

EVIDENCE INFORMED DECISION MAKING: THE USE OF “COLLOQUIAL EVIDENCE” AT NICE

Tarang Sharma

The Nordic Cochrane Centre-Rigshospitalet, Copenhagen, Denmark
ts@cochrane.dk, tarangs@gmail.com

Moni Choudhury, Bindweep Kaur, Bhash Naidoo, Sarah Garner

National Institute for Health and Care Excellence (NICE), London, United Kingdom

Peter Littlejohns

King's College, London, Primary Care and Public Health Sciences, London, United Kingdom

Sophie Staniszewska

University of Warwick, RCN Research Institute, Warwick Medical School, Coventry, United Kingdom

Objectives: Colloquial evidence (CE) has been described as the informal evidence that helps provide context to other forms of evidence in guidance development. Despite challenges around quality, and the potential biases, the use of CE is becoming increasingly important in assessments where scientific literature is sparse and to also capture the experience of all stakeholders in discussions, including that of experts and patients. We aimed to ascertain how CE was being used at the National Institute for Health and Care Excellence (NICE).

Methods: Relevant data corresponding to the use of CE was extracted from all NICE technical and process manuals by two reviewers and quality assured and analyzed by a third reviewer. This was considered in light of the results of a focused literature review and a combined checklist for quality assessment was developed.

Results: At NICE, CE is utilised across all guidance producing programmes and at all stages of development. CE could range from information from experts and patient/carers, grey literature (including evidence from websites and policy reports) and testimony from stakeholders through consultation. Six tools for critical appraisal of CE were available from the literature and a combined best practice checklist has been proposed.

Conclusions: As decisions often need to be made in areas where there is a lack of published scientific evidence, CE is employed. Therefore to ensure its appropriateness the development of a validated CE data quality check-list to assist decision makers is essential and further research in this area is a priority.

Keywords: colloquial evidence, checklist, evidence-based healthcare, decision making, guidelines, NICE

It is widely agreed that high quality evidence should underpin all healthcare guidance (1), yet there is debate about what constitutes as evidence. Almost always the scientific evidence is not complete, or does not address areas of importance for clinicians or patients to allow for a decision to be made solely on such evidence alone, necessitating the input from other sources within a deliberative process (2). There is often a need to contextualize this evidence and to understand how it should be implemented in healthcare practice, as without contextualization, guidance and policies may fail to produce the desired results. This shift from evidence-based to evidence-informed decision making has been reflected in the definition of evidence and methodological practices of leading guidance producing organizations such as the Health Evidence Network (HEN) of the World Health Organization that define evidence as “*findings from research and other knowledge that may serve as a useful basis for decision making in public health and health care*” (3).

Evidence has been conceptualized in a range of ways and one approach developed by Lomas and colleagues has described how three forms of evidence are used within a deliberative process for healthcare decision making, namely: “*scientific evidence on effectiveness, scientific evidence on context and colloquial evidence*” (3;4).

- *Scientific context-free evidence:* evidence that helps determine the potential benefit/ efficacy and safety of the health technology. This is likely to be universal and not subjective to specific geographical scenarios. The authors primarily refer to this being evidence from good quality randomized

controlled trials (RCTs) (3). One can argue, however, whether context-free evidence can truly exist. This is especially true with respect to the choice of comparator, as often the standard care or usual treatment differs between countries and, therefore, the geographical scenarios plays an important role in setting the context of any treatment alternatives. In addition, these RCTs may not include outcomes that are of importance to patients (5).

- *Scientific context-sensitive evidence:* evidence specific to particular real world scenarios and is less likely to be generalizable (3). It could be argued that to ensure the implementation of recommendations based on the best RCT evidence, some contextual evidence is needed to ensure effectiveness.
- *Colloquial evidence (CE):* evidence that helps support, supplement or refute the scientific evidence and is often used to augment an evidence landscape. CE is an umbrella term and consists of different types of data including informal expert opinion from clinicians and/ or patients, their views and narratives, electronic data from Web sites, policy documents, and other reports etc (3). “Colloquial” has been defined by the Oxford English dictionary as “*used in ordinary or familiar conversation; not formal or literary*” and its synonyms include “*informal*”, “*unofficial*”, and “*popular*” (6), and, therefore, can be understood in this context as informal evidence. It could be argued that CE should not be considered as evidence, as it may not be collected in a rigorous or systematic manner. However, we would assert that it is not appropriate to conceptualize CE from a research perspective in this way, as it has a different role, attributes, characteristics, and contribution. It has been suggested that CE should be understood as the additional “*knowledge*” or “*factors*” alongside scientific evidence, which is considered in a deliberative process (7).

