N community	CG users	CG	UpSG
Main outcome	–	Time to apply the tool	Participants' experiences with prioritisation processes for updating CGs	Items considered important to prioritise clinical questions for updating within a CG	Usefulness* rating for each item of the tool	Participants' views of prioritisation processes for updating CGs with the tool	Intraclass coefficient with 95% CI	–
Study size	–	Convenience sample	Sampling saturation	20–30 participants	250 organisations and individual members	Sampling saturation	Convenience sample	–

*Usefulness: The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.[29]
CG, clinical guideline; G-I-N, Guidelines International Network; UpSG, UpPriority Steering Group.

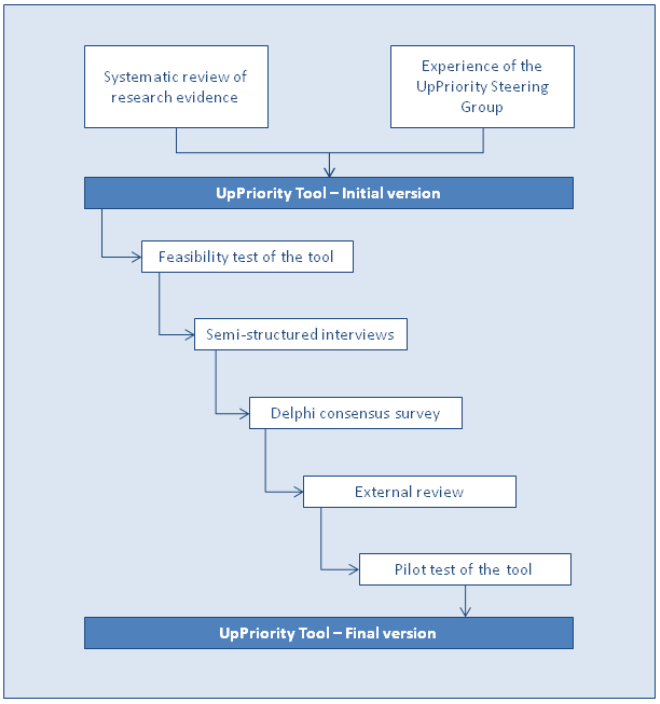


Figure 1 Multistep development process.

the UpSG. When someone does not respond or cannot participate, another contributor will be recruited.

Main outcome

Participants' experiences with prioritisation processes for updating CGs.

Other variables

Characteristics of participants and workplace, current practices in prioritisation processes for updating CGs, assessment of each item, assessment of the scoring calculation, assessment of the summary report and overall assessment of the tool (table 2).

Data collection

Interviews will be audiotaped and transcribed (each interview will last approximately 1 hour).

Bias

To minimise interviewer bias, semistructured interviews will be conducted using an interview guide.

Study size

We will recruit participants and collect data until information becomes repetitive and no new information emerges (sampling saturation).[20 21]

Data analysis

For quantitative data, we will calculate absolute frequencies and proportions. For qualitative data, we will use content analysis to summarise and draw conclusions (atlasti.com).[19] We will draft a final report, discuss results and refine the initial version of the tool with the UpSG.

Delphi consensus survey

Objective

The objective is to reach a consensus about the included items and refine the initial version of the tool.

Study design

Delphi consensus survey.

Table 2 Study variables in multistep development process

	Feasibility test	Semistructured interviews	Delphi consensus survey	External review with clinical guidelines developers	External review with clinical guidelines users	Pilot test
Response rate	X		X	X		X
Characteristics of participants and workplace	X	X	X	X	X	X
Characteristics of clinical questions	X					X
Priority scores	X					X
Current practices in prioritisation processes for updating CGs		X				
Assessment of each item		X	X (inclusion and understanding)	X (usefulness and understanding)	X (usefulness and understanding)	
Assessment of the scores calculation		X	X	X	X	
Assessment of the summary report		X	X	X	X	
Overall assessment of the tool	X	X	X	X	X	X

CG, clinical guideline.

Before the first Delphi round, we will provide the results of previous methodological research to Delphi panel members.

In the first Delphi round, we will ask participants to rate whether each item should be included in the tool and its clarity using a seven-point Likert scale (1=strongly disagree and 7=strongly agree).[22] We will calculate the median score for inclusion of each item and will classify them as (1) excluded (median score of 0–3 points), (2) review, modify and retest (median score of 4–5 points or with substantial comments) and (3) included (median score of 6 to 7 points and without substantial comments).

After each Delphi round, we will provide feedback to Delphi panel members (all responses will be anonymised prior to circulation). We will conduct additional Delphi rounds until consensus for inclusion or exclusion is reached and no more relevant comments were provided (two or three rounds, as needed).

Participants

CG methodological experts that (1) have methodological experience in CGs development and/or updating (defined as having participated in a CG technical team at least once in the past year and/or in methodological research) and (2) are fluent in English or Spanish. We will identify participants by contacting professionals associated with the Guidelines International Network (G-I-N) Updating Guidelines Working Group (http://www.g-i-n.net/working-groups/updating-guidelines) or authors of methodological research. Non-responders will not be invited to subsequent rounds.

Main outcome

Items considered important to prioritise clinical questions for updating within a CG.

Other variables (per round)

Characteristics of participants and workplace, assessment of each item (inclusion and understanding), assessment of the scoring calculation, assessment of the summary report and overall assessment of the tool (table 2).

Data collection

We will use online software to design the survey and collect responses (www.digestepiclin.com).

Bias

To minimise selection bias of Delphi panel members, all G-I-N Updating Guidelines Working Group members will be invited to participate. To minimise non-response bias, the survey will be available online for 1 month; weekly email reminders will be sent to reviewers.

Study size

Twenty to 30 participants.[23]

Data analysis

For quantitative data, we will calculate absolute frequencies and proportions. For qualitative data, we will use content analysis to summarise and draw conclusions (atlasti.com).[19] Questionnaires with no response in over 20% of the items will be withdrawn. We will draft a final report, discuss results and refine the initial version of the tool with the UpSG.

External review

External review with clinical guidelines developers

Objective

The objective is to assess the usefulness and understanding of each item and refine the initial version of the tool.

Study design

Survey.

Participants

CG developers that (1) have experience in CG development/updating (defined as having participated in GDG and/or GUG at least once in the past year) and (2) are fluent in English or Spanish. We will identify participants by contacting professionals associated with the G-I-N community (http://www.g-i-n.net).

Main outcome

Usefulness rating for each item of the tool.

Other variables

Characteristics of participants and workplace, assessment of each item (usefulness and understanding), assessment of the scoring calculation, assessment of the summary report and overall assessment of the tool (table 2).

Data collection

We will use online software to design the survey and collect responses (www.digestepiclin.com).

Bias

To minimise selection bias of survey participants, all G-I-N members will be invited to participate. To minimise non-response bias, the survey will be available online for 1 month; weekly email reminders will be sent to reviewers. Furthermore, the questionnaire will be pilot tested to improve wording and layout.

Study size

Currently, about 250 organisations and individual members are registered in the G-I-N community (http://www.g-i-n.net/membership/members-around-the-world).

Data analysis

For quantitative data, we will calculate absolute frequencies and proportions. For qualitative data, we will use content analysis to summarise and draw conclusions (atlasti.com).[19] Questionnaires with no response in over 20% of the items will be withdrawn. We will draft a final report, discuss results and refine the initial version of the tool with the UpSG.

External review with clinical guidelines users

Objective

The objective is to assess the usefulness and understanding of each item and refine the initial version of the tool.

Study design
Semistructured interviews (face-to-face, telephone or internet).

Participants
CG users (defined as healthcare professionals that use CGs on a regular basis) who are fluent in English or Spanish. We will identify participants with the help of the UpSG. When someone does not respond or cannot participate, a new contributor will be recruited.

Main outcome
Participants' views of prioritisation processes for updating CGs with the tool.

Other variables
Characteristics of participants and workplace, assessment of each item (usefulness and understanding), assessment of the scoring calculation, assessment of the summary report and overall assessment of the tool (table 2).

Data collection
Interviews will be audiotaped and transcribed (each interview will last approximately 1 hour).

Bias
To minimise interviewer bias, semistructured interviews will be conducted using an interview guide.

Study size
We will recruit participants and collect data until information becomes repetitive and no new information emerges (sampling saturation).[20 21]

Data analysis
For quantitative data, we will calculate absolute frequencies and proportions. For qualitative data, we will use content analysis to summarise and draw conclusions (atlasti.com).[19] We will draft a final report, discuss results and refine the initial version of the tool with the UpSG.

Pilot test of the tool

Objective
The objective is to explore the interobserver reliability of the final version of the tool and refine the initial version of the tool.

Study design
Methodological survey.

Participants
A CG developed within the Spanish National Health System Clinical Guideline Programme, published within the last 2 years and with <50 clinical questions.

Main outcome
Intraclass coefficient (ICC) with 95% CI for each item and overall.

Other variables
Response rate, characteristics of participants and workplace, characteristics of clinical questions and priority scores (single item) and overall assessment of the tool (table 2).

