ORIGINAL PAPER

From "getting things right" to "getting things right now": Developing COVID-19 guidance under time pressure and knowledge uncertainty

Marjolein Moleman MSc[1] | Fergus Macbeth Dr[2] | Sietse Wieringa Dr[3] | Frode Forland Dr[4,5] | Beth Shaw MSc[6] | Teun Zuiderent-Jerak Dr[1]

[1]Athena Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[2]Centre for Trials Research, Cardiff University, Cardiff, UK

[3]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

[4]Division for Infectious Diseases and Environmental Health, Norwegian Institute of Public Health, Oslo, Norway

[5]Institute for Community Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway

[6]Oregon Health & Science University, Portland, Oregon, USA

Correspondence
Teun Zuiderent-Jerak, VU University Amsterdam, Faculty of Science, Athena Institute, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.
Email: teun.zuiderent-jerak@vu.nl

Abstract

Background: At the start of the COVID-19 pandemic, guidance was needed more than ever to direct frontline healthcare and national containment strategies. Rigorous guidance based on robust research was compromised by the emergence of the pandemic and the urgency of need for guidance. Rather than aiming to "get guidance right", guidance developers needed to "get guidance right now".

Aim: To examine how guidance developers have responded to the need for credible guidance at the start of the COVID-19 pandemic.

Methods: An exploratory mixed-methods study was conducted among guidance developers. A web-based survey and follow-up interviews were used to examine the most pertinent challenges in developing COVID-19 guidance, strategies used to address these, and perspectives on the implications of the COVID-19 pandemic on future guidance development.

Results: The survey was completed by 46 guidance developers. Survey findings showed that conventional methods of guidance development were largely unsuited for COVID-19 guidance, with 80% ($n = 37$) of respondents resorting to other methods. From the survey and five follow-up interviews, two themes were identified to bolster the credibility of guidance in a setting of extreme uncertainty: (1) strengthening end-user involvement and (2) conjoining evidence review and recommendation formulation. 70% ($n = 32$) of survey respondents foresaw possible changes in future guidance production, most notably shortening development time, by reconsidering how to balance between rigour and speed for different types of questions.

Conclusion: "Getting guidance right" and "getting guidance right now" are not opposites, rather uncertainties are always part of guidance development and require guidance developers to balance scientific robustness with usability, acceptability, adequacy and contingency. This crisis points to the need to acknowledge uncertainties of scientific evidence more explicitly and points to mechanisms to live with such uncertainty, thus extending guidance development methods and processes more widely.

2 | WILEY— Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

MOLEMAN ET AL.

1 | INTRODUCTION

Throughout the COVID-19 pandemic, there has been a pressing need for reliable guidance, ranging from clinical guidance for frontline healthcare to public health guidance for national containment strategies. Guidance was needed to direct clinical care, ensure fair distribution of resources[1,2] and support public adherence to government policies.[3] Whereas guidance is often produced over a longer period of time,* guidance developers were faced with the task of evaluating a fleeting and complex body of knowledge within weeks or days.[4,5] This posed a challenge to guidance developing bodies (e.g., WHO, NICE, SAGE, NIPH, CDC) worldwide, since methods for developing guidance have increasingly emphasized basing recommendations on rigorously appraised[6] frequency-type evidence (such as systematic reviews of randomized controlled trials).[7] Frequency-type reasoning builds on the idea that events (e.g., response to a certain intervention) we *often* observe give the best prediction for future events, which would make frequently observed events count as the "best research evidence" to make causal inferences regarding the intervention in question. The prioritization of frequency-type reasoning no longer worked in the setting of this pandemic; a time of great uncertainty and urgency.[7] Due to epistemic uncertainty, in which knowledge or understanding is insufficient (see Reference [8], p. 507), prevailing grading methods such as GRADE that predominantly focus on evidence that is statistically reliable and privileges high internal validity,[9] often result in weak recommendations for strongly pressing issues.† Guidance institutions tried to streamline guidance development but found the emergence of a new disease and the urgency of need compromised the requirement for basing recommendations on prevailing standards of robust "high quality" research evidence.[10] Instead, guidance developers needed to rely on other types of knowledge,[11] including observational studies,[11] mechanistic studies[5,12] and indirect evidence on the effects on transmission of other viruses in non-pandemic conditions.[13] What is more, the urgency for guidance meant that developers could no longer aim to "get things right" but had to "get things right now". The style of reasoning became one of "taking care while the uncertain future unfolds" (see Reference [7], p. 89). Guidance production in this context could now better be described as "know-now": a form of knowing "used to interpret new situations, to establish what might be the problem, and how [one] could act" (see Reference [14], p. 80).

