

# Factors Associated with the Quality and Transparency of National Guidelines : A Mixed-Methods Study

---

**Kovačević, Tanja; Vrdoljak, Davorka; Jurić Petričević, Slavica; Buljan, Ivan; Sambunjak, Dario; Krznarić, Željko; Marušić, Ana; Jerončić, Ana**

*Source / Izvornik:* **International Journal of Environmental Research and Public Health, 2022, 19**

**Journal article, Published version**

**Rad u časopisu, Objavljena verzija rada (izdavačev PDF)**

<https://doi.org/10.3390/ijerph19159515>

*Permanent link / Trajna poveznica:* <https://urn.nsk.hr/urn:nbn:hr:266:435088>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2025-01-22**



**Klinički  
bolnički  
centar  
Split**

*Repository / Repozitorij:*

[University Hospital of Split Digital Repository](#)





Article

# Factors Associated with the Quality and Transparency of National Guidelines: A Mixed-Methods Study

Tanja Kovačević <sup>1,2</sup>, Davorka Vrdoljak <sup>3,4,†</sup>, Slavica Jurić Petričević <sup>5</sup>, Ivan Buljan <sup>3,6</sup>, Dario Sambunjak <sup>7</sup>, Željko Krznarić <sup>8,9,10</sup> , Ana Marušić <sup>3,6</sup>  and Ana Jerončić <sup>3,6,\*</sup>

- <sup>1</sup> Department of Pediatrics, University Hospital of Split, 21000 Split, Croatia; tkovacevic@kbsplit.hr  
<sup>2</sup> Department of Pediatrics, University of Split School of Medicine, 21000 Split, Croatia  
<sup>3</sup> Cochrane Croatia, University of Split School of Medicine, 21000 Split, Croatia; davorka.vrdoljak@mefst.hr (D.V.); ibuljan@mefst.hr (I.B.); ana.marusic@mefst.hr (A.M.)  
<sup>4</sup> Department of Family Medicine, University of Split School of Medicine, 21000 Split, Croatia  
<sup>5</sup> Department of Pulmology, University Hospital of Split, 21000 Split, Croatia; slavica.juric01@gmail.com  
<sup>6</sup> Department of Research in Biomedicine and Health, University of Split School of Medicine, 21000 Split, Croatia  
<sup>7</sup> Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, 10000 Zagreb, Croatia; dsambunjak@gmail.com  
<sup>8</sup> Department of Internal Medicine, Division of Gastroenterology and Hepatology, University Hospital Centre Zagreb, 10000 Zagreb, Croatia; zeljko.krznaric1@zg.ht.hr  
<sup>9</sup> Croatian Medical Association, 10000 Zagreb, Croatia  
<sup>10</sup> University of Zagreb School of Medicine, 10000 Zagreb, Croatia  
\* Correspondence: ajeronci@mefst.hr  
† Deceased.



**Citation:** Kovačević, T.; Vrdoljak, D.; Petričević, S.J.; Buljan, I.; Sambunjak, D.; Krznarić, Ž.; Marušić, A.; Jerončić, A. Factors Associated with the Quality and Transparency of National Guidelines: A Mixed-Methods Study. *Int. J. Environ. Res. Public Health* **2022**, *19*, 9515. <https://doi.org/10.3390/ijerph19159515>