Data collection
Two reviewers from the original GDG and two reviewers from the UpSG will apply the initial version of the tool. We will use online software to design the survey and collect responses (www.digestepiclin.com).

Bias
To minimise non-response bias, the survey will be available online for 1 month; weekly email reminders will be sent to reviewers. To minimise observer bias, two reviewers from outside the UpSG will apply the tool.

Study size
Convenience sample; the results of the pilot test will inform the sample size calculation for a subsequent main study.[24]

Data analysis
For quantitative data, we will calculate absolute frequencies and proportions. For qualitative data, we will use content analysis to summarise and draw conclusions (atlasti.com).[19] Questionnaires with no response in over 20% of the items will be withdrawn. We will calculate the ICC with 95% CI for each item and overall as an indicator of agreement among reviewers. According to the scale proposed by Landis and Koch, the degree of agreement between 0.00 and 0.20 is poor, from 0.21 to 0.40 is fair, from 0.41 to 0.60 is moderate, from 0.61 to 0.80 is substantial and from 0.81 to 1.00 is almost perfect.[25] We will draft a final report, discuss results and refine the initial version of the tool with the UpSG.

APPROVAL OF THE FINAL VERSION OF THE TOOL

Objective
The objective is to approve the final version of the tool (items, scoring calculation and summary report).

Method
The UpSG will participate in informal discussion and will approve the final version of the tool.

Participants
UpSG.

ETHICS AND DISSEMINATION
We have obtained a waiver of approval from the Clinical Research Ethics Committee at the Hospital de la Santa Creu i Sant Pau (Barcelona, Spain), since this study will not involve patients or biological samples.

The results of the study will be published in peer-reviewed journal and communicated to interested stakeholders (eg, via international conferences, electronic bulletin or website).

We will develop the UpPriority tool through a comprehensive development process, including the use of previous methodological evidence,[12 17] feasibility testing of the tool and engagement of the international CG community (semistructured interviews, Delphi consensus survey and external review) and finally a pilot testing of the tool.

Previous SRs on CG updating strategies found limited evidence on processes that could inform the decision of which CGs should be prioritised for updating.[9 10 26] There are, nevertheless, new studies that underscore the relevance of the prioritisation process in CG updating,[13 27] coinciding with a growing interest among developers to shift from developing to updating CGs.[28]

We recently systematically reviewed the available evidence on strategies to prioritise the updating of SRs, health technology assessments and CGs.[12] We observed that there is wide variability and suboptimal reporting of the methods used to develop and implement such prioritisation processes. Therefore, developers may have difficulties selecting and implementing a prioritisation method to optimise the updating process of CGs.

Agbassi *et al*[13] implemented an annual step-by-step prioritisation process of CGs for updating.[13] The authors reviewed CGs using two questionnaires; the process requires evidence search, evidence review and review approval.[13] We will build our proposal on this process while addressing some of its shortcomings. Following a comprehensive development process, we will develop a pragmatic survey based tool that will likely be less resource intensive and time consuming compared with formal approaches (based on step-by-step algorithm that generally includes literature searches). We will also publish detailed and explicit guidance to allow developers to implement the tool in their institutions and to adapt it, if needed, to their specific circumstances.

We expect to develop a pragmatic tool (items, scoring calculation and summary report) that will be applicable to all clinical questions within a CG and should be easy to uptake by CG developers. The UpPriority Tool could support the standardisation of prioritisation processes for updating CGs and therefore have important implications for a more efficient use of resources in the CG field.

Contributors LMG and PAC were involved in conception and study design. LMG, HPH, ENG and CS were involved in drafting of the first version of the article. LMG, HPH, ENG, CS, MB, EM, KP, MP, MRF, AJS, AS, RWMV and PAC were involved in critical revision of the article for important intellectual content. LMG, HPH, ENG, CS, MB, EM, KP, MP, MRF, AJS, AS, RWMV and PAC were involved in final approval of the article.

Funding This study has been funded by Instituto de Salud Carlos III through the project 'PI15/00325' (Cofunded by European Regional Development Fund/European Social Fund, 'Investing in your future'). Laura Martínez García is funded by a Juan Rodés contract from the Instituto de Salud Carlos III (JR15/00004). Pablo Alonso-Coello is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (MSII15/00034).

Competing interests None declared.

Ethics approval Clinical Research Ethics Committee (Hospital de la Santa Creu i Sant Pau, Barcelona, Spain).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Steinberg E, Greenfield S, Mancher M, *et al*. *Clinical practice guidelines we can trust*. National Academies Press, 2011.
2. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010;7:e1000326.
3. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485–514.
4. Martínez García L, Sanabria AJ, García Alvarez E, *et al*. The validity of recommendations from clinical guidelines: a survival analysis. *CMAJ* 2014;186:1211–9.
5. Shekelle PG, Ortiz E, Rhodes S, *et al*. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA* 2001;286:1461–7.
6. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. *J Clin Epidemiol* 2014;67:52–5.
7. Lyratzopoulos G, Barnes S, Stegenga H, *et al*. Updating clinical practice recommendations: is it worthwhile and when? *Int J Technol Assess Health Care* 2012;28:29–35.
8. Neuman MD, Goldstein JN, Cirullo MA, *et al*. Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations. *JAMA* 2014;311:2092–100.
9. Vernooij RW, Sanabria AJ, Solà I, *et al*. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. *Implement Sci* 2014;9:3.
10. Martínez García L, Arévalo-Rodríguez I, Solà I, *et al*. Strategies for monitoring and updating clinical practice guidelines: a systematic review. *Implement Sci* 2012;7:109.
11. Martínez García L, Pardo-Hernández H, Sanabria AJ, *et al*. *Continuous surveillance of a pregnancy clinical guideline: an early experience*. 2017.
12. Martínez García L, Pardo-Hernandez H, Superchi C, *et al*. Methodological systematic review identifies major limitations in prioritization processes for updating. *J Clin Epidemiol* 2017;86:11–24. Epub ahead of print.
13. Agbassi C, Messersmith H, McNair S, *et al*. Priority-based initiative for updating existing evidence-based clinical practice guidelines: the results of two iterations. *J Clin Epidemiol* 2014;67:1335–42.
14. Jamshidi A, Lamontagne M, Ait-kadi D, *et al*. *Developing a comprehensive Priority-based framework for updating clinical Practice guidelines (CPGs): A systematic review and international survey*. Québec, Canada: Huitième Colloque Étudiant du CIRRIS, 2016.
15. Amos Q, Chan W, Tom G. *Maximising efficiency in updating guidelines through prioritisation of clinical questions [abstract]. 10th G-I-N conference*. San Francisco, California, 2013.
16. Theobald S, Blanc-Vincent MP, Farsi F, *et al*. 1999. The identification of questions in the updating process of clinical practice guidelines for oncology [abstract]. *15th Annual Meeting of the International Society of Technology Assessment in Health Care*, Edinburgh (UK)
17. Vernooij RW, Alonso-Coello P, Brouwers M, *et al*. Reporting items for updated clinical guidelines: checklist for the reporting of updated guidelines (CheckUp). *PLoS Med* 2017;14:e1002207.
18. Arain M, Campbell MJ, Cooper CL, *et al*. What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Med Res Methodol* 2010;10:67.
19. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;8:45.
20. Grossoehme DH. Overview of qualitative research. *J Health Care Chaplain* 2014;20:109–22.
21. Guetterman TC. Descriptions of sampling practices within five approaches to qualitative research in education and the health sciences. *Forum Qual Soc Res* 2015;16:25.
22. Likert R. A technique for the measurement of attitudes. *Archives of psychology* 1932.
23. Akins RB, Tolson H, Cole BR. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC Med Res Methodol* 2005;5:37.
24. Thabane L, Ma J, Chu R, *et al*. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol* 2010;10:1.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.

26. Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. *J Clin Epidemiol* 2014;67:33–45.
27. Lord J, Willis S, Eatock J, *et al*. Economic modelling of diagnostic and treatment pathways in National Institute for Health and Care Excellence clinical guidelines: the Modelling Algorithm Pathways in Guidelines (MAPGuide) project. *Health Technol Assess* 2013;17:1–192.
28. Baker M. Maintaining the Currency of a Large Guidelines Portfolio; Necessity is the Mother of Invention. *Abstracts directory of the 13th G-I-N Conference*. Philadelphia, USA, 2016. http://www.g-in.net/document-store/g-i-n-conferences/philadelphia-2016/abstract-directory-2016.
29. van der Weegen S, Verwey R, Tange HJ, *et al*. Usability testing of a monitoring and feedback tool to stimulate physical activity. *Patient Prefer Adherence* 2014;8:311–22.