Greenhalgh notes that the logic of evidence-based medicine remains useful in some aspects of outbreak guidance (e.g., studying drug and vaccine efficacy), but that additional methods are needed on how to manage epistemic uncertainty, the unpredictability of events,[15] but also, the incorporation of patient's and public's values in the formulation of recommendations.‡ Whereas "political decision-making seems to have become technocratic backed by scientific signals" (see Reference [16], p. 614), it has shown that simply "following the science"

becomes problematic when it neglects other areas of science (e.g., economics, psychology, sociology, behavioural science) and the diversity of impacts on patients and society.[17] Whereas EBM provides detailed methods for appraising and including population-level research in guidance development, too narrow a focus on population-based research devalues the mechanistic and experiential knowledge[8]—which are often among the few sources of knowledge in the early phases of a health crisis. In contrast to the fine-grained assessment tools that exist for evidence that is high up in the frequentist "evidence hierarchy", the COVID-19 pandemic has shown that sustaining those prevailing standards of evidence would result in providing no guidance—a luxury few guidance producing bodies felt they could permit themselves. So how did developers cope in the meantime: how did guidance developers act while experiencing a "failure of routines?" (see Reference [14], p. 80). How did they substantiate recommendations based on best knowledge available? How did they accommodate uncertainty in guidance? And what consequences do they anticipate, based on their experience with pandemic guidance production, for guidance production post COVID-19?

This study reports on the practices of guidance development during the COVID-19 pandemic to illustrate how developers have responded to the need for credible guidance, in the face of uncertain knowledge and extreme time pressure. Drawing on those practices and experiences, we reflect on lessons that can be drawn for the future of guidance production.

2 | METHODS

2.1 | Study design

The study involved two phases: a web-based survey and qualitative follow-up interviews. The Institutional Review Board of the Faculty of Science (BETHCIE) of the Vrije Universiteit Amsterdam approved this research as exempt.

2.2 | Survey

The survey instrument (Supporting information) was drafted by M.M., J.R. and T.Z.J. Given the timing of survey administration in May 2020, during the aftermath of the first pandemic wave, we opted for a "rapid survey" design consisting of 12 items, taking into account that many developers were deeply involved in the hectic work of guidance production, while this was also a crucial time to capture their experiences. Eight closed-ended questions inquired about knowledge sources used to develop COVID-19 guidance and methods used for rapid knowledge appraisal and synthesis. Respondents could add comments

to clarify their answer. In addition, four open-ended survey questions inquired about the challenges in developing COVID-19 guidance, strategies used to address these and perspectives on the implications of the COVID-19 pandemic on future guidance development. F.M., S.W., B.S. and F.F. reviewed and revised the survey. The survey was sent out for external review to members of the Guidelines International Network (GIN), with one member providing comments. In May 2020, the survey was sent to 1333 members of GIN via email. The rapid survey was carried out among the members of the Guidelines International Network, since they comprise a broad group of guideline developers from around the world. Participants were informed about the purpose of the study and the anonymity of all study participants was guaranteed. Consent was confirmed upon agreement to participate in this study.

The survey consisted of closed questions that generated nominal data, entered into SPSS Statistics (version 26) software and analysed using descriptive statistics. Transcripts from open-ended survey questions were analysed with a grounded theory approach in which codes, categories and themes that emerged from the data were identified by using the ATLAS.ti software. Coding was conducted by two independent reviewers (M.M. and J.R.) based on the interview protocol and research questions. The most salient codes and relationships between codes were identified and aggregated into descriptive themes, using thematic synthesis.

