MICHÈLE FONFRÈDE, JEAN PAUL COUAILLAC, CHRISTINE AUGEREAU, JEAN PIERRE LEPARGNEUR, JOSEPH WATINE

CRITICAL APPRAISAL OF MICROBIOLOGY GUIDELINES ENDORSED BY TWO PROFESSIONAL ORGANISATIONS: SOCIÉTÉ FRANÇAISE DE MICROBIOLOGIE (SFM) AND AMERICAN SOCIETY OF MICROBIOLOGY (ASM)

Critical appraisal of microbiology guidelines endorsed by two professional organisations: Société Française de Microbiologie (SFM) and American Society of Microbiology (ASM)

Working Group "Guidelines and Evidence-Based Laboratory Medicine" of the Société Française de Biologie Clinique (SFBC)

Michèle Fonfrède[1], Jean Paul Couaillac[2], Christine Augereau[3], Jean Pierre Lepargneur[4], Joseph Watine[5]
[1]Biochimie Métabolique, Hôpital de la Pitié-Salpétrière, Paris
[2]Biologie polyvalente, Hôpital de Cahors
[3]Biochimie, Hôpital Européen Georges Pompidou, Paris
[4]Directeur Médical, Novescia, Paris
[5]Biologie polyvalente, Hôpital de la Chartreuse, Villefranche-de-Rouergue

Corresponding Author:

Joseph Watine
Laboratoire de biologie polyvalente
Hôpital de la Chartreuse
avenue Caylet, 12200 Villefranche de Rouergue
Tel: 05.65.65.30.00
e-mail: watine61@hotmail.com.

Conflicts of interest

Acknowledgements

Abstract

Medical practice guidelines (GLs) being tools that are mainly designed to evaluate medical professionals, it sounds logical, and fair, that professionals should in turn evaluate GLs. Microbiology being a medical discipline, we used the AGREE instrument, i.e. an established evaluation tool for GLs, in order to evaluate the quality of two major microbiology guidelines, i.e. the SFM GLs and the ASM GLs). Both guidelines remain sub-optimal in their levels of quality, and obtain scores that are not very different from the average scores obtained by many other guidelines in various medical disciplines. We therefore believe that both guidelines need to be modified before they can be recommended without provisos. A higher degree of multi-disciplinary work, including a more formal implication of methodologists, as well as of infectious disease clinicians, and of economists, might perhaps enable future editions of these guidelines to reach higher levels of quality.

Key-words

Evidence-Based-Medicine; practice guidelines; recommendations; SFM; ASM; AGREE Quality.

MICHÈLE FONFRÈDE, JEAN PAUL COUAILLAC, CHRISTINE AUGEREAU, JEAN PIERRE LEPARGNEUR, JOSEPH WATINE

CRITICAL APPRAISAL OF MICROBIOLOGY GUIDELINES ENDORSED BY TWO PROFESSIONAL ORGANISATIONS: SOCIÉTÉ FRANÇAISE DE MICROBIOLOGIE (SFM) AND AMERICAN SOCIETY OF MICROBIOLOGY (ASM)

INTRODUCTION

Medical practice guidelines (GLs) being tools that are mainly designed to evaluate medical professionals, it sounds logical, and fair, that professionals should in turn evaluate GLs. As far as we know, when it comes to appraise the quality of GLs, the AGREE criteria are the only ones that have been largely validated at an internationla level, in particular by the World Health Organisation and by the European Union, and most authors who publish in the field of GLs appraisal use the AGREE tool (1). Thus in a systematic review of 42 studies where 626 GLs were evaluated using the AGREE criteria, it was clearly shown that there is very much room for improvement of GLs (1).