The key issue for decision makers is to balance these diverse evidence types together, assess the weight to place on each, and more importantly let each specific type of evidence contribute appropriately to the final decision (2).

The National Institute for Health and Care Excellence (NICE) is an independent organization and is responsible for providing evidence-based guidance on the most effective ways to diagnose, treat and prevent disease and ill health for England and Wales (8;9). The Institute moved from considering evidence in the traditional hierarchies, to using the appropriate evidence for the question posed and uses evidence (including clinical, economic, and patient-based) from clinical trials, observational and qualitative studies (10;11). CE often takes the form of expressions of opinion by experts based on their practical experience or professional judgment (4;12). The deliberation of evidence by the advisory bodies at NICE is a systematic but discursive qualitative process that not only considers CE but also acts as a direct source of CE itself. The deliberative process at NICE has been defined as the “*careful, deliberate consideration and discussion of the advantages and disadvantages of various options*” used to elicit and combine evidence (2). The process can range from an explicit algorithmic formal quantitative method such as multi-criteria decision analysis (MCDA), to a very informal discursive process (2;12;13). It has been said that deliberative processes give a sense of valuing and weighting the different forms of evidence and help their systematic and explicit combination together (4). Therefore, CE plays a vital role in facilitating the creation of guidelines from evidence in a form that can inform practice, although this is rarely acknowledged.

NICE not only considers the appropriate evidence but undertakes a critical appraisal of the quality of the evidence to underpin its recommendations. Critical appraisal has been defined as “*a systematic process used to identify the strengths and weaknesses of a research article*” which is used to determine the utility and validity of the evidence being considered (14). As it stands, the definition is limited to published scientific research but one can argue that such a critical eye is needed even more when considering CE. Some critical appraisal tools for considering the quality of CE have been identified by a recent review (15) and are summarized in Table 1.

To better understand the role of CE in NICE’s processes, the Institute undertook this study to build upon preliminary work done by Reay et al. (15). The aims of this project were to understand the types of CE used across the guidance producing teams at NICE; the extent of its use and how it had been incorporated within their deliberative processes. The main areas for exploration were to identify the sources and nature of CE at NICE, whether there were any inconsistencies across the teams in their use of such evidence, any definitions used and where variations existed, to investigate whether these were warranted due to the nature of the different programs.

METHODS

In this study, CE was defined as the “*evidence about resources, opinion, political judgment, values, culture, and the particular pragmatics of a situation*” that is used to complement the scien-

tific evidence as originally defined by Lomas and colleagues (3). This definition includes information that comes informally from different stakeholders including practitioners and patients and is also consistent with its use in other such studies (15;25;26).

All of NICE’s technical methods and process manuals (8) were considered independently by one of two reviewers from August 2011 to March 2012. Any text relating to CE was systematically searched for and identified using search terms like “advice” or “colloquial” or “comment” or “consultation” or “database” or “electronic*” or “expert” or “grey literature” or “informal” or “narrative” or “patient” or “policy” or “professional” or “public” or “report” or “specialist” or “testimony” or “views” or “web*” etc, based on the search criteria developed by Reay et al. (15). The relevant sections were extracted verbatim using data extraction forms designed to capture guidance on the use of CE (i.e., the source of CE; methods of collection of CE and its purpose). The raw data was then quality assured by a third independent reviewer who resolved any disagreements by discussion and consensus and analyzed and summarized the data. The results were also assessed for whether any quality assessment or formal critical appraisal tools were used.

RESULTS

The guidance development processes at NICE can be broadly grouped into the five stages namely scoping, evidence review and economic modeling, deliberative process, stakeholder consultation and implementation tool development, and CE is used throughout these stages.

Mapping of the use of CE over the development of guidance is summarized in Figure 1.

The following sources of CE were identified: (i) CE1: Evidence from experts (professionals/ clinicians) and patients/carers; (ii) CE2: Evidence from grey literature; (iii) CE3: Evidence from all stakeholders through public consultation

CE1: Evidence from Experts and Patients/Carers

This type of CE is used in a varying degree and at different time-points by the different teams at NICE. This can be done directly in person or indirectly through written advice or testimonies. All programs use the evidence from experts (such as clinicians or other NHS professionals) and patients/ carers (who are experts in relation to their own condition), from an early stage in their scoping processes to help inform their research questions. Advice is sought directly through stakeholder workshops and/or indirectly through development of briefing notes or papers.