2.3 | Interviews

Trained research interviewers (M.M. and J.R.) conducted the semi-structured follow-up interviews using Zoom videoconferencing between June and August 2020 to further understand challenges and changes in guidance production. We invited survey respondents who indicated willingness to participate in follow-up interviews via email (4 participated, 4 declined participation, 3 did not respond, 1 was not eligible). People who declined indicated being too busy to participate. Furthermore, one interview participant was recruited using snowballing as a supplementary approach to the main recruitment method. The participants and the interviewers did not have any contact before the interview other than to set the appointment. Participants were contacted via email and were informed that they were going to participate in a 45-minute interview about their experiences with developing COVID-19 guidance. Anonymity was guaranteed and interviewees were asked to agree to the terms of the interview verbally.

The interview guide (Supporting information) was developed on the basis of the survey results. The guide inquired about how the challenges of the COVID-19 pandemic affected guidance procedures. Probing questions were used to get a better sense of the changes in practice that were identified through the survey. Finally, participants were asked to reflect on the implications of the COVID-19 pandemic on future guidance development. With participant consent, interviews were audio recorded and transcribed verbatim. Analysis was performed following the same methods as for qualitative analysis of the survey.

3 | RESULTS

The survey was completed by 76 members from 26 countries, of whom 46 members were currently involved in developing COVID-19 guidance. Demographic characteristics are summarized in Table 1. Follow-up interviews were conducted with two developers/information specialists and three guidance department directors.

TABLE 1 Survey respondent characteristics ($n = 46$)

Characteristics	Data[a]
Specialty (MC)	
Chronic care	35% (16)
Emergency care	15% (7)
Long-term care	9% (4)
Mental healthcare	17% (8)
Nursing	11% (5)
Occupational health	9% (4)
Oncology	13% (6)
Paediatrics	20% (9)
Primary care	28% (13)
Public health	20% (9)
Surgery	9% (4)
All of the above	13% (6)
Experience in guidance production	
<2 years	11% (5)
2–5 years	17% (8)
5–10 years	11% (5)
10–20 years	37% (17)
>20 years	24% (11)
Guidance developed for	
National use	65% (30)
International use	15% (7)
Local or regional use	20% (9)
Region of practice	
Africa	13% (6)
Asia	2% (1)
Australia	2% (1)
Europe	37% (17)
United States of America	39% (18)
Central and South America	2% (1)
Standardized methodology used within institutions (MC)	
National standard for guideline development	43% (20)
Inhouse developed guideline development methodology	41% (19)
GRADE methodology	33% (15)
Other	15% (7)

[a]Data are the number (percentage) of survey respondents. Percentages may not total to 100 owing to rounding and/or multiple-choice questions (MC).

4 | WILEY— Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

MOLEMAN ET AL.

3.1 | Knowledge appraisal in a setting of limited evidence

The survey inquired about knowledge sources used to develop COVID-19 guidance and methods used for rapid knowledge appraisal and synthesis. Conventional methods of guidance development were largely unsuited to the production of recommendations in this new context, with 80% ($n = 37$) of survey respondents resorting to different procedures for appraising and including knowledge. In the comments section, respondents described that conventional methods were too lengthy, too reliant on systematic reviews of empirical research evidence and poorly suited to the appraisal of other types of knowledge. Conventional methods were also not designed to respond rapidly to new knowledge; respondents highlighted flexibility as crucial, producing rapid updates to guidance, occasionally many times. Figure 1 shows that developers often relied on expert opinion (76.6%, $n = 36$) and other sources of information otherwise considered as being at too high a risk of bias on which to base recommendations.

3.2 | Responding to the need for credible guidance in COVID-19 times

This section reports on findings from the survey and the interviews regarding the greatest challenges in developing COVID-19 guidance and strategies used to address these. Findings indicated that developers were working with a very limited number of studies, mostly with small study samples and concomitant risk of bias. In addition, many experts, who usually can delegate parts of the everyday clinical work to more junior staff and dedicate time to serve on guidance working groups that evaluate evidence and develop recommendations, were extremely busy with clinical duties at the frontline. One respondent notes that *"knowledge regarding diagnosis, treatment was evolving rapidly [in the field], while the evidence base is still limited"*. In addition, other respondents reported lacking rigorous methods to appraise and include these wide-ranging types of knowledge to come to robust guidance in the absence of frequentist (RCT-based) evidence. Guidance institutions were confronted with the challenge of

maintaining their reputation, while the rapidly changing knowledge about COVID-19 and its clinical management meant that guidance was often only valid for a short time - weeks, days (or even hours). As one survey respondent described, the most pressing challenge was the development of: *"credible guidelines that folks want to adhere to, [but] that are stale a week after they are published"*.