Although microbiology is a medical discipline, few microbiology GLS are available in the litterature, and this might perhaps explain why, to our knowledge, no GLs appraisal study has ever been published in this field, except for a preliminary work of ours that was published two years ago. In this preliminary work we evaluated, with the help of the 1st edition of the AGREE tool (2), the quality of the only French GLs that were available, i.e. the 3rd edition of the Rémic (microbiology guidelines - bacteriology and mycology) of the *Société Française of Microbiologie* (SFM) (3). One of the main limitations of this preliminary work was that we did not compare the quality of the Rémic with guidelines published by others. Also the authors of the Rémic told us that our evaluation would help them to improve future editions of their guidelines (4). Since then a 4th edition of the Rémic (5) as well as a 2nd edition of the AGREE tool (6) have been released. We therefore used the 2nd edition of the AGREE tool in order to compare the methodological quality of the 3rd and of 4th editions of the Rémic, with particular emphasis on the recommendations dedicated to the laboratory diagnosis of urinary tract infections, and in comparison with the similar recommendations of the *American Society of Microbiology* (ASM), ie CUMITECH (7).

MATERIAL AND METHODS

We used the Appraisal of Guidelines for Research and Evaluation (AGREE) tool, which is composed of 23 items grouped in 6 domains (for more details, see table 1 and references 2 and 5).

Five members of our group (JPC, CA, MF, JW, JPL) independently attributed scores, from 1 to 7, to each of the 23 items. This part of our work was facilitated by the fact that we previously made a similar work (where we used the 1st version of the AGREE tool in order to evaluated the 2007 edition of the Rémic (4)). Then we had some discussions *via* telephone calls or *via* emails where we managed to eliminate differences above 2 in our scores. We then calculated mean scores for each item, and then scores for each domain (6). The results thus obtained (table 1) enable us to conclude, using the AGREE terminology, that the three guidelines need to be modified before they can be recommended without provisos.

RESULTS AND DISCUSSION

In the afore-mentioned systematic review of 42 studies where 626 guidelines were evaluated using the AGREE criteria (1), the mean scores were (table 1):
- 64% (95% CI 62 to 66) for 'Scope and purpose' (domain 1);
- 35% (95% CI 34 to 37,5) for 'Stakeholder involvement' (domain 2);
- 43% (95% CI 41 to 45) for 'Rigour of development' (domain 3);
- 60% (95% CI 58 to 62) for 'Clarity and presentation' (domain 4);
- 22% (95% CI 20 to 24) for 'Applicability'. (domain 5);
- 30% (95% CI 28 to 32) for 'Editorial independence' (domain 6).

We can see (table 1) that our scores are quite close to these scores above, except perhaps for:
- Domain 2, where ASM scores much lower than 35% (ie 16%)
- Domain 4, where the SFM's latest edition, and ASM scores higher than 60% (ie respectively 84% and 86%)
- Domain 5, where the SFM's latest edition, and ASM scores higher than 22% (ie respectively 52% and 55%)

The figures in Table 1 suggest that Rémic 2010 has improved compared to Rémic 2007. It is also good news to see that guidelines dedicated to laboratory diagnosis tend to score higher in 'Clarity and presentation' (domain 4) and in 'Applicability' (domain 5) than do guidelines in other fields of medicine (1). However we can see in Table 1 that there is still room for improvement in most items and domains.

We discuss below the areas where it seems to us efforts should be made in priority in future editions of the SFM guidelines as well as in future editions of CUMITECHs.

Although items and domain scores are rather high in domain 1 (Scope and Purpose), the main problem that we had with this domain is that the recommendations in these guidelines do often not answer to structured clinical questions according to the classical PICO (*Patient, Intervention, Comparator, Outcome*) frame (8). For example the Rémic guidelines recommends in its 2010 edition: "antibiograms have to be performed for all isolates responsible for urinary tract infections".