In certain cases, such as the Evidence Review Groups of the Technology Appraisals (TA) program, the use of specialist advisors (clinicians, NHS commissioning experts, or patients) to help understand and interpret the clinical evidence throughout the period of development can be undertaken. Expert elicitation methods can also be used to obtain data for certain parameters of economic models, when it is unavailable through the

Table 1. List of Studies Using Critical Appraisal Techniques or Grading of Colloquial Evidence as Identified from the Review by Reay et al (2010) (15)

| Study | Critical appraisal tools and grading of Colloquial evidence |
|-----------------------------|---|
| Benzies et al (2006) (16) | <p>This study included CE evidence and appraised it alongside other forms of evidence. It considered '<i>opinions of respected authorities or that of an expert committee as indicated in published conferences or guidelines; or of those individuals who have knowledge in one particular field and are applying that knowledge to another field; or summaries of the collective wisdom or experience of others in the field; or of those individuals who have written and reviewed the guidelines, based on their experience, knowledge or the relevant literature, and discussion with their peers</i>'.</p> <p>The tool considered CE within the hierarchy of evidence model as lower levels (levels IV and V) after experimental evidence (levels I-III).</p> <p>Level I - Randomised controlled trials (RCTs) or meta-analysis of such trials, of adequate size and power.</p> <p>Level II - RCTs that are too small to provide Level I evidence</p> <p>Level III - Non-randomised controlled trials or cohort studies, case series, case controlled studies or cross-sectional studies</p> <p>Level IV – Opinion of respected authorities or that of an expert committee as indicated in published conferences or guidelines</p> <p>Level Va - Opinion of those individuals who have knowledge in one particular field and are applying that knowledge to another field; or summaries of the collective wisdom or experience of others in the field</p> <p>Level Vb - Opinion of those individuals who have written and reviewed the guidelines, based on their experience, knowledge of the relevant literature, and discussion with their peers</p> |
| Coad et al (2006) (17) | <p>This study discussed a tool for the critical appraisal of grey literature. The tool suggested that when using any grey literature one should be prepared to carefully evaluate three key areas:</p> <ul style="list-style-type: none"> - when it was written - by whom and - its purpose |
| Haig and Dozier (2003) (18) | <p>This study discusses a criterion for appraising the quality of the evidence obtained from websites and all electronic sources by Wilson (2002) (19) by extending to other forms of CE.</p> <p>They suggest five key domains for the reviewers to critically question:</p> <ul style="list-style-type: none"> - Authority - accuracy - currency - scope - objectivity |
| Haig and Dozier (2003) (20) | <p>The authors discuss an adaption of the DISCERN tool for use in appraising health information in general. The DISCERN tool (21) (developed by University of Oxford and the British Library) is aimed at consumers and information providers to enable them to judge the quality of information on treatment choices. The tool asks the reviewer to consider sixteen questions, all with a rating from 1 or 'No' to 5 'Yes' on a 5-point scale. The authors suggest that though the questions themselves may need amending, but the principles and methods may have some use in appraising CE and some questions that maybe useful are:</p> <ul style="list-style-type: none"> - Are the aims clear? Does it achieve its aims? - Is it relevant? - Is it clear what sources of information were used to compile the publication? - Is it clear when the information used or reported in the publication was produced? - Is it balanced and unbiased? - Does it provide details of additional sources of support and information? - Does it refer to areas of uncertainty? |
| Rundall et al (2007) (22) | <p>This study discusses a criterion for evaluating evidence which they feel is a critical step in the decision-making process and describe a set of key areas that reviewers should use to appraise information:</p> <ul style="list-style-type: none"> - accurate - applicable - actionable - accessible |

Table 1. Continued.

| Study | Critical appraisal tools and grading of Colloquial evidence |
|---------------------------|---|
| Shpilko et al (2005) (23) | <p>The study discusses the criterion for evaluating health-related information from websites using the format from the National Network of Libraries of Medicine (24) with the domains of:</p> <ul style="list-style-type: none"> - accuracy - authority - bias/ objectivity - currency/ timeliness - coverage |