Based on our analysis of responses, we identified two strategies in guidance development to cope with these challenges. The urgency of need and the rapidly changing evidence base worked as catalysts to strengthen two mechanisms for bolstering the credibility of guidance: (1) Pre- and post-publication end-user involvement and (2) conjoining evidence review and recommendation formulation.

3.3 | Pre- and post-publication end-user involvement

Developers reported that the conventional phases of guideline development were followed, albeit shortened, as there was simply so little relevant evidence that was sufficiently robust to review. A survey respondent described: *"Rapid rate of publication of findings, lack of rigour of most published findings, many papers are based on limited sample sizes (..) and/or on other countries' data with different patient demographics and resources"*. The poor quality of evidence and the need to translate evidence to local contexts underlined the importance of expert input during guidance development and tracking the validity of published guidance in practice.

Getting diverse clinical input into the guidance working groups was reported by three interviewees as one of the biggest challenges. One interviewee described that their institution tried to make working groups as diverse as possible, recognizing that guidance was largely based on consensus—the credibility of which depends heavily on the representation of diverse views. In addition, feedback mechanisms were strengthened to move more towards the concept of *living guidelines* to allow developers to provide guidance that was "good enough" to respond to immediate crises while keeping closely in touch with practice and changes in knowledge. For example, an interviewee described designated COVID-19 channels through which clinicians could submit feedback and requests for guidance.
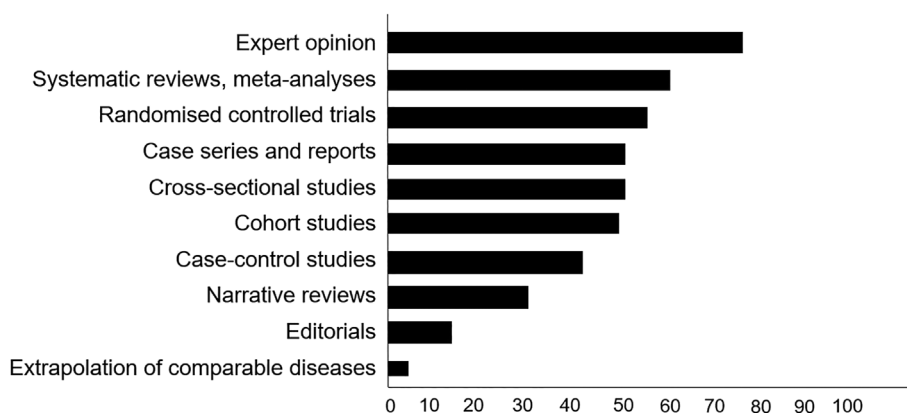

FIGURE 1 Knowledge types used for COVID-19 guidance. Figure 1 indicates what knowledge types respondents (% of $n = 46$) used for developing COVID-19 guidance. *Editorials* came up as a new category. Many journals have published *editorials* to allow for swift publication of experiences with COVID-19. In addition, the extrapolation of evidence from comparable diseases was added as a source of knowledge

[After the first wave we] build in this 'living guideline' concept and that helped us with the speed of development, because what we could say to the teams and (…) the clinicians: "don't expect perfect". The issue here is having something that is good enough for what was clearly an emergency (…) [building in] a rapid update in the process so as more evidence emerges, we could tweak and change (…) and having mechanisms for clinicians to feedback (Interviewee 4).

These feedback loops were based on principles of transparency and flexibility in order to develop credible guidance in a context of limited and uncertain knowledge.

The need to speed up the processes of collating evidence and developing guidance thus forced developers to make pragmatic choices. Where mechanisms for feedback by clinicians were installed rapidly, similar involvement of patients was largely abandoned in the early stages of the pandemic. An interviewee mentioned for example that patients' levels of involvement dropped drastically and later were somewhat restored. "*At the very beginning there was no patient involvement at all. We're now at a stage where there is a lot of patient involvement but not at the same level as before*". The quote illustrates how initial choices to bolster pragmatism were later revised to rebalance considerations of rigour and pragmatism. Furthermore, where feedback by clinicians was instantly seen as central to creating living guidelines in times of uncertainty, the same did not apply to patients.