MICHÈLE FONFRÈDE, JEAN PAUL COUAILLAC, CHRISTINE AUGEREAU, JEAN PIERRE LEPARGNEUR, JOSEPH WATINE

CRITICAL APPRAISAL OF MICROBIOLOGY GUIDELINES ENDORSED BY TWO PROFESSIONAL ORGANISATIONS: SOCIÉTÉ FRANÇAISE DE MICROBIOLOGIE (SFM) AND AMERICAN SOCIETY OF MICROBIOLOGY (ASM)

Table 1
Critical appraisal of three guidelines (3, 5, 7) according to AGREE (6)

DOMAINS	ITEMS		Item scores SFM3/ SFM4/ ASM	*Domain scores SFM3/ SFM4/ ASM
DOMAIN 1. Scope and purpose	1)	The overall objective(s) of the guideline is (are) specifically described	5.2/ 5.6/ 5.4	**69%/77%/73%**
	2)	The health question(s) covered by the guideline is (are) specifically described	5.2/ 5.8/ 5.4	64%
	3)	The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described	5.0/ 5.4/ 5.4	
DOMAIN 2. Stakeholder involvement	4)	The guideline development group includes individuals from all relevant	2.8/ 2.6/ 2.4	**40%/41%/16%**
	5)	The views and preferences of the target population (patients, public, etc.) have	1.2/ 1.2/ 1.4	35%
	6)	The target users of the guideline are clearly defined	6.2/ 6.6/ 2.0	
DOMAIN 3. Rigor of development	7)	Systematic methods were used to search for evidence	4.4/ 4.6/ 4.6	**37%/46%/53%**
	8)	The criteria for selecting the evidence are clearly described	3.4/ 3.6/ 4.6	43%
	9)	The strengths and limitations of the body of evidence are clearly described	3.0/ 3.8/ 4.0	
	10)	The methods for formulating the recommendations are clearly described	3.0/ 3.4/ 3.6	
	11)	The health benefits, side effects, and risks have been considered in formulating the recommendations	3.2/ 4.6/ 5.6	
	12)	There is an explicit link between the recommendations and the supporting evidence	3.4/ 4.6/ 5.2	
	13)	The guideline has been externally reviewed by experts prior to its publication	2.0/ 2.4/ 2.8	
	14)	A procedure for updating the guideline is provided	3.2/ 3.4/ 2.8	
DOMAIN 4. Clarity of presentation	15)	The recommendations are specific and unambiguous	5.2/ 6.4/ 6.4	**66%/84%/86%**
	16)	The different options for management of the condition or health issue are clearly presented	4.6/ 5.6/ 6.6	60%
	17)	Key recommendations are easily identifiable	5.0/ 6.2/ 5.4	
DOMAIN 5. Applicability	18)	The guideline describes facilitators and barriers to its application	3.0/ 4.8/ 5.4	**26%/52%/55%**
	19)	The guideline provides advice and/or tools on how the recommendations can be put into practice	3.2/ 4.6/ 5.8	22%
	20)	The potential resource implications of applying the recommendations have been considered	2.0/ 2.7/ 3.2	
	21)	The guideline presents monitoring and/or auditing criteria	2.2/ 4.8/ 2.7	
DOMAIN 6. Editorial independence	22)	The views of the funding body have not influenced the content of the guideline	4.4/ 4.2/ 4.4	**28%/27%/32%**
	23)	Competing interests of guideline development group members have been recorded and addressed	1.0/ 1.0/ 1.4	30%

*Legends: Scores items: from 1 to 7 (from weakest to strongest). Scores domains: from 0 to 100% (from weakest to strongest). *In red characters: results of a review of 42 studies where 626 guidelines were evaluated using the AGREE criteria (1).*

This example of recommendation above from the Rémic 2010 answers to a question which is not very specific about the "P" (do all patients with, or without, particular symptoms or particular risk factors need to be tested the same?), and even less specific about the "O" (what are the expected benefits, or risks, or costs of performing antibiotic susceptibility testing in all populations of patients?) and the "C" (is antibiotic susceptibility testing the only possible option? couldn't the probabilistic approach be another option for treating some patients?). This recommendation lacks of medical, as well as of economical, and of ecological, pertinence.