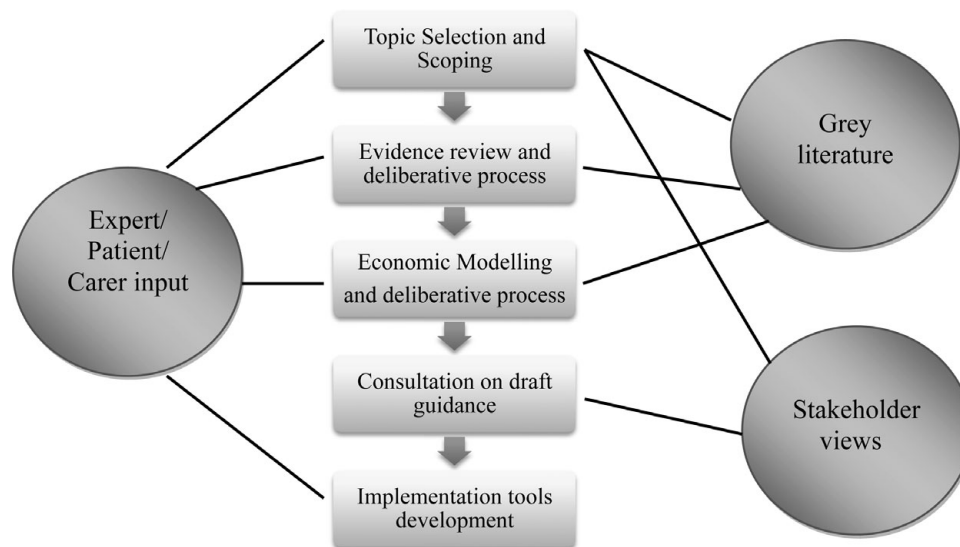


Figure 1. Use of colloquial evidence (CE) in guidance development at NICE.

published scientific literature, to help construct a plausible pathway of care for modeling. These “consultees” give advice on various stages of the guidance development process, starting with the draft scope. The Consultee organizations (nonmanufacturers) nominate patient experts or clinical specialists while the manufacturer or sponsor consultees can only nominate clinical specialists, who then help feed into the process. A similar case is seen in other programs such as the Interventional Procedures (IP), Medical Technologies Evaluation Program (MTEP) and Diagnostics Assessment Program (DAP). Specifically in the IP program a “commentary” from the specialist advisor can also be produced which summarizes their “*opinion, and/or information about an interventional procedure, or to the peer-reviewed literature relating to the interventional procedure of interest*”.

In the Clinical Guidelines (CG) program, advice from clinical experts and patients/ carers is obtained throughout the guidance development process through guidance development groups (GDGs) that include health professionals and patient/carers representatives with relevant expertise and experi-

ence of the specific guideline topic at hand. In the Public Health (PH) program “expert testimonies” can once again be used as evidence, which are defined as “*short papers (with references to any relevant published work)*” that reflect opinions of certain experts in the field when there are either significant gaps in the evidence, significant conflicts amongst available evidence, or in instances where there is the need for the “*views and experiences of specific groups*”. Some PH guidance can also undertake primary research as “field work” to inform practice and test the feasibility for implementation of their draft recommendations with “*policy makers, commissioners and practitioners (including members of the community, volunteers, parents, and carers as well as professionals such as GPs, nurses, and teachers)*”.

All programs have either a NICE standing committee or a temporary advisory body such as a GDG that considers the evidence in a deliberative process that further generates CE through its deliberation. A deliberative process has been defined as a process that “*provides guidance informed by relevant scientific evidence, interpreted in a relevant context*”.

wherever possible with context-sensitive scientific evidence and, where not, by the best available colloquial evidence” (2). The standing committees have a general expertise and are not specialists in the condition of interest and so often get specialists, professionals, relevant commissioners and patients/ carers to participate in the deliberation process by presenting their views at committee meetings. The GDGs for the CG program have topic specific membership but can still have further co-opted experts if required for the deliberative processes. These deliberations are summarized within the “considerations” or “evidence to recommendations” sections (depending on the program) of the final guidance and act as a primary direct source of CE.

Additionally, all guidance production at NICE follows the patient and public involvement (PPIP) policy, which sets the platform for the contribution of lay people, and organizations representing their interests, to the work of NICE. This enhances the NICE guidance, giving them a greater patient, carer, or community focus and relevance (27). Finally, NICE’s Citizens Council (group of 30 ordinary members of the public, representing the country) also have their views captured through reports that feed into the methods and processes across the Institute. Even though the decisions reached by the Citizens Council do not directly affect any individual piece of guidance; their views are responsible for ensuring the “*public perspective on overarching moral and ethical issues*” and can be considered as another direct source of CE (28).