Academic Editor:  
Paul B. Tchounwou

Received: 29 June 2022  
Accepted: 29 July 2022  
Published: 3 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** We assessed the methodological quality and transparency of all the national clinical practice guidelines that were published in Croatia up until 2017 and explored the factors associated with their quality rating. An in-depth quantitative and qualitative analysis was performed using rigorous methodology. We evaluated the guidelines using a validated AGREE II instrument with four raters; we used multiple linear regressions to identify the predictors of quality; and two focus groups, including guideline developers, to further explore the guideline development process. The majority of the guidelines (N = 74) were developed by medical societies. The guidelines' quality was rated low: the median standardized AGREE II score was low, 36% (IQR 28–42), and so were the overall-assessments. The aspects of the guidelines that were rated best were the “clarity of presentation” and the “scope and purpose” (median ≥ 59%); however, the other four domains received very low scores (15–33%). Overall, the guideline quality did not improve over time. The guidelines that were developed by medical societies scored significantly worse than those developed by governmental, or unofficial working groups (12–43% per domain). In focus group discussions, inadequate methodology, a lack of implementation systems in place, a lack of awareness about editorial independence, and broader expertise/perspectives in working groups were identified as factors behind the low scores. The factors identified as affecting the quality of the national guidelines may help stakeholders who are developing interventions and education programs aimed at improving guideline quality worldwide.

**Keywords:** knowledge; education; public health; national guidelines; guideline development; clinical practice guidelines; methodological quality; focus group

## 1. Introduction

At a national level, evidence-based medicine (EBM) is introduced into clinical practice primarily through clinical practice guidelines (CPGs) [1]. CPGs include “recommendations to optimize patient care based on a systematic review of evidence and an assessment of the benefits and harms of alternative care options” [2]. They have been found to improve treatment outcomes, the cost-effectiveness of a healthcare system, and the maintenance of

its quality standards [3–6]. However, to achieve the goals of improved resource utilization and health care delivery, CPGs must be high-quality in both content and form. Poor-quality CPGs have been shown to negate their potential benefits and can lead to suboptimal, ineffective, or harmful practices [7,8]. Still, because of the complexity and the demanding nature of guideline development (GD) and guideline implementation (GI) processes, the quality of the evidence-based information that is translated into clinical practice varies widely [9–13].

The development of high-quality CPGs, including *de novo*, adopted, adapted, or contextualized guidelines, is required for successful clinical practice guideline implementation and has been identified as critical in closing the gap between international recommendations and realistic best practice in low- and middle-income countries [14]. Studies that have examined the quality of CPGs for a specific clinical topic have found that it varies considerably between countries [9,11,15–19]. However, to gain an insight into the overall quality of national CPGs, a representative sample of national CPGs (all, or a random sample of the CPGs published within a certain time period) should be assessed. So far, few studies have examined an entire body of national guidelines, and they have found that countries, on average, produce either low- [20–23] or high-quality guidelines [24,25]. This means that in countries where guideline developers produce typically poor-quality guidelines, the chances of providing evidence-informed and cost-effective health care are significantly reduced. This is especially concerning given that it is often necessary to make recommendations at a national level due to the uniqueness of each country's healthcare system—in epidemiological, social, organizational, economic, and political contexts—and due to different patients' expectations across countries, making the use of international guidelines in such cases insufficient.

While numerous studies have investigated the barriers and facilitators of the guideline implementation process, the factors influencing the guideline development process are relatively underexplored [26], particularly at a country level.

The mapping of factors that are potentially affecting the global guideline development process, which were inferred from qualitative studies that discussed and evaluated GD steps, revealed that flaws in the process—such as a lower rigor of evidence synthesis; the omission of a transparent and usable summary of evidence, made available for discussion; flaws in considering the evidence, balancing it with panel members' input, and opportunities to interpret the evidence; inappropriate panel member selection or conduct; and conflicts of interest, among other factors—may contribute to the lower quality of a CPG [27]. At a country level, only one study has been conducted, which was in India. This study focused on the GD of a specific set of CPGs and revealed several barriers of the process, in the form of attitudes towards the use of evidence, a lack of methodological capacity, an inadequate governance structure, and funding [28].