3.4 | Conjoining evidence review and recommendation formulation

In addition to fostering close connections between guidance institutions and end users, teams working within guidance institutions also sought closer collaboration. As an interviewee described, separate teams whose main activities were either evidence surveillance or guidance development were joined to allow for rapid updating of COVID-19 guidance.

[Normally] we have a surveillance team that is continually trawling the evidence [and] topic experts [judging] "we should or should not update the guideline" (..) with Covid we're bringing those closer together or making them concurrent because of the repetitiveness of the evidence changes and the need to be as up to date as we can (interviewee 4).

Another interviewee described how evidence appraisal teams (systematic reviewers) became more closely involved in formulating recommendations, removing boundaries between activities traditionally performed separately by systematic reviewers and developers with topic expertise. The interviewee described how their institution had tried for years to shift the culture and bring the work of systematic reviewers and guidance developers closer.

This shift was catalysed by the pandemic and the need for faster guidance production. Systematic reviewers became closer to the translation of evidence into practice recommendations, while relying more heavily on sources of evidence at higher risk of bias than conventionally accepted and on consensus. This led to a better understanding and more respect for the work of different teams.

3.5 | Future perspectives

This section provides an overview of survey and interview findings regarding possible future implications of the COVID-19 pandemic on the future of guidance production beyond the pandemic. The vast majority of survey respondents (70%, $n = 37$) thought it possible that the failure of existing routines and the new practices that emerged while developing guidance in response to the pandemic could lead to future changes in conventional guidance development, with 44% ($n = 20$) of respondents actually expecting this change to occur. The remaining 30% ($n = 14$) considered the COVID-19 pandemic to be an exceptional situation and emphasized the need to return to traditional, "gold standard" procedures as soon as possible.

The majority of respondents expected or hoped that this approach to guidance development might influence future guidance production, most notably by shortening development time. An interviewee said that developers, systematic reviewers and end-users need to be willing to acknowledge the inevitable uncertainty and the associated risks of this in all guidance. Guidance recommendations are developed primarily to guide clinical, public health or social action and to minimize risk to patients and the public. Nonetheless, an approach to guidance that is too risk-averse and prolongs development is in itself a huge risk.

Another interviewee noted that it is currently difficult to determine when conventional methods for maximizing the quality of the research evidence are needed, and when other mechanisms to produce credible guidance can be used. They called for a discussion within the guidance community to develop better criteria to strike the right balance between the two modes of guidance production.

In cases of poor evidence, the superiority of GRADE as opposed to other, less extensive, methods for grading, is not straightforward (…) especially when you know in advance that there is no good quality evidence. We need to design criteria regarding when to resort to GRADE or not (interviewee 1).

This quote highlights the need to explain uncertainties inherent in guidance development. The need to identify when it is wise to resort to GRADE or not is also supported by other studies. Mercuri and Gafni[18] have for example argued that although transparency about weighing the evidence is considered one of the main advantages of using GRADE, in practice choices often remain insufficiently substantiated

6 | **WILEY**— Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

MOLEMAN ET AL.

and explained, leaving considerable space for interpretation regarding the operationalisation of the GRADE framework. One survey respondent remarked: *"For all the wrong reasons, I am happy that at last our greatest critics are now faced by the same problems we had to deal with all of our careers"*. The 'critics' referred to are those who argue that guidance should rely only on "high quality" evidence, regardless of the question being asked. This respondent was expecting the COVID-19 pandemic to challenge such attitudes and help find ways to include other sources of knowledge in guidance development.