CE2: Evidence through Grey Literature and Web Sites

The term “grey literature” was first understood in 1978, with the creation of the “System for Information on Grey Literature” database, now “OpenGrey” managed by the European Association for Grey Literature Exploitation (EAGLE) and partners. They now define it as “*information produced on all levels of government, academia, business and industry in electronic and print formats not controlled by commercial publishing, that is, where publishing is not the primary activity of the producing body*”. Examples include conference abstracts, research reports, unpublished data, dissertations, policy documents, personal correspondence, electronic publications, online publications, online resources, open access research, ePrints, digital documents, and so on (29;30). At NICE, some types of grey literature are also commonly used to inform guidance and act as a secondary indirect source of evidence.

Across the programs, if data for all the parameters of an economic model is not available through the scientific published literature, CE in the form of data from various electronic sources can also be used. Data from the Health and Social Care Information Center Web site for data on Hospital Episodes Statistics (HES) and Healthcare Resource Groups (HRG) is commonly used for this purpose.

In the CG program, evidence from Web sites (such as “health talk online”) is also routinely used, where Web sites

are searched manually for any additional relevant patient experience information. An example is the clinical guideline for Motor Neurone Disease (CG105), where to get a complete picture of the information and support needs for patients and carers, alongside the evidence from eleven published studies, narratives from interviews taken directly from the health talk online Web site were also considered (31).

Grey literature such as policy documents and local council reports are used actively in the PH program, for the formal “mapping of local practice” to give a snapshot of current practices or policy scenarios. Raw data from registries is often used in the IP program and this is similar for the DAP and MTEP programs due to the lack of robust trial data in the areas of procedures, diagnostics and devices. Proponents of real world data would, however, argue that evidence from routine sources such as databases and audits should be considered valid sources of scientific information (32) rather than informal or CE.

CE3: Evidence from All Stakeholders through Public Consultation

All programs have a public consultation of their draft scope, and draft guidance, where registered stakeholders (which include professional groups and societies, patient groups and charities, other NICE staff, the NHS etc) and industry are able to comment and submit their views on the questions (draft scope) or recommendations (draft guidance) proposed.

The technical teams present these comments to the standing committees or temporary advisory bodies, who then consider them through their deliberations, for the final document, to ensure that all comments have been taken into consideration. Therefore, the stakeholder comments can have a direct impact on the final recommendations of any guidance.

Critical Appraisal of Colloquial Evidence

All programs use formal critical appraisal techniques for considering scientific evidence but none had an explicit appraisal checklist for reviewing CE. The mechanism of appraisal of CE at NICE was informal and through deliberative consideration and through stakeholder consultation of those considerations. It could be argued that formal checklists may be less helpful with CE, as it is conceptually different. There may, however, be the need for development of some form of evaluation of CE and its contribution to the deliberative process.

DISCUSSION

The study identified that different forms of evidence including CE were used in different ways and for different reasons for guidance development at NICE. Although each piece of NICE guidance is developed following different processes depending on the nature of guidance program, all these processes share common features based on the key procedural principles – “*scientific rigor, inclusiveness, transparency, independency, challenge, review, support for implementation, and timeliness*”

(8). The process of developing guidance also draws on different forms of evidence, which this study acknowledges and attempts to understand. As such, we hope it provides an important contribution to the understanding of CE and its use at NICE.

All guidance producing programs at NICE aim to use the best available evidence to inform their decisions of clinical and cost effectiveness, but these judgments often require different forms of evidence including CE, to support the gaps in the scientific evidence and to help contextualize the scientific evidence that is available. CE could be considered vital in understanding the implications and utility of any guidance.

On the whole, three broad sources of CE were identified at NICE: evidence from experts and patients/ carers, evidence from grey literature and evidence from all stakeholders through public consultation. There is some variability across programs to the degree of their CE use and the manner in which it is considered in a deliberative process. Some differences are warranted and dependent on the nature of the program. For example, you would expect a RCT to be the foundation of a TA recommendation because the process considers drugs that have been through a regulatory body and, therefore, by definition are supported by RCT evidence of effectiveness. However, RCT evidence may be unavailable or inappropriate for other questions raised by for example a PH guidance, where the question may be less focused on the specifics of a particular intervention, and concerned more with structural complex issues, for which an RCT design may not adequately address the problem. In such cases CE is often needed routinely through expert testimonies or policy analysis or fieldwork to develop recommendations that are appropriate. One may argue that CE may also exert influence at other levels of the knowledge production process in a latent oblique manner. Therefore, there is also a possibility of a “type 2” error that could be impossible to measure. It could also be that CE and other traditionally “lower” forms of evidence, act as the first seeds of what will eventually become more robust forms of evidence and are actually part of the natural process of maturation of evidence.