BMJ Open

Development of a prioritisation tool for the updating of clinical guideline questions: the UpPriority Tool protocol

Laura Martínez García, Hector Pardo-Hernandez, Ena Niño de Guzman, Cecilia Superchi, Monica Ballesteros, Emma McFarlane, Katrina Penman, Margarita Posso, Marta Roqué i Figuls, Andrea Juliana Sanabria, Anna Selva, Robin WM Vernooij and Pablo Alonso-Coello

BMJ Open 2017 7:
doi: 10.1136/bmjopen-2017-017226

APPENDIX V

Article: "Guideline on terminology and definitions of updating clinical guidelines: the Updating Glossary."

Journal of Clinical Epidemiology 95 (2018) 28−33

Guideline on terminology and definitions of updating clinical guidelines: The Updating Glossary

Laura Martínez García[a,*], Hector Pardo-Hernández[a,b], Andrea Juliana Sanabria[a], Pablo Alonso-Coello[a,b], Katrina Penman[c], Emma McFarlane[c], on behalf of the G-I-N Updating Guidelines Working Group

[a]*Iberoamerican Cochrane Centre, Barcelona, Spain*
[b]*CIBER de Epidemiología y Salud Pública (CIBERESP), Spain*
[c]*National Institute for Health and Care Excellence, Manchester, UK*

Accepted 28 November 2017; Published online 6 December 2017

Abstract

Objective: The Guidelines International Network (G-I-N) Updating Guidelines Working Group launched an initiative to develop a glossary (the Updating Glossary) with domains, terms, definitions, and synonyms related to updating of clinical guidelines (CGs).

Study Design and Setting: The steering committee developed an initial list of domains, terms, definitions, and synonyms through brainstorming and discussion. The panel members participated in three rounds of feedback to discuss, refine, and clarify the proposed terms, definitions, and synonyms. Finally, the panel members were surveyed to assess their level of agreement regarding the glossary.

Results: Eighteen terms were identified and defined: (1) continuous updating, (2) decision to update, (3) fixed updating, (4) full updating, (5) impact of the new evidence, (6) partial updating, (7) prioritization process, (8) reporting process, (9) signal for an update, (10) surveillance process, (11) time of validity, (12) timeframe, (13) tools and resources, (14) up to date, (15) update cycle, (16) update unit, (17) updated version, and (18) updating strategy. Consensus was reached for all terms, definitions, and synonyms (median agreement scores ≥ 6); except for one term.

Conclusions: The G-I-N Updating Guidelines Working Group assembled the Updating Glossary to facilitate and improve the knowledge exchange among CGs developers, researchers, and users.

Key words: Classification; Clinical guidelines; Methodology; Terminology; Updating

1. Introduction

The volume of scientific information is increasing at an exponential rate. It is estimated that approximately 75 clinical trials and 11 systematic reviews are published every day [1]. Nevertheless, the peak in publishing production has not yet been reached [1,2].

To address the increasing volume of information and to guide decision-making with the best evidence available, resources such as clinical guidelines (CGs, also known as clinical practice guidelines or practice guidelines) acquire significant relevance. However, CGs need to remain up to date to guarantee the validity of their recommendations and maintain their usefulness for patients, health-care providers, and other stakeholders [3−7].

The updating of CGs should be based on the same systematic and transparent approaches as for de novo development. However, little attention has been paid to strategies for updating CGs. Further research is needed to develop, implement, evaluate, optimize, and standardize CG-updating strategies [8−12].

One of the challenges in the CG-updating field is the lack of standards on terminology (what do we call it?) and definitions (what does it mean?). This makes it difficult to share methods and experiences efficiently, retrieve research

Funding: Laura Martínez García is funded by a Juan Rodés contract from the Instituto de Salud Carlos III (JR15/00004). Pablo Alonso-Coello is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (MSII15/00034).

Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of National Institute for Health and Care Excellence.

* Corresponding author. Iberoamerican Cochrane Centre, Hospital de la Santa Creu i Sant Pau, Sant Antoni Maria Claret 167, 08025, Barcelona, Spain. Tel.: +34 93 553 78 08; fax: +34 93 553 78 09.

E-mail address: laura.martinez.garcia@cochrane.es (L. Martínez García).

https://doi.org/10.1016/j.jclinepi.2017.11.023

What is new?

Key findings

- G-I-N Updating Guidelines Working Group consensus on domains, terms, definitions, and synonyms in the clinical guideline (CG)–updating field.
- Eighteen terms were identified and defined; 11 synonyms were identified and linked to six terms.

What this adds to what was known?

- We developed three conceptual domains: time (when?), method (how?), and unit (what?).
- We developed four strategic domains: approach, strategy or method, process, and task.

What is the implication and what should change now?

- The Updating Glossary should facilitate and improve the knowledge exchange among CGs developers, researchers, and users.
- The Updating Glossary could support future methodological research (e.g., retrieving previously published research, communicating research findings, and identifying research gaps).
- As CGs, the Updating Glossary needs to remain up to date to guarantee the validity of their domains, terms, and maintain their usefulness for CGs developers, researchers, and users.

evidence previously published, communicate research findings, or identify research gaps [8–13]. The Guidelines International Network (G-I-N) Updating Guidelines Working Group (http://www.g-i-n.net/working-groups/updating-guidelines) launched an initiative to develop a glossary (the Updating Glossary) with domains, terms, definitions, and synonyms related to updating of CGs.

2. Methods

2.1. Participants

An Updating Glossary steering committee was convened to design and coordinate this initiative. The steering committee was responsible for the development of the first version of the glossary and the analysis of the feedback provided by the Updating Glossary panel members.

The Updating Glossary panel members were assembled from institutions that develop CGs belonging to the G-I-N Updating Guidelines Working Group. The panel members were responsible for the review of the proposed glossary and provide of feedback.

2.2. Development process

Based on systematic reviews of methodological research evidence in updating field [9,10,12], the steering committee developed an initial list of domains, terms, definitions, and synonyms through brainstorming and discussion.

The panel members participated in three rounds of feedback to discuss, refine, and clarify the proposed terms, definitions, and synonyms.

One member of the steering committee reviewed and summarized the panel members' feedback and suggested, if necessary, modifications to the terms, definitions, and synonyms. The steering committee then discussed the results and agreed a new version of terms, definitions, and synonyms.

Finally, the steering committee surveyed the panel members to assess their level of agreement using a 7-point Likert scale (from 1: strongly disagree to 7: strongly agree) for each term, definition, and synonym. We used online software to design the survey and to collect the responses (http://www.digestepiclin.com).

2.3. Data analysis

Descriptive statistics were used to calculate participation rates (frequencies and percentages) and to assess levels of agreement (medians and ranges).

3. Results

3.1. Panel members

All members of the G-I-N Updating Guidelines Working Group were invited to participate. Thirteen (13/23; 56.5%) members participated in the first round of feedback (June 2016), 17 (17/33; 51.5%) members in the second round (December 2016), 23 (23/38; 60.5%) members in the third round (March 2017), and 22 (22/39; 56.4%) members in the consensus survey (June 2017).

Consensus was reached for all terms, definitions, and synonyms (median agreement scores ≥ 6), except for one term (''time of validity'' with median agreement score of 5) (Table 1).

3.2. Updating taxonomy

Two classifications were devised to contextualize the proposed terms: a conceptual domains (time, methods, and unit) and a strategic domains (approach, strategy or method, process, and task).

1. Conceptual domains: The terms can be outlined within three conceptual domains: time (when?), method (how?), and unit (what?) (Fig. 1).

In the development of the Updating Glossary, CGs have been used as the update unit. However, definitions can be modified depending on whether the updating strategy is implemented in sections of a CG, clinical questions, or recommendations (Fig. 2).

Table 1. Levels of agreement for each term, definition and synonyms

Term	Term scores median (range)	Definition scores median (range)	Synonyms scores median (range)
1. Continuous updating	6 (2–7)	6 (2–7)	6 (1–7)
2. Decision to update	7 (1–7)	7 (5–7)	-
3. Fixed updating	6 (2–7)	6 (2–7)	-
4. Full updating	6,5 (2–7)	7 (2–7)	-
5. Impact of the new evidence	7 (4–7)	7 (5–7)	-
6. Partial updating	7 (2–7)	7 (4–7)	-
7. Prioritization process	7 (4–7)	7 (2–7)	-
8. Reporting process	6 (2–7)	6 (3–7)	-
9. Signal for an update	6 (3–7)	6,5 (3–7)	6 (2–7)
10. Surveillance process	7 (5–7)	6 (2–7)	7 (2–7)
11. Time of validity	5 (1–7)	6 (1–7)	-
12. Timeframe	6 (2–7)	6 (3–7)	6 (2–7)
13. Tools and resources	7 (3–7)	7 (3–7)	-
14. Up to date	7 (4–7)	7 (1–7)	6 (2–7)
15. Update cycle	6,5 (3–7)	6 (3–7)	-
16. Update unit	6,5 (4–7)	6,5 (5–7)	-
17. Updated version	7 (6–7)	7 (4–7)	7 (4–7)
18. Updating strategy	7 (5–7)	6 (3–7)	-

2. Strategic domains: The terms can be outlined within four strategic domains: approach (perspective for dealing with a certain situation), strategy or method (a plan of action designed to achieve an overall aim), process (series of actions to achieve a particular aim), and task (a piece of work to be done) (https://en.oxforddictionaries.com) (Fig. 3). These domains aim to facilitate the design of an updating strategy.