4 | DISCUSSION

This study describes how guidance developers coped with the very specific problem of extreme knowledge uncertainty and urgent time pressures during the COVID-19 pandemic. Some respondents considered COVID-19 guidance the "messy exception to the gold standard". However, most developers thought that the COVID-19 outbreak uncovered real anxieties about the complete absence of empirical evidence while working with tools that did not help to deal with such major uncertainties. The development of methods for ensuring the statistical reliability and internal validity of the evidence,[9] has been an important driving force of the EBM movement[19] to safeguard the trustworthiness of guidance. The quality of the evidence in this context is equated with the level of statistical confidence in study outcomes.[20] Baigrie and Mercuri[9] posit that this tradition insufficiently acknowledges that what is considered "best available research evidence" may be different from field to field, or question to question. The authors describe that the prioritization of the "goodness" of evidence in evidence-based medicine defined as preferencing study designs that provide the most reliable/valid basis for studying causation, insufficiently acknowledges that the *goodness* of evidence and the *relevance* of the evidence for the question at hand are distinct criteria. Perillat and Baigrie[5] note that the evaluation of methodological rigour is an essential element for supporting evidence-based decision-making, but also argue that other criteria including relevance, feasibility, adaptability and acceptability should be taken into account - considerations that may be better addressed using different kinds of studies. Hence, "best research evidence" is not merely defined by the (statistical) "confidence" in the evidence, but also needs to speak to the situation for which guidance is being developed and the objective(s) of the guidance. In an opinion piece entitled "controversial policies on COVID-19 stem from a deeply rooted view of evidence"[21] Sager analyses how maintaining high evidence-thresholds for COVID-19 policies in Sweden, resulted in guidance that ironically bears close resemblance to those of the anti-scientific approach of populists like Donald Trump: no lockdown or mask mandates in the pandemic's first 6 months. Sager notes that the risk of such a technocratic approach is that "a lack of positive evidence can easily be misread as negative evidence" and that inflexible evidence thresholds "ignore the merits of a broader approach to evidence".[21] Guidance developers in other locations almost invariably ended up abandoning prevailing standards of evidence for guidance development as decisions had to be made even

without gold-standard evidence. This study's findings describe which mechanisms guidance developers mobilized to balance "goodness" and "relevance", a process we paraphrase as shifting the focus from "getting guidance right" to "getting guidance right now".

By studying the practices of guidance developers during the pandemic, we have identified two changes in their working processes. First, the COVID-19 outbreak highlights the need to strengthen expert feedback before and after publication to keep guidance connected and relevant in case of a changing knowledge base,[22] the emergence of crucial first-hand clinical experience and in a setting where the political and social acceptability of public health guidance cannot be drawn from having used rigid epistemic criteria. Furthermore, the involvement of end-users helps to ensure that questions are matched to the most pertinent needs from the field.[5] Second, linking more closely the evidence appraisal and recommendation formulation can short-circuit the usual, more lengthy guideline development processes. The focus is on publishing rapid guidance that, supported by pragmatic mechanisms, could be updated as new insights emerged, rather than guidance intended to be used for a number of years. This allows the focus to shift from getting recommendations that are based on the statistical "goodness" of evidence to getting those that are predominantly "relevant"; that is, less-than-perfect but when needed. Through these mechanisms developers avoid needing to choose between harms from issuing incorrect or incomplete guidance and the ethical risk of issuing no guidance at all. These mechanisms offer ways to manage epistemic uncertainty and unpredictability, by balancing scientific robustness with usability, acceptability, adequacy[23] and contingency of the guidance[22] in more explicit ways.

A strength of this study is that the timing of the survey administration coincided with the aftermath of the first pandemic wave, making the findings particularly timely and pertinent. In addition, the exploratory mixed-methods design allowed us to gain a better understanding of the challenges reported in the survey and discuss why participants expected changes in future guidance production.

A limitation of this study was that due to the timing and pertinence, we needed to adjust to the hectic situations many developers found themselves in. Given these circumstances, we opted for a "rapid survey" design consisting of limited number of items. Although the survey instrument was reviewed and tested by three researchers experienced in the field of guidance development, it was not feasible to fully validate it. In addition, since we wanted to do the interviews while memories of guidance development were still recent and current, we were able to recruit no more than the limited number of five respondents for the follow-up interviews. These study design choices lowered methodological rigour, but enabled us to capture aspects of the worldwide experiences regarding guidance development practices during the first pandemic wave, as it was happening. Thus, this study provides exploratory impressions of what guidance development during the COVID-19 outbreak looked like, rather than seeking full data saturation.

To conclude, this study draws lessons about guidance development during the COVID-19 crisis and points to the need to acknowledge and communicate uncertainties of scientific evidence more explicitly,[13,24] by mobilizing important mechanisms that allow

guidance developers to live with such uncertainty.[25] Fostering these mechanisms, especially those of strengthening pre- and post-publication end-user involvement and more tightly conjoining evidence review and recommendation formulation, may well prove important contributions to extending guidance development methods and processes more widely.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS

Marjolein Moleman and Teun Zuiderent-Jerak designed the study methodology with Fergus MacBeth, Sietse Wieringa, Frode Forland and Beth Shaw contributing significantly to revising this. Marjolein Moleman drafted the initial manuscript with all authors contributing significantly to revising this for submission. All authors agreed on the final version for submission to the journal.