Few studies have looked at the correlations between the quality of global and national CPGs, covering a wide range of health-care topics—as measured by the Appraisal of Guidelines for Research and Evaluation II tool (AGREE II) and the guideline characteristics. Burgers et al. studied the CPGs published between 1992–1999 and discovered that the quality of the AGREE domain, “rigor of development”, was related to the level of care (primary/secondary or tertiary/all levels), the scope (prevention/diagnosis, treatment, combination), and the guideline program (part of the program/not part of the program) of the CPGs [29]. The authors also observed that the guidelines developed by government-funded agencies had the highest overall scores across all the domains, although statistical significance was reached only for the “editorial independence” domain. They additionally observed that time had no effect on the AGREE domains, except in the “clarity of presentation” domain. On the other hand, Alonso-Coello et al. examined the CPGs published between 1980–2007 and discovered that time had a significant effect on the quality of all the AGREE II domains, except for the “editorial independence” domain, and that geographic region had an effect on the quality of the “scope and purpose” and the “clarity of presentation” domains [12]. They also found that medical societies scored lower than gov-

ernmental and international institutions in all domains, except “editorial independence”. Finally, Armstrong et al. assessed the CPGs published between 1992 and 2014 and found a significant improvement in the quality over time in all of the domains of the AGREE II tool [30].

With a few exceptions, studies that have assessed the quality of a representative sample of national CPGs did not examine the factors influencing their quality [20–22,24,31–34]. In Argentina, the quality of the national CPGs produced between 1994 and 2004 did not change over time and was not related to the type of guideline dissemination (published vs. not published); the level of healthcare provided by the organization involved in the GD (macro—Ministry of Health; meso—individual health care providers, organizations of providers, and health insurance institutions; or micro—individual health professionals); the type of guideline (prevention vs. treatment/diagnosis CPG); or the scope (national vs. regional/local guidelines) [23]. However, the authors did find that the guidelines produced by institutions belonging to more than one level of the healthcare system had higher scores than the guidelines produced by institutions belonging to only one level, and that prevention CPGs had higher scores than the treatment/diagnosis management guidelines. In China, the authors determined an increase in quality over time for the guidelines produced between 2014 and 2018, with subgroup analyses revealing that both the methodological and reporting quality were generally higher among the CPGs that used evidence grading systems or reported receiving funding [35]. Finally, in Japan, the quality of all the AGREE II domains improved over time for the guidelines produced between 2000 and 2014 [24], whereas for those published between 2018 and 2019, the effect of time was not tested, but the authors discovered that the number of guideline developers, the clinical question ratio, and the adopted guideline development methods (GRADE and Minds vs. others) were significantly related to the quality of the “rigor of development” domain [36].

Identifying the factors that systematically affect the overall quality of national guidelines and, consequently, the quality of health care in a country is of both national and international importance. Such knowledge would provide a solid foundation for identifying the issues in the process of developing CPGs that are shared among different countries, and would help to improve the quality of future guidelines [37]. To gain this knowledge, an in-depth, systematic investigation is needed, which not only assesses the quality of the CPGs in a country by using a representative sample, but also investigates the possible reasons for such quality by interviewing the country’s guideline developers and users, and examines which CPG characteristics are associated with the level of quality.

The main objective of our study was to assess the methodological quality and transparency of all the CPGs published in Croatia up until 2017 and to systematically investigate the factors affecting this quality using a quantitative regression analysis, with guideline characteristics as predictors, and a qualitative analysis with two focus groups, comprising guideline developers and users.

## 2. Materials and Methods

### 2.1. Quantitative Study

We aimed to identify all CPGs developed in Croatia up until 2017 and found 81 documents published as CPGs in the Croatian Medical Association’s (CMA) official journal, *Liječnički vjesnik*, which is a target journal, publishing all Croatian CPGs [38]. Seven documents were excluded because they were narrative reviews, clinical pathways, translations of foreign guidelines, or training manuals.