3.3. Updating glossary

Eighteen terms were identified and defined (listed in alphabetical order); 11 synonyms were identified and linked to six terms.

1. Continuous updating: Continuous updating involves a prospective approach and active processes that use continuous surveillance and a rapid response to include new relevant evidence identified into a CG. Synonyms: Living or dynamic CG.
2. Decision to update: Judgment on the need for updating a CG, taking into account signals for an update.
3. Fixed updating: Fixed updating involves a decision to update a CG in a fixed timeframe. The identification of new evidence and evaluation of the impact of the new evidence could be continuous (surveillance process) or fixed.
4. Full updating: When a full update of a CG is undertaken; the updating processes leads to a change of the entire CG document.
5. Impact of the new evidence: Changes in a CG in response to the new relevant evidence. Examples (at recommendation level) include the following factors: modified recommendations, new recommendations, deleted recommendations, or valid recommendation (in this last case, where the new evidence identified does not have an impact on the recommendation).
6. Partial updating: When a partial update of a CG is undertaken; the updating processes may include only certain sections, clinical questions, or recommendations.
7. Prioritization process: Assessment and ranking of CGs, within a defined collection of CGs, according to the need for updating.
8. Reporting process: Methods of communicating to users the changes in an updated version of a CG.
9. Signal for an update: New relevant evidence that may have an impact on a CG. New relevant evidence may be related with new studies, clinical expert input, or safety alerts. The potential changes may be related to clinical questions components (patients, intervention, comparisons, or outcomes) or factors that influence the formulation of recommendations (e.g., quality of the evidence, balance between benefits

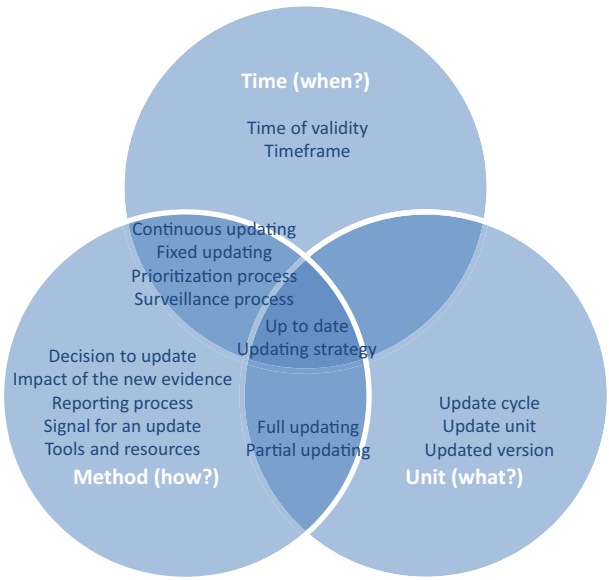


Fig. 1. Clinical guidelines updating conceptual domains.

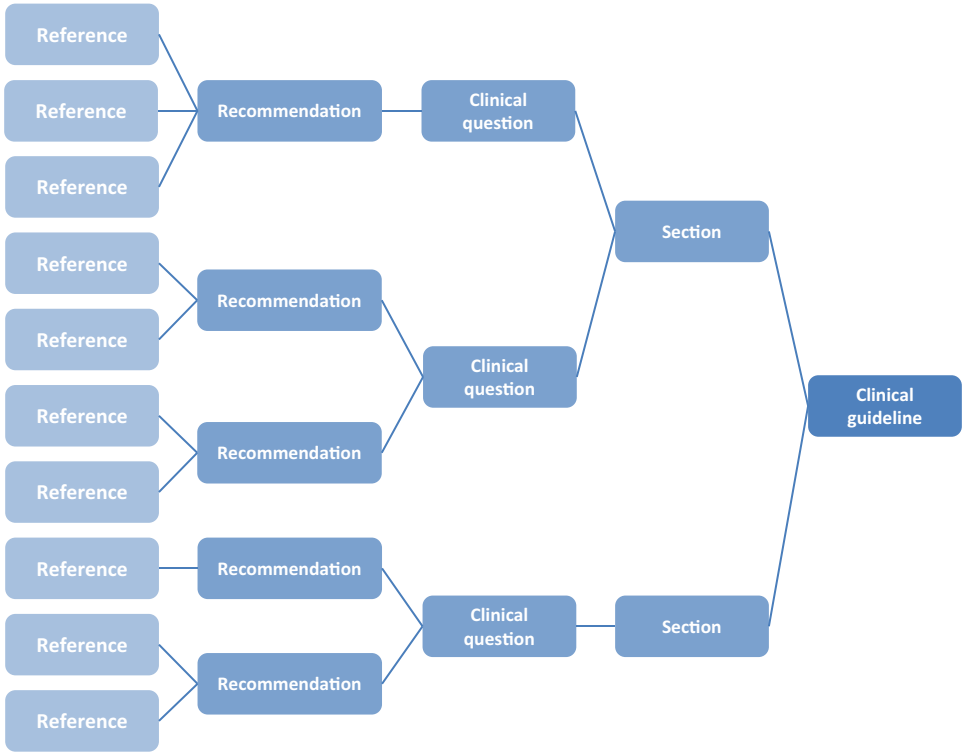


Fig. 2. Clinical guidelines update units.

and harms, values and preferences, or use of resources and costs). The potential impact may be assessed qualitatively or quantitatively. Synonyms: Trigger for an update and new relevant evidence.

10. Surveillance process: This process comprises the detection of new evidence and the evaluation of its impact on a CG. Detection of new evidence may include literature searches, surveys of clinical experts, or alert systems. Evidence surveillance is a continuous, prospective, and active process. Synonyms: Monitoring process.
11. Time of validity: Time between the publication of a CG and the identification of at least one signal for updating.
12. Timeframe: Time interval between successive updating processes. Synonyms: Time point and time interval.
13. Tools and resources: All available elements for organizing and optimizing the updating processes.
14. Up to date: CG that still includes valid recommendations to guide decision-making. Synonyms: Current, keep up, and valid.
15. Update cycle: Each updating strategy, which is implemented in a CG, corresponds to an update cycle and results in an updated version of the CG.
16. Update unit: The component of a CG upon which decisions to update are made. Examples include the whole CG document or selection of subsections, clinical questions, or recommendations (Fig. 2).
17. Updated version: CG version with the results of the updating strategy. It is important to differentiate between a full or partial updated version. Synonyms: Updated edition.

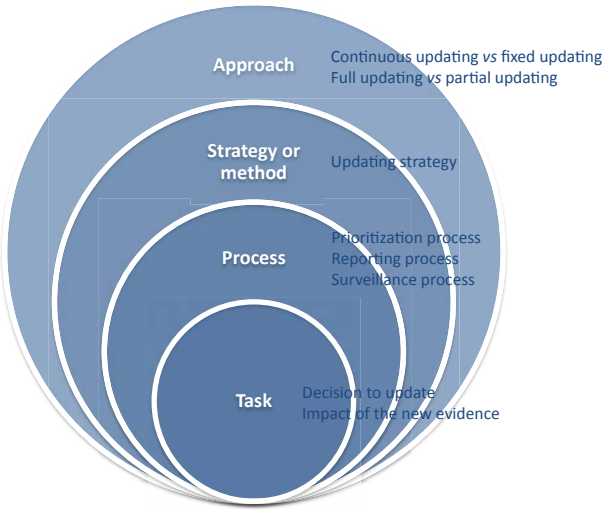


Fig. 3. Clinical guidelines updating strategic domains.

18. Updating strategy: Iterative set of processes with a systematic and explicit methodology that involves identifying and reviewing new evidence that had not been included in the current version of a CG. If new relevant evidence is identified and it is considered to have an impact on the current CG, the CG should be modified, if necessary.

The methodology and reporting of CGs should also be improved if necessary during the updating processes (e.g., introducing a new system for grading the quality of the evidence and the strength of recommendations [14,15], incorporating editorial changes or improving readability).

The three essential processes in a CG-updating strategy are (1) identifying relevant new evidence, (2) assessing whether the new evidence has an impact on the current CG and whether updating is required, (3) reviewing and, if necessary, modifying the CG. If the identification of new relevant evidence and the assessment of the impact are continuous, this is considered a surveillance process.

4. Discussion

4.1. Main findings

Through this initiative, the G-I-N Updating Guidelines Working Group has reached a consensus on domains, terms, definitions, and synonyms in the CG-updating field. Three conceptual domains (time, method, and unit) and four strategic domains (approach, strategy or method, process, and task) were identified; although most terms were classified under more than one domain. Eighteen terms were identified and defined; 11 synonyms were identified and linked to six terms.

The most controversial issue was determining the difference between ‘‘signal for an update’’ and ‘‘impact of the new evidence’’. The first term was only associated with identifying new relevant evidence and the second with whether this evidence has an effect on a given recommendation. In addition, the term ‘‘time of validity’’ received a low agreement score, even though its definition appeared to be clear.