ENDNOTES

* The development of clinical guidance, for example, generally takes between 6 months up to 2 years.

† Policies on the use of face coverings are a telling example of the mismatch between the high standards of evidence that warrant action and the urgency of policy needs to act in a situation of uncertainty. At first, the World Health Organization did not support the implementation of face coverings in public, since the observed effect of facemasks was classified as "low" certainty based on the GRADE framework. Evidence on face coverings drawn from laboratory science, mathematical modelling and policy studies, for example, fail to generate a quantitative estimate of effect size to support and legitimize policy action.[26] Greenhalgh[15] notes that the overuse of methodological quality as a criterion for action is especially problematic in a setting in which the assumptions of EBM cannot feasibly be met, nor is such evidence relevant or desirable for the question at hand. Even the Cochrane Collaboration published an article to underline the need to act on incomplete evidence during the COVID-19 pandemic, noting that we may never have strong evidence on public health interventions like individual behavioural measures, nor are trials able to capture the effects of behavioural measures "when deployed a general population living in the complex, diverse circumstances of individuals' everyday lives" (see Reference [13], p. 1). In this setting, the authors note, "waiting for strong evidence is a recipe for paralysis" (see Reference [13], p. 1).

‡ Incorporating the experiences and values by patients, professionals, and citizens in COVID-19 guidance, given its high time pressure, may only become possible through methodological innovation, for example through deploying artificial intelligence for harvesting such experiences and value-considerations. For first attempts at such inclusion of experience and values under time pressure, see: https://www.zonmw.nl/nl/onderzoek-resultaten/kwaliteit-van-zorg/programmas/project-detail/kwaliteit-van-zorg-ondersteuning-kwaliteitsinstituut/evidence-in-actie-dynamisch-standaardiseren-van-ervaringskennis-en-waarden-in-richtlijnen/.

DATA AVAILABILITY STATEMENT

The anonymised data and materials of this study are available from the corresponding author on reasonable request.

ORCID

Marjolein Moleman https://orcid.org/0000-0002-5289-2341
Teun Zuiderent-Jerak https://orcid.org/0000-0003-3008-1297

REFERENCES

1. Emanuel EJ, Persad G, Upshur R, et al. Fair allocation of scarce medical resources in the time of COVID-19. *N Engl J Med*. 2020;382(21): 2049-2055. doi:10.1056/nejmsb2005114
2. Burman R, Cairns R, Canestrini S, et al. Making ordinary decisions in extraordinary times. *BMJ*. 2020;370:1-3. doi:10.1136/bmj.m3268
3. Steens A, Freiesleben de Blasio B, Veneti L, et al. Poor self-reported adherence to COVID-19-related quarantine/isolation requests, Norway, April to July 2020. *Eurosurveillance*. 2020;25(37):2001607. doi:10.2807/1560-7917.ES.2020.25.37.2001607
4. Munn Z, Twaddle S, Service D, et al. Developing guidelines before, during, and after the COVID-19 pandemic. *Ann Intern Med*. 2020; 173(12):1012-1014. doi:10.7326/m20-4907
5. Perillat L, Baigrie BS. COVID-19 and the generation of novel scientific knowledge: evidence-based decisions and data sharing. *J Eval Clin Pract*. 2021;27(3):708-715. doi:10.1111/jep.13548
6. Knaapen L, Cazeneuve H, Cambrosio A, Castel P, Fervers B. Pragmatic evidence and textual arrangements: a case study of French clinical cancer guidelines. *Soc Sci Med*. 2010;71(4):685-692. doi:10.1016/j.socscimed.2010.05.019
7. Wieringa S, Dreesens D, Forland F, et al. Different knowledge, different styles of reasoning: a challenge for guideline development. *BMJ Evid Based Med*. 2018;23(3):87-91. doi:10.1136/bmjebm-2017-110844
8. Tonelli MR, Upshur REG. A philosophical approach to addressing uncertainty in medical education. *Acad Med*. 2019;94(4):507-511. doi: 10.1097/ACM.0000000000002512
9. Baigrie B, Mercuri M. Relevance, validity, and evidential reasoning in clinical practice. *J Eval Clin Pract*. 2020;26(5):1341-1343. doi: 10.1111/jep.13474
10. Fretheim A, Brurberg KG, Forland F. Rapid reviews for rapid decision-making during the coronavirus disease (COVID-19) pandemic, Norway, 2020. *Eurosurveillance*. 2020;25(19):2000687. doi:10.2807/1560-7917. ES.2020.25.19.2000687
11. Kodadek LM, Berger JC, Haut ER. Guidance vs. guidelines: the role of evidence-based medicine in the COVID-19 pandemic. *J Patient Saf Risk Manage*. 2020;25:1-3. doi:10.1177/2516043520965844
12. Aronson JK, Auker-Howlett D, Ghiara V, Kelly MP, Williamson J. The use of mechanistic reasoning in assessing coronavirus interventions. *J Eval Clin Pract*. 2020;27(3):684-693. doi:10.1111/jep.13438