We used the AGREE II tool to assess the methodological quality and transparency of guidelines [39,40]. The tool was released in 2009 [39], and is now the most widely used and verified guideline evaluation tool. It is comprised of 23 items, grouped into the following six domains: “scope and purpose”, “stakeholder involvement”, “rigor of development”, “clarity of presentation”, “applicability”, and “editorial independence”. It is rated on a 7-point scale (from 1—“strongly disagree”, to 7—“strongly agree”). Based on this ranking, a standardized domain score, ranging from 0% to 100% (low- to high-quality), is computed

for each of the six domains. Standardized overall score (0–100%) can also be reported [24]. AGREE II has two additional global ratings—overall assessments. The overall quality of the guidelines is scored on a 7-point scale (1—“lowest possible quality”, to 7—“highest possible quality”), while the second evaluation provides a recommendation on the guideline use (“yes”, “yes, with modifications”, “no”). In this study, each of the four raters independently rated all 74 CPGs. For more details see Methods—Quantitative study, Additional File S1.

Categorical data distributions were presented with absolute numbers and percentages, and quantitative with median and interquartile range (IQR), due to asymmetrical distribution. Multiple linear regressions were used to determine predictors of CPG’s quality and transparency. The standardized AGREE II scores (overall, and domain scores) were dependent variables in these models, while independent variables were as follows: type of organization leading GD, number of entities/organizations involved in GD, and the year of publication. Conclusions were drawn at a significance level of 0.05, using two-sided tests. Confidence intervals (CIs) of 95% were also considered. SPSS 24.0 software (IBM Corp., Armonk, NY, USA) was used for statistical analysis.

The VOSviewer software was used for visualization of the co-citation network of Croatian guideline authors [41].

### Comparison with Other Countries

We conducted a systematic search in PubMed to find studies that investigated the quality of national CPGs based on a representative sample, using the AGREE tool. Section S1 of the Supplementary Files (Methods—Quantitative research) contains the search filter and details on inclusion and exclusion criteria. If the publication period of the analyzed CPGs included more recent years—2012 and later—median standardized domain scores for a country were extracted from each domain for that study and plotted on the radial graph.

### 2.2. Qualitative Study

To further investigate factors affecting the quality of Croatian guidelines, we conducted two focus groups, including the following: guideline developers (N = 6) and potential guideline users (N = 5). In total, 10 physicians and 1 dentist took part, working for primary, secondary, and tertiary health care organizations or Croatian Insurance Fund. Interviews were conducted by DS at the Psychology Research Laboratory of the Catholic University of Croatia, in Zagreb. The sessions, which lasted 1 h, were audio recorded. The conversations in both groups started with the question, “Why are clinical guidelines important in Croatian healthcare?”, after which we let the discussion run its course. We had prepared additional questions in case of a silence, but the discussions were lively.

The discussions were then transcribed into an electronic document and analyzed using grounded theory approach. We used an iterative process for reading the answers, open coding, and constant comparison to generate themes and patterns [42]. For more details see Methods—Qualitative study, Additional File S1. The checklist for reporting guideline COREQ is provided in Supplementary Table S2.