Regarding the usefulness of performing antibiograms in case of urinary tract infection, an example of structured clinical questions could have been:
"In an young female patient with a first episode of uncomplicated cystitis treated by a quinolone (P), is antibiotic-susceptibility-testing in order to change the therapy in case of in vitro resistance to the quinolone (I), increasing the likelyhood of curing the patient? and/or increasing the financial costs? and/or increasing the risk of selecting bacteria that are resistant to antibiotics? (O), compared to just keeping the bacterial isolate in the lab for a few days, or weeks, in order to be able to perfom later antibiotic-susceptibility-testing in case therapeutic failure would occur (C)"?

MICHÈLE FONFRÈDE, JEAN PAUL COUAILLAC, CHRISTINE AUGEREAU, JEAN PIERRE LEPARGNEUR, JOSEPH WATINE

CRITICAL APPRAISAL OF MICROBIOLOGY GUIDELINES ENDORSED BY TWO PROFESSIONAL ORGANISATIONS: SOCIÉTÉ FRANÇAISE DE MICROBIOLOGIE (SFM) AND AMERICAN SOCIETY OF MICROBIOLOGY (ASM)

In our own example above, there are in fact three sub-questions that would each need a systematic review of the litterature. Based on the evidence thus examined, one would expect that the authors of the guidelines make an explicit judgment where both clinical benefits, and clinical costs, would be weighted against their financial, and ecological, counterparts. We cannot see how such a judgment could be wisely made by microbiologists alone. For example, should it be for micobiologists alone to decide how the resources of the collectivity have to be spent?

Regarding Domain 2 (Stakeholder involvement), all of the guidelines score rather low for items 4 and 5 (Table 1) because:
- we do not always know the names, discipline/content expertise, institution, geographical location of each member of the guideline development group, together with a description of the member's role in the guideline development group;
- some relevant professions are missing among those recommended by AGREE, i.e. relevant clinicians, researchers, policy makers, clinical administrators, and funders, methodology experts (e.g. systematic review expert, epidemiologist, statistician, library scientist);
- statements are not made regarding type of strategy used to capture patients'/public's' views and preferences (eg, participation in the guideline development group, literature review of values and preferences), methods by which preferences and views were sought (eg, evidence from literature, surveys, focus groups), outcomes/information gathered on patient/public information, description of how the information gathered was used to inform the guideline development process and/or formation of the recommendations.

This situation is regrettable because it is well established that multi-specialty groups can take into account a wider range of knowledge and opinion than single-specialty groups, while allowing members of the team to learn from each other (9). The multidisciplinary nature of the guideline development team, if associated with formal consensus development methods, may also be a good way of reducing professional prejudice or controversy and conflicts of interest (see below domain 6).

An even more worrying finding is that all of the guidelines obtain scores barely above average for domain 3 (Rigour of development) which is perhaps the more important step in evidence-based guideline development (10). Such low scores may be due to a lack of methodological expertise in guideline developing teams or a lack of resources needed to perform a well-documented systematic literature search. Another explanation for low scores on rigour of development is that the methods used are poorly reported in the guidelines. This could be improved by using addenda for including search strategies, literature selection process or evidence tables. In electronic documents, hyperlinks to these addenda and to methodology sections can be helpful (1).

Regarding domains 4 (Clarity and Presentation) and 5 (Applicability), although, as stressed above, Rémic guidelines and CUMITECH obtain better scores than the average scores obtained by other guidelines in other fields of medicine (Table 1) there are still some important problems with these domains.

Regarding domain 4 in particular it is often difficult to understand what are the purpose of the recommended actions (e.g. to improve quality of life, to decrease side effects, etc), or what are the patients or conditions for whom the recommendations would not apply, or what are the other possible options (see also the PICO examples above).

Scores in domain 5 (Applicability) are also quite worrying because it is often difficult to understand what types of facilitators and barriers were considered. The low scores on 'Applicability' could be due to considering guideline development and guideline implementation as separate activities. Guideline development groups may feel that it is not for them to discuss potential organisational barriers and cost implications in applying the recommendations. Organisations need to integrate these issues, and involve professionals with the relevant expertise, in the early stages of the development of the guidelines. At least guidelines should clearly inform users about the need to consider these issues (barriers, costs, indicators or criteria for monitoring) locally when implementing or adapting a particular guideline.