It could also be that at times the three different types of evidence overlap and interact with each other. For example expert opinion could be based on knowledge of credible scientific evidence and at other times, in the absence of good external evidence, be limited to biased personal opinions. There could also be circumstances where credible objective scientific evidence may exist but be over-ruled by poorly quantified and biased personal opinions, due to personal agendas and beliefs. The hope at NICE is that by having standing committees that hear arguments from both sides, or topic specific advisory groups that consist of members with opposing professional viewpoints, that individual biases are minimized. Therefore, CE like any other type of evidence can be of varying quality and with certain level of uncertainty associated with it. So, therefore, maybe CE should also be critically examined before its inclusion into any decision-making model.

Though some critical appraisal tools for assessing the quality of CE are available from the literature (15), our study found that no formal method was being used at NICE. The use of such critical appraisal tools would allow all the evidence both scientific and CE that is presented before the independent bodies, to have gone through a similar rigorous process of critical appraisal. The toolkits available from the literature identified by Reay and colleagues (15) range from adaptations of standard levels of evidence hierarchies, to specific checklists for the different types of CE, with key questions to be posed to the evidence before incorporating them into a decision model. They do however, share some common features that could be considered in a systematic manner, in a proposed “SART” system, where a reviewer could score either “yes”, “no”, or “unclear” using the standard Cochrane critical appraisal system (33) as described in Table 2.

The SART instrument is, however, still very much a crude tool and needs to be developed further and tested for robustness. To create a formal checklist a thorough process such as an expert elicitation Delphi Panel or other such formal consensus methodology, should be undertaken. Moreover, as these concepts are solely derived from the limited literature identified by one scoping literature review, this may not represent all the factors that need to be considered and so, therefore, any interpretations should be made very cautiously.

For this project, we used the conceptual framework developed by Lomas and colleagues (3), however, other frameworks and approaches of understanding evidenced- informed decision making have been developed and are also widely used by policy makers. One such tool, the SUPPORT tool *describes processes to help ensure that relevant research is identified, appraised and used appropriately to inform health policy making* and also considers a wider evidence base that includes CE, and could have also been a useful framework for analyzing this research (34).

With NICE now developing quality standards for social care and the introduction of value-based pricing for new pharmaceuticals for the technology appraisals program, further changes are expected to their methods and processes (35). The three programs of technology appraisals, clinical guidelines and public health at NICE also recently went through a process of updating their technical manuals in 2012 and further work is being done to harmonize the processes and methods for public health, clinical guidelines, and the new social care program. This will have a direct effect on the type of evidence used and the way it is used in their decision-making processes in the future. These changes also may have an impact on how CE will be used at the Institute going forward, but one can expect it to only grow in importance in the coming years. A focus on developing our understanding of its contribution will be an important future research area.

Table 2. Proposed Key Areas of Questioning for Critical Appraisal of Colloquial Evidence Using the “SART” System

| Area of critical appraisal | Detailed questions | Yes | Unclear | No | Comments | |
|----------------------------|--------------------|--|---------|----|----------|--|
| S | Source | <p>Is the source of the CE credible?</p> <ul style="list-style-type: none"> - by whom it was written/ spoken and its purpose - determine to what extent one could trust the author(s) - are there certain conflicts of interests to be noted - did it appear to be objective or were there any potential biases present - who was the sponsor/ funder of the source - if the information only presented one point of view | | | | |
| A | Accuracy | <p>Is the information accurate?</p> <ul style="list-style-type: none"> - consider how plausible are the claims, in terms of whether they could be verified by other sources or testimonies - if they contain any footnotes or state any references that can be checked and how reliable are the sources cited - are the aims available and information presented clearly - are the methods used for collecting that information transparent and clearly presented - are there any inconsistencies present in the content - does it highlight any areas of uncertainty | | | | |
| R | Relevance | <p>Is the evidence with respect to the scope of the guidance, relevant?</p> <ul style="list-style-type: none"> - consider how applicable the information really is to the question at hand - is all the information present or it seems to be incomplete with respect to your question (is it discussing only part of the equation) - is the setting and context clearly defined and how well does it fit with the research question | | | | |
| T | Timeliness | <p>How timely is the information?</p> <ul style="list-style-type: none"> - how current is the information - are any dates associated with the information (the last date when the website was updated if an electronic source or recent or old experiences if testimonies etc) - how regularly is the information updated - Is there any more recent research or information available | | | | |

CONCLUSION

CE's role is vital in shaping and influencing guidance development and in many respects it provides the architecture within which specific forms of clinical, economic and patient-based evidence are considered together (36).