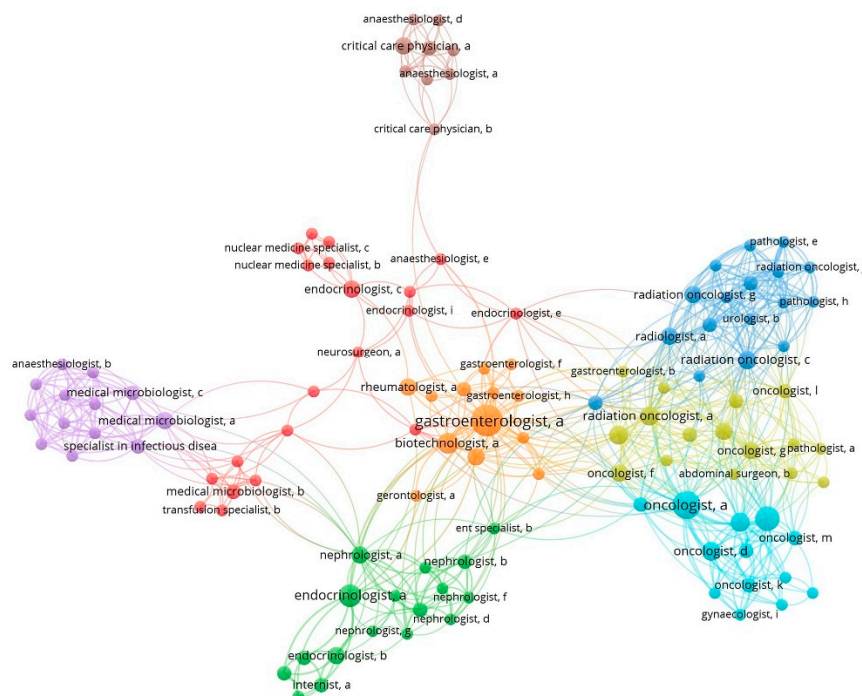
## 3. Results

### 3.1. Quantitative Study

Seventy-four Croatian CPGs were published up until 2017. The number of published CPGs increased linearly with time, from just one (1%) CPG published in 2004, to 12 (16%) in 2017 (see Supplementary Figure S1).

The development of the CPGs was led by CMA’s medical societies (81% of the 74 guidelines), the Ministry of Health working groups (WGs) (7%), and unofficial WGs, comprising physicians-volunteers, and, in one case, a patient association (12%). Five out of 163 CMA societies [43] (two different oncology societies and societies for gastroenterology, urology, and clinical nutrition) were involved in the development of 43% of the guidelines. The authorship landscape of the Croatian guidelines demonstrates considerable intercon-

nection, with eight clusters apparently topic-oriented (three belonging to oncology topics and one to gastroenterology) and some strongly connected authors that also connect to other clusters (Figure 1).



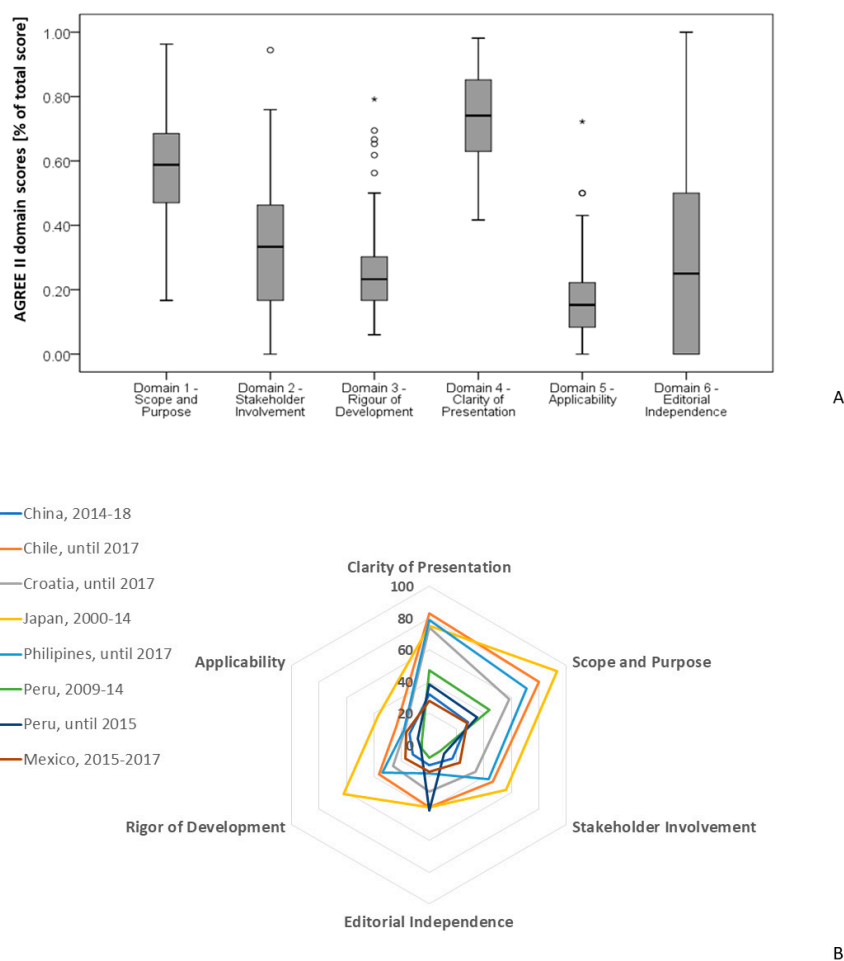
**Figure 1.** The authorship landscape of 74 Croatian clinical practice guidelines (649 unique authors, median—11 authors per guideline). Authors are identified by specialty, followed by an index letter after the comma that uniquely identifies a person. The size of the circle reflects the number of CPGs ( $\geq 2$ ) published by this person.