Scores obtained in domain 6 (Editorial independence) are perhaps the most worrying of all because they suggest that the guideline development teams are not aware of the ample evidence available that shows that individuals' judgments on best practice are often biased and can represent self-interest, beliefs and prejudices based on one's own experience (9). The motivations for this behavior can be financial (e.g. drug companies, the diagnostics industry, granting or sponsoring bodies, health insurance organizations), academic or administrative (e.g. pressure to publish, pressure to increase or to decrease the work-flow of departments), political (e.g. pressure from patients advocates), personal (e.g. pressure from friends and colleagues, intellectual passion, will-power, vanity), and cultural or educational (e.g. religion). To allow users of guidelines to judge the trustworthiness of a guideline, sponsorship or the lack of it, together with the names, affiliations, and the precise role of all contributors to the guideline development process should be stated explicitly. Even if such precautions are taken, it is extremely

MICHÈLE FONFRÈDE, JEAN PAUL COUAILLAC, CHRISTINE AUGEREAU, JEAN PIERRE LEPARGNEUR, JOSEPH WATINE

CRITICAL APPRAISAL OF MICROBIOLOGY GUIDELINES ENDORSED BY TWO PROFESSIONAL ORGANISATIONS: SOCIÉTÉ FRANÇAISE DE MICROBIOLOGIE (SFM) AND AMERICAN SOCIETY OF MICROBIOLOGY (ASM)

difficult, if not impossible, for a guideline development team to get rid of all their conflicts of interest and prejudices. Therefore some subjective element of judgment, that can distort the evidence, is bound to be present in most guidelines. Multidisciplinarity and formal consensus development methods can reduce this element and make guidelines more transparent (see domains 2 and 3 above).

In conclusion a higher degree of multi-disciplinary work, including a more formal implication of methodologists, as well as of infectiologists, and of economists, might perhaps enable future editions of those microbiology guidelines to reach higher levels of quality. This is mainly why we believe that these guidelines need to be modified before they can be recommended without provisos.

References

1. Alonso-Coello P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, Tort S, Bonfill X, Burgers J, Schunemann H (2010). The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. Qual Saf Health Care 19:e58.
2. AGREE Collaboration (2001). Appraisal of Guidelines for Research and Evaluation (AGREE) Instrument. www.agreecollaboration.org/
3. Société Française de Microbiologie (2007). Rémic: référentiel en microbiologie médicale - bactériologie et mycologie, de la Société Française de Microbiologie (SFM), 3ème édition, Vivactis Plus Ed, Paris.
4. Groupe de travail "Guides de pratique clinique (guidelines) et Evidence-Based Laboratory Medicine" de la Société Française de Biologie Clinique (2011). Evaluation de la qualité méthodologique du Rémic (référentiel en microbiologie médicale - bactériologie et mycologie) de la Société Française de Microbiologie (SFM), 3ème édition (2007). Ann Biol Clin 69:239-245.
5. Société Française de Microbiologie (2010). Rémic: référentiel en microbiologie médicale, 4ème édition. Société Française de Microbiologie (SFM), Paris.
6. AGREE Trust (2009). Appraisal of Guidelines for Research and Evaluation (AGREE) Instrument. http://fhswedge.csu.mcmaster.ca/pebc/agreetrust/
7. American Society of Microbiology (2009). Cumitech 2C: Laboratory Diagnosis of Urinary Tract Infections, American Society of Microbiology (ASM), Washington. http://estore.asm.org/
8. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ (2011). GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 64:395-400.
9. Hutchings A, Raine R (2006). A systematic review of factors affecting the judgments produced by formal consensus development methods in health care. J Health Serv Res Policy 11:172-9.
10. Oosterhuis WP, Bruns DE, Watine J, Sandberg S, Horvath AR (2004). Evidence-Based Guidelines in Laboratory Medicine: Principles and Methods. Clin Chem 50:806-818.