Although this piece of scoping work has several limitations, as it is only a snapshot of the scenario at NICE in 2012 and only considered limited literature on the subject, some useful conclusions can be drawn from it. There is a need to have a clear definition of CE with explicit information of what should and should not be included within the umbrella term. For this project, another weakness was that we kept to the definition of CE as described by Lomas and colleagues (3), but it ideally should be debated and deliberated more widely so a common definition amongst different practitioners is agreed. More work is also needed to identify how it complements and fits with other forms of evidence as there is a lack of appropriate methodology for integrating the different types of evidence together.

However, it is clear that while CE is used throughout the guidance development process at NICE, no formal consistent mechanism of appraisal was used. There are different appraisal tools available in the literature and information could be taken from them to develop a single checklist for use, after a full and comprehensive systematic review of the literature. A great deal of further research is needed to develop and validate any single checklist proposed before it could be routinely used for guidance development. Policy changes in the remit of NICE's role may also mean that CE could play a more influential role in the future.

CONFLICTS OF INTEREST

This is to attest that all named authors have contributed towards the conception, design and interpretation of the study and the writing of the manuscript. All have approved the final version being submitted; and the content has not been published nor

is being considered for publication elsewhere. Preliminary results were previously presented as an oral presentation at the HTAi conference in 2012 in Bilbao, Spain: *Gaceta Sanitaria* 26 (Espec Congr 2):69 and a sub-set of the complete results related to clinical guidelines were presented as a poster at the G-I-N conference in 2012 in Berlin. This project did not receive any source of funding. The authors understand the terms of the conflict of interests and do not have any to declare. PL and TS were employees of NICE at the time of the study and BK, BN, MC and SG are currently employed by NICE and were so at the time of the study. SS had advised NICE as a patient expert and has previously chaired a NICE clinical guideline.