4.2. Our results in the context of previous research

Up to now, a few methodological systematic reviews about updating CGs have been published [8−12,16]. Occasionally, there are inconsistencies with the terms used between these reviews; although, some of them were developed by the same authors. This situation highlights the need to a standardized terminology.

Other glossaries on research methods have been published (e.g., for scoping studies or for measurement of properties [COnsensusbased Standards for the selection of health Measurement INstruments study]) [17,18]. As in the CG-updating field, these glossaries were developed in light of new and complex methodological research [19−21].

4.3. Strengths and limitations

The Updating Glossary is the result of an international multidisciplinary collaborative initiative involving the G-I-N Updating Guidelines Working Group members. Their engagement and active participation in this initiative may facilitate the uptake of the glossary among CG developers. The open discussion conducted during the glossary development enriched the final outcome of this initiative.

Our work is subject to some limitations. First, we did not provide a methodological guidance, so the user may miss the potential connections between different terms or more practical information. Second, a more structured design to reach consensus among experts could have been adopted (e.g., Delphi consensus process).

4.4. Implications for practice and research

As CG developers reach their maximum guidelines production capacity, and given the limitation of resources, updating is becoming a global challenge [22]. In this context, it is essential to improve communication among CG developers to share methods and experiences in CG-updating field [13].

The implementation of Updating Glossary terms could facilitate retrieving previously published research, communicating research findings, and identifying research gaps in the CG-updating field [23].

To continue improving the Updating Glossary, it will be necessary to assess its applicability in other related health decision-making tools (e.g., systematic reviews or health technology assessments) and to test its usability among developers, researchers, and users. In the future, it is also warranted to update the glossary (new terms, reviewed and modified terms, reviewed and not modified terms, or withdrawn terms) through a formal consensus process, and to include practical examples.

5. Conclusions

The implementation of the Updating Glossary could facilitate and improve sharing methods and experiences, retrieving previously published research, communicating research findings, and identifying research gaps in the CG-updating field.

Acknowledgments

Members of the G-I-N Updating Guidelines Working Group: Martínez García L., Pardo-Hernández H., Sanabria A.J., Alonso-Coello P., Penman K., McFarlane E., Blanchard S., Brereton L., Browers M., Dean V., Flórez Gómez I.D., Fuentes C., Grimmer K., Harris J., Haynes C., Iorio A., James R., Kwong J., Lynch R., Nolan K., Ogunremi T., Okechukwu K., Prabhu Iyer N., Qaseem A., Rey M.,

Schorr S., Selva A., Shaw B., Shearn P., Shin E.S., Stapon C., Tam I., Thornton J., Uhl S., Vandvik P.O., and Vernooij R.W.M.

The authors would like to thank María Victoria Leo for her help in editing the final version of the manuscript.

G-I-N is formally constituted as a Scottish guarantee company under Company Number SC243691 and recognized as a Scottish charity under Scottish Charity Number SC034047 (Registered Office at J. & H. Mitchell W.S., 51 Atholl Road, Pitlochry, Perthshire PH16 5BY, Scotland).

References

[1] Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med 2010;7(9):e1000326.

[2] Tsay MY, Yang YH. Bibliometric analysis of the literature of randomized controlled trials. J Med Libr Assoc 2005;93:450−8.

[3] Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. J Clin Epidemiol 2014;67:52−5.

[4] Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. Updating clinical practice recommendations: is it worthwhile and when? Int J Technol Assess Health Care 2012;28(1):29−35.

[5] Neuman MD, Goldstein JN, Cirullo MA, Schwartz JS. Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations. JAMA 2014;311: 2092−100.

[6] Martínez García L, Sanabria AJ, Garcia Alvarez E, Trujillo-Martin MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ 2014;186(16):1211−9.

[7] Shekelle P, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001;286:1461−7.

[8] Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol 2014;67:33−45.

[9] Martínez García L, Arevalo-Rodriguez I, Sola I, Haynes RB, Vandvik PO, Alonso-Coello P. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci 2012; 7:109.

[10] Martínez García L, Pardo-Hernandez H, Superchi C, Nino de Guzman E, Ballesteros M, Ibargoyen Roteta N, et al. Methodological systematic review identifies major limitations in prioritisation processes for updating. J Clin Epidemiol 2017;86:11−24.

[11] Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. A systematic review identified few methods and strategies describing when and how to update systematic reviews. J Clin Epidemiol 2007;60:1095−104.

[12] Vernooij RW, Sanabria AJ, Sola I, Alonso-Coello P, Martínez García L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci 2014;9:3.

[13] Alonso-Coello P, Martinez Garcia L, Carrasco JM, Sola I, Qureshi S, Burgers JS. The updating of clinical practice guidelines: insights from an international survey. Implement Sci 2011;6:107.

[14] Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook. Edinburgh: Scottish Intercollegiate Guidelines Network; 2011.

[15] Scottish Intercollegiate Guidelines Network (SIGN). SIGN 50: a guideline developer's handbook. Edinburgh: SIGN; 2015: (SIGN publication no. 50). [November 2015]. Available at: http://www.sign.ac.uk.

[16] Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ 2014;186(3):E123−42.

[17] O'Brien KK, Colquhoun H, Levac D, Baxter L, Tricco AC, Straus S, et al. Advancing scoping study methodology: a web-based survey and consultation of perceptions on terminology, definition and methodological steps. BMC Health Serv Res 2016;16:305.

[18] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010; 63:737−45.

[19] Gentles SJ, Charles C, Nicholas DB, Ploeg J, McKibbon KA. Reviewing the research methods literature: principles and strategies illustrated by a systematic overview of sampling in qualitative research. Syst Rev 2016;5(1):172.

[20] Scientific Resource Centre (SRC) Information Resource Services. Effective Health Care Program. USA: Agency for Healthcare Research and Quality. Available at: https://effectivehealthcare.ahrq.gov/index.cfm/who-is-involved-in-the-effective-health-care-program1/about-the-scientific-resource-center1/src-information-resource-services/#SRCML.

[21] Wulff HR. The language of medicine. J R Soc Med 2004;97(4): 187−8.

[22] Baker M. Maintaining the Currency of a Large Guidelines Portfolio; Necessity is the Mother of Invention. Philadelphia, USA: 13th G-I-N Conference; 2016.

[23] Martínez-García L, Pardo-Hernandez H, Hidalgo Armas L, Sanabria AJ, Solà I, Alonso-Coello P. Updating DataBase: an open access living repository of methodological documents about updating systematic reviews and clinical guidelines. Cape Town, South Africa: Global Evidence Summit; 2017.

APPENDIX VI

Article: "Continuous surveillance of a pregnancy clinical guideline: an early experience."

Martínez García et al. Systematic Reviews (2017) 6:143
DOI 10.1186/s13643-017-0506-7

Systematic Reviews

METHODOLOGY Open Access

Continuous surveillance of a pregnancy clinical guideline: an early experience

Laura Martínez García[1*], Hector Pardo-Hernández[1,2], Andrea Juliana Sanabria[1], Pablo Alonso-Coello[1,2] on behalf of the Pregnancy Clinical Guideline Updating Working Group

Abstract

Background: To date there is no consensus about the optimal strategy for keeping clinical guidelines (CGs) up-to-date. The aims of this study were (1) to develop a continuous surveillance and updating strategy for CGs and (2) to test the strategy in a specific CG.

Methods: The main steps were as follows: (1) recruiting members for the CG Updating Working Group, (2) mapping the CG, (3) identifying new evidence from the CG Updating Working Group, (4) designing and running restricted literature searches, (5) reviewing drugs and medical devices alerts, (6) screening and assessing the new evidence, (7) reviewing and, if necessary, modifying clinical questions and recommendations, and (8) updating the CG document.

Results: The Pregnancy CG Updating Working Group consisted of 29 members, including clinicians, patients and caregivers, and clinical guideline methodology experts. We selected 69 clinical questions (123 recommendations) from the "Assistance during pregnancy" section.
For the first update cycle (32-month duration), 9710 references were identified. Of these, 318 were pertinent, 289 were relevant, and 55 were classified as potential key references. For the second and third update cycles (6-month duration each), 2160 and 2010 references were retrieved, respectively. The continuous surveillance and updating strategy has not yet been completely implemented.

Conclusions: Further resources are needed in updating the CG field, both for implementing updating strategies and for developing methodological research.

Keywords: Diffusion of innovation, Dissemination and implementation, Evidence-based medicine, Methodology, Practice guidelines, Updating

Background

Clinical guidelines (CGs) are useful tools to help patients, health care providers, and policymakers make evidence-based decisions about health care. Consequently, they need to be updated in order to guarantee the validity of their recommendations.

Time of validity

Several studies have assessed the time of validity of CGs and their recommendations (defined as "time between the publication of a CG and the identification of new relevant evidence") [1–5]. Data showed that recommendations quickly became outdated (about 20% of the recommendations were out of date within 3 years) [2]. Based on this evidence, 40% (14/35) methodological handbooks typically recommend reviewing and updating guidelines within 2 to 3 years of their publication [6]. Nevertheless, CG developers acknowledge that it is challenging to maintain these predetermined timeframes [7].