The CPGs most often targeted neoplasms (38%), followed by endocrine (8%), and infectious (8%) diseases. None of the CPGs focused on children’s health, and no guideline was developed for patients.

Only seven guidelines (9%) envisaged updates, either in a three- or five-year period, or in the case of new scientific evidence. Still, only one guideline from 2011 was updated in 2017, and this update did not specify whether a new update is planned.

Regarding the methodological quality and transparency of Croatian CPGs, the median standardized overall AGREE II score was 36% (IQR 28–42%), with all the guidelines obtaining less than 50%, except for seven (9%) guidelines, which achieved between 63–76%. However, the domains differed considerably in the level of quality that was assessed (see 95% CIs in Supplementary Figure S2), with “clarity of presentation” and “scope and purpose” scoring by far the highest (the median standardized domain scores were 74% and 59%, respectively) and “applicability” scoring the lowest (15%). “Rigor of development”, “editorial independence”, and “stakeholder involvement”, with medians of 23–33%, were comparable (Figure 2A; Supplementary Figure S2). In terms of the variability in the quality assessment, the “applicability” and “rigor of development” ratings were lower and more homogeneous than those of the other domains.

The guidelines’ overall AGREE II assessments on the 7-point scale were mainly rated low to moderate (76% of CPGs received up to four points, see Supplementary Table S1). Overall, the raters did not recommend 31 (42%) of the guidelines, they recommended 38 (51%) with modifications, and only five (7%) were recommended for use in clinical practice.



**Figure 2.** Standardized domain score distribution across six AGREE II domains: (A) among the Croatian clinical practice guidelines, N = 74. Outliers and extreme outliers are represented by circles and asterix, respectively; (B) among countries that used AGREE II to evaluate national guidelines. The period studied and the country of origin are indicated in the legend of the figure. Japan was the only country reported to have high-quality guidelines [24], whereas all others were classified as those with low-quality guidelines [20–22,33–35].

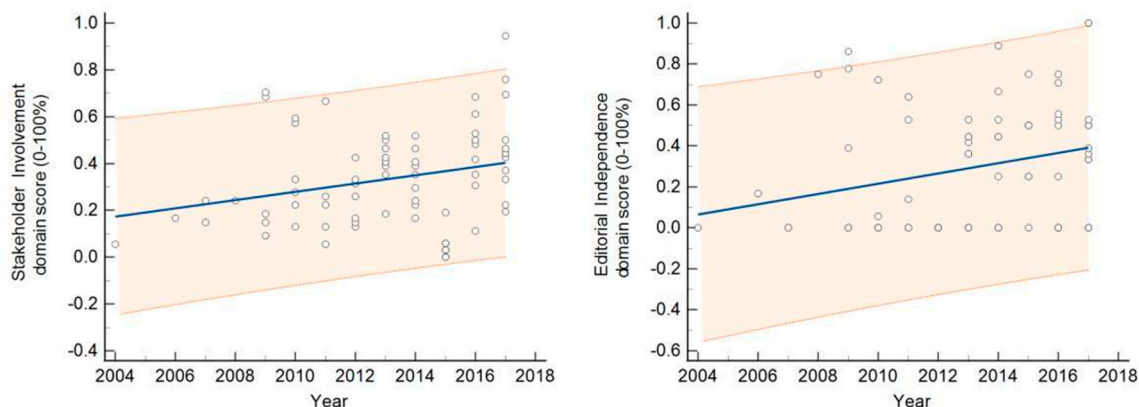
### 3.1.1. Comparison with Other Countries