8 | WILEY— Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

MOLEMAN ET AL.

13. Soares-Weiser K, Lasserson T, Jorgensen KJ, et al. Policy makers must act on incomplete evidence in responding to COVID-19. *Cochrane Database Syst Rev*. 2020;11:ED000149. doi:10.1002/14651858.ED000149

14. Pols J. Knowing patients: turning patient knowledge into science. *Sci Technol Hum Values*. 2014;39(1):73-97. doi:10.1177/0162243913504306

15. Greenhalgh T. Will COVID-19 be evidence-based medicine's nemesis ? *PLoS Med*. 2020;17(6):e1003266. doi:10.1371/journal.pmed.1003266

16. Van Dooren W, Noordegraaf M. Staging science: authoritativeness and fragility of models and measurement in the COVID-19 crisis. *Public Adm Rev*. 2020;80(4):610-615. doi:10.1111/puar.13219

17. Mercuri PDM. Just follow the science: a government response to a pandemic. *J Eval Clin Pract*. 2020;26(6):1575-1578. doi:10.1111/jep.13491

18. Mercuri M, Gafni A. The evolution of GRADE (part 3): a framework built on science or faith? *J Eval Clin Pract*. 2018;24(5):1223-1231. doi:10.1111/jep.13016

19. Saarni SI, Gylling HA. Evidence based medicine guidlines: a solution to rationing or politics disquised as science? *J Med Ethics*. 2004;30(2):171-175. doi:10.1136/jme.2003.003145

20. Graham R, Mancher M, Wolman DM. *Clinical Practice Guidelines We Can Trust*. Washington D.C.: National Academies Press; 2011. doi:10.17226/13058

21. Sager M. Why Sweden chose a different path through the pandemic—Research Professional News. Research Professional News; 2021. https://www.researchprofessionalnews.com/rr-news-europe-views-of-europe-2021-2-why-sweden-chose-a-different-path-through-the-pandemic/. Accessed 15 March 2021.

22. Dagens A, Sigfrid L, Cai E, et al. Scope, quality, and inclusivity of clinical guidelines produced early in the COVID-19 pandemic: rapid review. *Br Med J*. 2020;369:m1936. doi:10.1136/bmj.m1936

23. Moreira T. Diversity in clinical guidelines: the role of repertoires of evaluation. *Soc Sci Med*. 2005;60(9):1975-1985. doi:10.1016/j.socscimed.2004.08.062

24. Palmer S, Jansen A, Leitmeyer K, Murdoch H, Forland F. Evidence-based medicine applied to the control of communicable disease incidents when evidence is scarce and the time is limited. *Eurosurveillance*. 2013;18(25):20507. doi:10.2807/1560-7917.ES2013.18.25.20507

25. Jerak-Zuiderent S. Certain uncertainties: modes of patient safety in healthcare. *Soc Stud Sci*. 2012;42(5):732-752. doi:10.1177/0306312712448122

26. Science Media Centre. Expert Reaction to Review of Evidence on Face Masks and Face Coverings. Science Media Centre; 2020. https://www.sciencemediacentre.org/expert-reaction-to-review-of-evidence-on-face-masks-and-face-coverings-by-the-royal-society-delve-initiative/. Accessed 30 August 2021.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.