Updating strategy

The updating of CGs is an iterative process that involves an explicit and systematic methodological approach for the identification and assessment of new evidence not included in the original CG [6, 8, 9]. If new relevant evidence is identified and it was considered to have an

* Correspondence: laura.martinez.garcia@cochrane.com
[1]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Sant Antoni Maria Claret 167, 08025 Barcelona, Spain
Full list of author information is available at the end of the article

impact on the current CG, the CG should be modified, if necessary [6, 8, 9]. Moreover, the updating strategies provide an opportunity to improve the overall methodology and edition of the document (e.g. correction of mistakes or enhancement to the writing) [6, 8, 9].

To date, there is no real consensus on the optimal strategy for updating CGs [6, 7, 10]. Most of the available methodological research focuses on identifying new relevant evidence. Research suggests that pragmatic search strategies (with the aim of favouring precision over sensitivity) are efficient and feasible for retrieving new evidence that triggers a recommendation update [6, 11, 12].

Living clinical guidelines

To address the updating of CGs, the majority of CG developers support the concept of living CGs [7], generally defined as "prospective and active processes that use continuous surveillance and a rapid response to include new relevant evidence identified" [9]. However, until now, no guidance has been developed to put this concept in practice [6], and a few empirical experiences were published [13–17]. CG developers considering the transition to living CGs will have to address challenges to operationalize the process [18].

Objectives

In order to address some of the challenges related with living CGs, this study aimed to (1) develop a continuous surveillance and updating strategy for CGs and (2) test the strategy in a specific CG.

Methods

We conducted a cohort study of recommendations from the Assistance During Pregnancy and Puerperium CG included in the Spanish National Health System CG Programme [19].

Living strategy

The strategy was developed based on published methodological research and the experience of the CG Updating Working Group technical team [2, 6–8, 10, 11, 20, 21]. The processes included in the strategy were as follows: (1) recruitment of members for the CG Updating Working Group, (2) mapping of the CG, (3) identification of evidence from the CG Updating Working Group, (4) designing of restricted literature search strategy, (5) running of restricted literature searches, (6) reviewing alerts for drugs and medical devices, (7) development of reference database, (8) first reference screening, (9) second reference screening (assessment of new evidence impact), (10) development of a clinical questions database, (11) classification of clinical questions, (12) review and, if necessary, modification of clinical questions and recommendations, and (13) update of the CG manuscript (Table 1). Following, we provided a detailed description of the most complex processes: (1) identification of new evidence, (2) reference screening, and (3) classification of clinical questions.

Identification of new evidence

Three different strategies were used to identify new evidence: (1) a questionnaire sent to the CG Updating Working Group, (2) a restricted literature search strategy, and (3) a revision of available drugs and medical devices alerts.

The questionnaire sent to the CG Updating Working Group aimed to identify any new relevant evidence that could have an impact on the CG (questionnaire available from the authors upon request). The questionnaire covered the different areas of the CG including the scope, new potential aspects not included in the original version, or new relevant evidence assessing the effectiveness and safety of the interventions. The survey also included questions about other relevant factors such as changes in the relative importance of the outcomes, changes in the resource use and cost of the interventions, equity, acceptability, or feasibility issues that might have arisen since the publication of the CG. Information about ongoing research studies was also sought in the survey.

Restrictive literature search strategies were developed for each clinical question in MEDLINE (PubMed) following a validated methodology described elsewhere [11]. In summary, the minimum number of Medical Subject Headings (MeSH) terms and text words required from the original exhaustive search strategies were selected. The strategies were validated checking that all key references supporting the recommendations in the original CG were retrieved and were refined if needed. Once the strategies were validated, PubMed Clinical Queries filters were applied (www.ncbi.nlm.nih.gov/pubmed/clinical). Search strategies by topic were also developed using a specific filter to identify studies on how patients and other stakeholders value health outcomes and economic studies [22, 23].

Finally, drugs and medical devices alerts published by the Spanish Agency for Medicines and Health Products were reviewed (www.aemps.gob.es/en/home.htm).

Reference screening

The references were sequentially classified in the following:

- Pertinent references: topic-related references that met the study design criteria
- Relevant references: pertinent references that could be used when considering an update to a recommendation, but that would not necessarily trigger a potential update
- Potential key references: relevant references that could potentially trigger an update

Table 1 Description of the continuous surveillance and updating strategy

Process	Description	Participants
1. Recruitment of members for the CG Updating Working Group	– Contact the CG Development Group to invite them to participate in the implementation of the strategy. – Replace non-respondents or those who declined with new members.	– Technical team
2. Mapping of the CG	– Identify clinical questions, recommendations, and references in the CG. – Compile original documentation (searches, references, evidence syntheses, and GRADE evidence profiles).	– Technical team
3. Identification of evidence from the CG Updating Working Group	– Distribute a questionnaire via email among the CG Updating Working Group for identifying new evidence.	– Clinical team – Patients and carers team
4. Design of restricted literature search strategy	– Design and validate restricted search strategies per clinical question. – Design search strategy for costs and resources use and for patients' values and preferences.	– Technical team
5. Running of restricted literature searches	– Conduct restricted searches in MEDLINE (through PubMed).	– Technical team
6. Review of alerts for drugs and medical devices	– Identify alerts for drugs and medical devices issued by the Spanish Agency for Medicines and Health Products.	– Technical team
7. Development of references database	– Develop references database and identify duplicates among the different information sources and between the original and updated CGs.	– Technical team
8. First reference screening	– Identify *pertinent references* (topic-related references with a fitting study design).	– Technical team
9. Second reference screening (assessment of new evidence impact)	– Develop a questionnaire to identify: (*1*) *relevant references*: references that are pertinent for updating a recommendation but that actually do not trigger an update and (*2*) *potential key references*: references that could potentially trigger an update of a recommendation.	– Clinical team – Technical team
10. Development of a clinical questions database	– Select clinical questions with pertinent, relevant, and key references.	– Technical team
11. Classification of clinical questions	– Analyse clinical questions database to identify: (*1*) *clinical questions to be reviewed*: with potential key references and with different relevant references or important pharmacological alerts, (*2*) *valid clinical questions*: without potential key references associated and (*3*) *new clinical questions*.	– Technical team
12. Review and, if necessary, modification of clinical questions and recommendations	– Assessment of the potential key references. – Update recommendations if necessary. – Identify key references (references that have triggered changes in one or more recommendations). – Reach a consensus with the CG Updating Working Group on the suggested updates.	– Clinical team – Patients and carers team – Technical team
13. Update of the CG manuscript	– Incorporate updates in the previous version of the CG manuscript.	– Technical team

CG clinical guideline

A specific questionnaire for each clinical question was used to identify relevant and key references (questionnaire available from the authors upon request). The questionnaire included the clinical question, the recommendations, and the references considered pertinent in the first screening to that clinical question. If the reference was considered relevant for that particular question by the reviewer, then it was deemed necessary to assess if the reference could potentially trigger an update (key reference). If it was the case, it was necessary to explicitly state which part of the question and/or recommendations was affected (population, intervention, comparator, outcomes, resource use and costs, equity, acceptability, feasibility, strength or direction of the recommendation).

Classification of clinical questions

Each clinical question was classified in one of the following categories:

- Clinical question to be reviewed: question with potential key references or with alerts
- Valid clinical question: question without potential key references or without alerts
- New clinical question

Once the questions were classified, we planned to update them following a similar method used in the development of the original recommendations but taking into account the new evidence identified and the evidence used to develop the recommendations.

Update cycle

Conducting the 13 processes was considered a one update cycle. The first update cycle included new evidence since the last search date in the CG development process up to the first search date in the CG surveillance process; subsequent update cycles were scheduled every 6 months [2].

Data analysis

We performed a descriptive analysis of the data: literature search time periods, number of identified references, number of screened references, and number of classified references (pertinent, relevant, and key). We described narratively the steps achieved.

Results

Clinical Guideline Updating Working Group

All members of the Pregnancy CG Development Working Group were initially contacted (20 members). However, since only five agreed to participate in this study, 30 additional candidates were contacted. The Pregnancy CG Updating Working Group finally consisted of 29 members: (1) clinical team: three medical specialists in gynaecology and obstetrics, three medical specialists in family and community medicine, and three midwives; (2) patients and caregivers team: three patients or patient representatives; and (3) technical team: 17 CG methodologists.

Mapping process

We identified 89 clinical questions and 201 recommendations in the Assistance During Pregnancy and Puerperium CG [19]. We focused specifically on the "Assistance during pregnancy" section, which contained 69 clinical questions and 123 recommendations (36 strong recommendations, 49 weak recommendations, and 38 good clinical practice statements). We also retrieved the references used to support recommendations, original literature search strategies, evidence syntheses, and GRADE evidence profiles.