REFERENCES

- Rycroft-Malone J, Seers K, Titchen A, et al. What counts as evidence in evidence-based practice? *J Adv Nurs*. 2004;47:81-90.
- Culyer AJ, Lomas J. Deliberative processes and evidence-informed decision making in healthcare: Do they work and how might we know? *Evid Policy*. 2006;2:357-371.
- Lomas J, Culyer AJ, McCutcheon C, McAuley L, Law S. *Conceptualizing and combining evidence for health system guidance*. Ottawa: Canadian Health Services Research Foundation; 2005.
- Dobrow MJ, Chafe R, Burchett HED, Culyer AJ, Lemieux-Charles L. *Designing deliberative methods for combining heterogeneous evidence: A systematic review and qualitative scan. A report to the Canadian Health Services Research Foundation*. Ottawa: Canadian Health Services Research Foundation; 2009.
- Staniszewska S, Brett J, Mockford C, Barber R. The GRIPP checklist: Strengthening the quality of patient and public involvement reporting in research. *Int J Technol Assess Health Care*. 2011;27:391-399.
- Oxford dictionaries online. <http://www.oxforddictionaries.com/> (accessed August 22, 2014).
- Canadian Health Services Research Foundation. *Weighting up the evidence: Making evidence informed guidance accurate, achievable, and acceptable*. A summary of the workshop held on September 29, 2005. Ottawa: Canadian Health Services Research Foundation; 2006.
- National Institute for Health and Care Excellence (NICE). <http://www.nice.org.uk/> (accessed September 12, 2013).
- Sharma T, Doyle N, Garner S, Naidoo B, Littlejohns P. NICE supporting England and Wales through times of change. *Eurohealth*. 2011;17:30-31.
- Littlejohns P, Chalkidou K, Wyatt J, Pearson SD. Assessing evidence and prioritizing clinical and public health guidance recommendations: The NICE way. In: Killoran A, Kelly MP, eds. *Evidence-based public health*. Oxford: Oxford University Press; 2009.
- Staniszewska S, Boardman F, Gunn L, et al. The Warwick Patient Experiences Framework: Patient-based evidence in clinical guidelines. *Int J Qual Health Care*. 2014;26:151-157.
- Culyer AJ. *Deliberative processes in decisions about health care technologies: Combining different types of evidence, values, algorithms and people*. London: Office of Health Economics; 2009.
- Littlejohns P, Sharma T, Jeong K. Social values and health priority setting in England: "Values" based decision making. *J Health Organ Manag*. 2012;26:363-373.
- Young JM, Solomon MJ. How to critically appraise an article. *Nat Clin Pract Gastroenterol Hepatol*. 2009;6:82-91.
- Reay CA, Colechin ES, Bousfield DR, Sims AJ. *Review of published literature relating to methods for identifying, synthesis and integration of colloquial evidence*. Regional Medical Physics Department Newcastle upon Tyne Hospitals NHS Foundation Trust, Final Report RSGT405. 2010.
- Benzies KM, Premji S, Hayden KA, Serrett K. State-of-the-evidence reviews: Advantages and challenges of including grey literature. *Worldviews Evid Based Nurs*. 2006;3:55-61.
- Coad J, Hardicre J, Devitt P. How to search for and use 'grey literature' in research. *Nurs Times*. 2006;102:35-36.
- Haig A, Dozier M. BEME Guide no 3: Systematic searching for evidence in medical education—Part 1: Sources of information. *Med Teach*. 2003;25:352-363.
- Wilson PR. How to find the good and avoid the bad or ugly: A short guide to tools for rating quality of health information on the Internet. *BMJ*. 2002;324:598-602.
- Haig A, Dozier M. BEME Guide no 3: Systematic searching for evidence in medical education—Part 2: Constructing searches. *Med. Teach*. 2003;25:463-484.
- The DISCERN Instrument. <http://www.discern.org.uk/index.php> (accessed October 7, 2014).
- Rundall TG, Martelli PF, Arroyo L, et al. The informed decisions toolbox: Tools for knowledge transfer and performance improvement. *J Healthc Manag*. 2007;52:325-341.
- Shpilko I. Locating grey literature on communication disorders. *Med Ref Serv Q*. 2005;24:67-80.
- National Network of Libraries of Medicine. *Evaluating health websites*. <http://nnlm.gov/outreach/consumer/evalsite.html> (accessed October 7, 2014).
- Fournier M. Knowledge mobilization in the context of health technology assessment: An exploratory case study. *Health Res Policy Syst*. 2012;10:10.
- Watt A, Hiller J, Braunack-Mayer A, et al. The ASTUTE Health study protocol: Deliberative stakeholder engagements to inform implementation approaches to healthcare disinvestment. *Implement Sci*. 2012;7:101.
- National Institute for Health and Care Excellence (NICE). *Patient and public involvement policy*. http://www.nice.org.uk/media/B41/22/NICE_PPIP_policy_-_final_for_PDF_-_June_2011.pdf (accessed September 17, 2013).
- National Institute for Health and Care Excellence (NICE). *Citizen's council*. http://www.nice.org.uk/aboutnice/howwework/citizenscouncil/citizens_council.jsp (accessed September 17, 2013).
- Hopewell S, McDonald S, Clarke MJ, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev* 2007;2:MR000010.
- Lawrence A. Electronic documents in a print world: Grey literature and the internet. *Media Int Aust*. 2012;143:122-131.
- National Institute for Health and Care Excellence (NICE). *Motor neurone disease: The use of non-invasive ventilation in the management of motor neurone disease*. London: National Institute for Health and Clinical Excellence; 2010 July. Report No: NICE CG105.
- Garrison LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: The ISPOR Real-World Data Task Force Report. *Value Health*. 2007;10:326-335.
- Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.cochrane-handbook.org (accessed August 22, 2013).
- Lavis J, Oxman A, Lewin S, Fretheim A. SUPPORT Tools for evidence-informed health Policymaking (STP) 9: Assessing the applicability

- of the findings of a systematic review. *Health Res Policy Syst.* 2009;16(Suppl 1):S9.
35. Department of Health. *Equity and Excellence: Liberating the NHS.* http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_117353 London. (accessed 12 September 2013).
36. Staniszewska S, Crowe S, Badenoch D, et al. The PRIME project: Developing a patient evidence-base. *Health Expect.* 2010;13:312-322.