Continuous surveillance process

We contacted a total of 13 members of the Pregnancy CG Updating Working Group for the baseline survey and received 11 responses (84.6% response rate). We developed one search strategy per clinical question (a total of 62, as 7 clinical questions were clustered together), as well as topic searches for studies on patients' values and preferences and for costs and resource use.

We identified 26 recommendations (26/123; 21.1%) related to drugs or dietary supplements. We consulted drugs and medical devices alerts from the Spanish Agency for Medicines and Health Products, searching by CG as topic.

For the first literature search cycle (32-month period), 9710 references were identified. Of these, the technical team classified 318 as pertinent, 289 as relevant, and 55 as potential key references (Table 2).

For the second and third literature search cycles (each a 6-month period), 2160 and 2010 references were retrieved, respectively (Table 2).

The surveillance process lasted 1 year, from November 2014, when the CG development institution was contacted for establishing the Pregnancy CG Updating Working Group, until November 2015, when the study was stopped early due to budgetary constraints.

The continuous surveillance and updating strategy has not yet been completely implemented. We have not assessed the results of the second and third cycles of the literature search or gauged the effect on recommendations of potential key references. As such, we have not reviewed and, if necessary, updated the CG recommendations.

Discussion

Main findings

We designed a step-by-step process for continuous surveillance and updating CGs. We were able to implement a continuous and restricted literature search strategy for the "Clinical Practice Guideline on Assistance during

Table 2 Preliminary results of the continuous surveillance implementation

	First update cycle	Second update cycle	Third update cycle
Literature search			
–Search dates	01/01/2012 31/08/2014	01/09/2014 28/02/2015	01/03/2015 31/08/2015
–Time period included (months)	32	6	6
Results of the literature search			
–Evidence identified from the CG Updating Working Group	19	NC	NC
–References on efficacy	9191	2089	1946
–References on costs and resource use	116	51	19
–References on patients' values and preferences	384	10	39
–Drug alerts	NA	10	6
Total	9710	2160	2010
Results of reference screening			
–Pertinent references	318	NC	NC
–Relevant references	289	NC	NC
–Potential key references (≥1 participants)	184	NC	NC
–Potential key references (≥2 participants)	31	NC	NC
–Potential key references (CG methodology experts)	55	NC	NC

NA not available, *NC* not completed

Pregnancy and Puerperium" for a 1-year period. In the first update cycle we identified 9710 references (318 pertinent, 289 relevant, and 55 potential key references). For the second and third update cycles 2160 and 2010 references were retrieved, respectively.

The continuous surveillance and updating strategy has not yet been completely implemented due to budgetary constraints.

Our results in the context of previous research

Only one previous study, published in 2003, assessed a continuous surveillance and updating strategy for CGs, specifically for cancer guidelines. This approach included a continuous and exhaustive literature search strategy, evaluation of the newly found evidence, review and updating of recommendations, and dissemination of the new evidence and modified recommendations among stakeholders. Similarly to our experience, the authors of this study highlighted the considerable resources required [17].

Other initiatives have ventured the implementation of new technologies to facilitate the CG updating process. One of them, called MAGIC (Making GRADE the Irresistible Choice) provides a publication platform where the main content of CGs can be disseminated. MAGIC also facilitates uploading modifications, including any potential updates, which would be available to users instantly [24]. Similarly, "Kidney Disease: Improving Global Outcomes (KDIGO)" recently published a series of recommendations for a continuous, dynamic strategy for maintaining their CGs current. Their model heavily relies on the availability and processing of new evidence using integrated electronic platforms [25]. Unfortunately, these new technologies have not yet been formally implemented and evaluated.

Strengths and limitations

We were able to retrieve, organise, and map the original documentation related to the development of the assessed CG, including the clinical questions, recommendations and references, original literature search strategies, evidence syntheses, and GRADE evidence profiles. We also adopted a systematic and continuous approach (every 6 months) to identify new evidence and to assess its impact on the CG recommendations. Lastly, we introduced evidence searches for patients' values and preferences and for costs and resource use in the surveillance process.

However, our work is subject to some limitations. First of all, we have not been able to assess the impact of the new evidence on clinical questions and recommendations for either the second or third update cycles. In addition, we have not reviewed or modified clinical questions and recommendations based on the identified new evidence in any of the update cycles.

Second, we had difficulties assembling the Pregnancy CG Updating Working Group. The vast majority of the CG Developing Group did not take part in the implementation of the strategy. Hence, an almost new working group had to be set up for this purpose. On the other hand, some members of the Pregnancy CG Updating

Working Group withdrew during the study, probably due to an excessive study-related workload (appraisal of a high volume of publications in the first update cycle, inadequate training related to the implementation of the strategy, and/or a lack of knowledge of the content of the original CG).

Third, the first surveillance cycle was quite resource-intensive (from the last search date in the original CG development process to the first search date in the CG surveillance process) and required the retrieval, mapping, and classification of the documentation generated during the development of the original CG. The process was optimised in the subsequent update cycles (second and third) given that (1) the process had already started and (2) the time between cycles (6 months) and, consequently, the volume of references (approx. 2000 references) were smaller.

Lastly, and related to the previous limitation, we did not have adequate funding to take on the management and development of the completely continuous surveillance and updating strategy we originally intended to implement. CG developers should consider using different surveillance and updating strategies to maintain their CGs up-to-date (a living strategy might not be suitable for all CGs). More research is needed to identify which CGs, topics, or areas could benefit from this or other approaches.

Conclusions

Implementing a continuous and restricted literature search process is a potentially feasible approach for the surveillance of new evidence. A continuous surveillance and updating strategy (as living CG) requires long-term substantial resources for its adoption. Further resources are needed in the updating CG field, both for implementing updating strategies and for developing methodological research.

Abbreviations

CG: Clinical guideline

Acknowledgements

The authors would like to thank Maria Victoria Leo for her help in editing the final version of the manuscript.

The authors would like to acknowledge the members of the Pregnancy Clinical Guideline Updating Working Group: Laura Martínez García, Hector Pardo-Hernández, Andrea Juliana Sanabria, Pablo Alonso-Coello, Longinos Aceituno-Velasco, Ignacio Araya, Emilia Bailon Muñoz, Petra Díaz del Campo, Itziar Etxeandia-Ikobaltzeta, Manuel Fillol Crespo, Lucía Fuertes Luis, Elvira García Álvarez, Laura García Carrascosa, Natalia Gómez-Gómez, Blanca Herrera Crerizo, Jacinta Landa Goñi, Dario López, Arturo Louro González, María Dolores Martínez-Romero, Juliana Ester Martín-López, Adoración Montejo Ráez, Dimelza Osorio, Isabel Roman Romera, Flavia Salcedo-Fernandez, Anna Selva, Ivan Solà, Helen Strivens, Rafael Torrejón-Cardoso, Mar Trujillo-Martín, and Robin WM Vernooij.

Funding

This project is funded by research grants from the Instituto de Salud Carlos III through the project "PI14/02006" (co-funded by European Regional Development Fund/European Social Fund "Investing in your future"). LMG is funded by a Juan Rodés contract from the Instituto de Salud Carlos III (JR15/00004). PAC is supported by a Miguel Servet investigator contract from the Instituto de Salud Carlos III (CPII15/0034).

Availability of data and materials

The datasets generated and/or analysed during the current study are available from the corresponding author upon request.

Authors' contributions

LMG and PAC contributed to the conception and design of the study. The technical team participated in the data extraction and in the design and running of literature searches and in the data extraction. LMG, HPH, JAS, and PAC are responsible for the writing of the manuscript. All authors participated in the analysis and interpretation of data and contributed by commenting on and editing the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Publisher's Note

Author details

[1]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Sant Antoni Maria Claret 167, 08025 Barcelona, Spain. [2]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain.

Received: 16 March 2017 Accepted: 24 May 2017
Published online: 14 July 2017

References

1. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. J Clin Epidemiol. 2014;67(1):52–5.
2. Martínez García L, Sanabria AJ, Garcia Alvarez E, Trujillo-Martin MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ. 2014;186(16):1211–9.
3. Neuman MD, Goldstein JN, Cirullo MA, et al. Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations. JAMA. 2014;311(20):2092–100.
4. Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. Updating clinical practice recommendations: is it worthwhile and when? Int J Technol Assess Health Care. 2012;28(1):29–35.
5. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001; 286(12):1461–7.
6. Vernooij RW, Sanabria AJ, Sola I, Alonso-Coello P, Martinez GL. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implement Sci. 2014;9:3.
7. Alonso-Coello P, Martinez Garcia L, Carrasco JM, Sola I, Qureshi S, Burgers JS. The updating of clinical practice guidelines: insights from an international survey. Implement Sci. 2011;6:107.
8. Martínez García L, Arevalo-Rodriguez I, Sola I, Haynes RB, Vandvik PO, Alonso-Coello P. Strategies for monitoring and updating clinical practice guidelines: a systematic review. Implement Sci. 2012;7:109.
9. Guidelines International Network. G-I-N Updating Guidelines Working Group. Scotland: Scottish Charity No: SC034047. Available from: http://www.g-i-n.net/working-groups/updating-guidelines.
10. Martínez García L, Pardo-Hernandez H, Superchi C, Niño de Guzman E, Ballesteros M, Ibargoyen Roteta N, et al. Methodological systematic review identified major limitations in prioritisation processes for updating. J Clin Epidemiol. 2017.
11. Martínez García L, Sanabria AJ, Araya I, Lawson J, Solà I, Vernooij RWM, et al. Efficiency of pragmatic search strategies to update clinical guidelines recommendations. BMC Med Res Methodol. 2015;15:57.

12. Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Assessing the need to update prevention guidelines: a comparison of two methods. Int J Qual Health Care. 2004;16(5):399–406.
13. Fochsen G, Zingmark P, Mild M. Living guidelines: A new model for guideline maintenance in Sweden. 13th G-I-N Conference; Philadelphia, USA; 2016.
14. Venhorst K, Ploegmakers M, Holtus A, Van Barnelveld T. What is needed to realize a dynamic process of guideline updating. 11th G-I-N Conference; Melbourne, Australia; 2014.
15. Robbins C, Tom G. Dynamic Updating of Clinical Practice Guidelines (CPGs). 9th G-I-N Conference; Berlin, Germany; 2012.
16. Zelman Lewis S. Sustainable Living Guidelines: A Model for the Future. 8th G-I-N Conference; Seoul, South Korea; 2011.
17. Johnston ME, Brouwers MC, Browman GP. Keeping cancer guidelines current: results of a comprehensive prospective literature monitoring strategy for twenty clinical practice guidelines. Int J Technol Assess Health Care. 2003;19(4):646–55.
18. Totten A, Zelman Lewis S, Patel S, Shorte V, Kolakowshy-Hayner S. Making the Transition to "Living Guidelines": Challenges and Opportunities. 13th G-I-N Conference; Philadelphia, USA; 2016.
19. Grupo de trabajo de la Guía de práctica clínica de atención en el embarazo y puerperio. Guía de práctica clínica de atención en el embarazo y puerperio. Ministerio de Sanidad, Servicios Sociales e Igualdad. Agencia de Evaluación de Tecnologías Sanitarias de Andalucía; 2014. Guías de Práctica Clínica en el SNS: AETSA 2011/10.
20. Martínez García L, McFarlane E, Barnes S, Sanabria AJ, Alonso-Coello P, Alderson P. Updated recommendations: an assessment of NICE clinical guidelines. Implement Sci. 2014;9:72.
21. Vernooij RW, Alonso-Coello P, Brouwers M, Martinez GL. Reporting items for updated clinical guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp). PLoS Med. 2017;14(1):e1002207.
22. Selva A, Sola I, Zhang Y, Pardo-Hernández H, Haynes RB, Martínez García L, Navarro T, Schünemann H, Alonso-Coello P, et al. Development and use of a content search filter for studies on how patients and other stakeholders value health outcomes. Health Qual Life Outcomes. 2017 [accepted].
23. Craig D, Rice S. NHS Economic Evaluation Database Handbook. 3rd ed. York: Centre for Reviews and Dissemination, University of York; 2007.
24. Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. Chest. 2013;144(2):381–9.
25. Uhlig K, Berns JS, Carville S, Chan W, Cheung M, Guyatt GH, et al. Recommendations for kidney disease guideline updating: a report by the KDIGO Methods Committee. Kidney Int. 2016;89(4):753–60.

APPENDIX VII

CheckUp: Spanish translation.

Lista de verificación para evaluar guías de práctica clínica actualizadas (CheckUp)

Ítem	Valoración	Recogido en la página número	Notas
1. La versión actualizada se diferencia con facilidad de la versión original de la guía de práctica clínica.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
2. Incluye la justificación para actualizar la guía de práctica clínica.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
3. Los cambios en el alcance y los objetivos de la versión actualizada con respecto a la original están descritos y justificados.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
4. Describe las secciones revisadas en el proceso de actualización.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
5. Las recomendaciones están claramente presentadas y identificadas como nuevas, modificadas o sin cambios. Las recomendaciones retiradas están identificadas de forma clara.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
6. Describe y justifica los cambios de las recomendaciones.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
7. Menciona a los miembros del grupo de actualización de la guía.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
8. Recoge la declaración sobre los posibles conflictos de interés del grupo de actualización de la guía.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		

9.	El organismo financiador o patrocinador de la versión actualizada figura de forma explícita y detallada.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
10.	Describe los métodos utilizados para la búsqueda e identificación de la nueva evidencia en el proceso de actualización.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
11.	Describe los métodos utilizados para seleccionar la evidencia en el proceso de actualización.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
12.	Describe los métodos utilizados para evaluar la calidad de la evidencia incluida en el proceso de actualización.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
13.	Describe los métodos utilizados para sintetizar la evidencia en el proceso de actualización.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
14.	Describe los métodos para la revisión externa de la versión actualizada.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
15.	Describe los métodos para implementar en la práctica los cambios de la versión actualizada.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		
16.	Describe los métodos para actualizar la nueva versión.	☐ Sí ☐ No ☐ Incierto ☐ No aplicable		

APPENDIX VIII

CheckUp: Chinese translation.

更新临床实践指南需要报告的条目：更新版指南的报告清单

条目	评估	指南中报告相关内容所在的页码	备注
1. 临床指南的更新版本与以前的版本能够被区分开	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
2. 报告了更新临床指南的缘由	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
3. 描述并解释了指南更新版本与以前版本之间范围和目的的变化，并提供了支撑材料	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
4. 描述了在指南更新过程中被修订和审阅过的章节	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
5. 清楚地呈现出指南的推荐意见，并标记为新的、修改的或无变化的；清楚地标注了被删除的推荐意见	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
6. 报告了被更改了的指南推荐意见，并提供了支撑材料	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
7. 报告了参与指南更新的专家组成员	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
8. 记录了负责指南更新版本群组的利益披露	☐ 是的 ☐ 不是 ☐ 不清楚		

	☐ 不适用		
9. 确定和描述了指南更新版本的供资机构所发挥的作用	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
10. 描述了在指南更新中搜索和识别新证据的方法	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
11. 描述了在指南更新过程中用于选择证据的方法	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
12. 描述了在指南更新过程中评估证据质量的方法	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
13. 描述了在指南更新过程中证据合成的方法	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
14. 描述了更新版指南的外部评审方法	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
15. 描述了在实践中实施更新版指南变更的推荐意见的方法和计划	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		
16. 报告了未来更新指南的计划和方法	☐ 是的 ☐ 不是 ☐ 不清楚 ☐ 不适用		

APPENDIX IX

CheckUp: Dutch translation.

CheckUp voor het rapporteren van geactualiseerde richtlijnen (CheckUp)

Item	Beoordeling	Gerapporteerd op paginanummer	Opmerkingen
1. De geactualiseerde versie van de richtlijn kan onderscheiden worden van de vorige versie van de richtlijn.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
2. De reden voor het actualiseren van de richtlijn is aangegeven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
3. Een wijziging in het doel van de geactualiseerde richtlijn ten opzichte van de vorige versie van de richtlijn is beschreven en gerechtvaardigd.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
4. De secties die beoordeeld zijn voor het actualiseren van de richtlijn zijn beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
5. Van de verschillende aanbevelingen in de richtlijn is duidelijk aangegeven welke wel en welke niet zijn gewijzigd tijdens het actualiseren. Tevens zijn verwijderde aanbevelingen duidelijk vermeld.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
6. Wijzigingen in de aanbevelingen van de richtlijn zijn beschreven en verantwoord.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
7. De deelnemers van de werkgroep die betrokken zijn bij de actualisatie van de richtlijn zijn beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
8. Informatie over belangenconflicten van de leden van de werkgroep die verantwoordelijk zijn voor de actualisatie van de richtlijn zijn gerapporteerd.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		

9.	De financierende partijen en hun rol in de actualisatie van de richtlijn is aangegeven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
10.	De methoden die gebruikt zijn voor het zoeken naar nieuwe bewijsvoering in het actualiseren van de richtlijn worden beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
11.	De methoden die gebruikt zijn voor het selecteren van de wetenschappelijke bewijsvoering in het actualiseren van de richtlijn worden beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
12.	De methoden die gebruikt zijn om de kwaliteit van het geïncludeerde data voor de wetenschappelijke bewijsvoering in het actualiseren van de richtlijn te beoordelen worden beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
13.	De methoden die gebruikt zijn voor het synthetiseren van de data voor de wetenschappelijke bewijsvoering in het actualiseren van de richtlijn worden beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
14.	De manier waarop externe beoordeling van de geactualiseerde richtlijn is verlopen wordt beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
15.	Er wordt een plan om de wijzigingen van de geactualiseerde versie in de praktijk te implementeren gepresenteerd.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		
16.	Een concreet plan voor een volgende actualisatie van de richtlijn in de toekomst wordt beschreven.	☐ Ja ☐ Nee ☐ Onduidelijk ☐ Niet van toepassing		