Figure 2B shows a radial graph comparing the standardized AGREE II domain scores in Croatia to the scores in other countries, where studies analyzed the whole body of national guidelines published during a specified period, assessed the quality using the AGREE II instrument, and reported the domain scores. The findings in Croatia were comparable to the findings in other countries. Except for Japan—which has high-quality guidelines and scored above the cutoff of 60% on the “rigor of development” domain as well as the “clarity of presentation” and “scope and purpose” domains—all other countries developed, on average, low-quality guidelines that followed the same pattern as Croatia. The highest scoring domains were “clarity of presentation” and “scope and purpose”, which were only actually rated as low-quality in those countries with, on average, extremely low overall AGREE scores (<30%). The domain with the lowest score was “applicability”, which was followed by “rigor of development” and “editorial independence”.

### 3.1.2. Predictors of Methodological Quality and Transparency

The predictor of the methodological quality and transparency of the guidelines, measured by the AGREE II standardized score—overall and per domain—was the type of entity/organization that led GD, not the total number of participating organizations

(Table 1). Compared with the guidelines developed by medical societies, the guidelines developed by a governmental entity performed 22–43% better in all the domains of AGREE II. The guidelines developed by unofficial WGs/non-governmental organizations performed 12% better in the “rigor of development” domain. These latter guidelines also appeared to perform better than those produced by medical societies in the areas of “clarity of presentation” and “applicability” by 7–8%, as assessed at the 0.1 significance level. The year of publication was a weak predictor for the domains “stakeholder involvement” and “editorial independence”, increasing 2–3% per year. For all other domains and for the overall score, the quality did not change over the time period (Table 1, Figure 3).



**Figure 3.** Croatian guidelines’ AGREE II scores in the domains “stakeholder involvement” and “editorial independence” increased over time. Shown here are the individual scores, as well as the simple regression line and its accompanying 95% prediction interval.

### 3.2. Qualitative Study

In Table 2, we list the themes identified in the focus group discussions, with representative quotes for each topic. In the paragraph below, we describe the qualitative findings in detail.

#### 3.2.1. Guideline Development—Framework

When discussing which entity should lead the development of Croatian guidelines, both developers and users preferred a bottom-up approach, with medical societies as CPGs’ main developers. One developer stated that even if the topic of the guideline and the call from it came from the Ministry of Health (top-down approach), it was still medical societies that should answer that call and develop them. Reference centers were not seen as feasible in the Croatian system by either developers or users, due to the inability to have a sufficient number of centers and experts to cover all the necessary areas of care in a small country such as Croatia. One top-down approach that was believed possible was the establishment of a dedicated national agency, responsible for developing guidelines.

The topic of prioritizing GD emerged among developers but was not discussed extensively. One developer believed that the criteria for decisions about which CPGs to develop should be based on disease prevalence and another believed that the criteria should be based on affordability and the complexity of the guidelines.



**Table 1.** Effects of different factors on quality of Croatian guidelines, shown by AGREE II domains.

Factors	AGREE II: All Domains		AGREE II Domains											
	Overall		Scope and Purpose		Stakeholder Involvement		Rigor of Development		Clarity of Presentation		Applicability		Editorial Independence	
	B *	p-Value	B	p-Value	B	p-Value	B	p-Value	B	p-Value	B	p-Value	B	p-Value
Year	-	0.292	-	0.522	0.02	<0.001	-	0.726	-	0.131	-	0.421	0.03	0.005
Total No. of entities	-	0.594	-	0.427	-	0.79	-	0.156	-	0.58	-	0.833	-	0.996
Entity primarily responsible for CPG's development														
Led by unofficial WGs vs. led by medical society(ies)	0.08	0.017	-	0.699	-	0.138	0.12	0.002	0.08	0.065 †	0.07	0.050 †	-	0.466
Led by governmental health WGs vs. led by medical society(ies)	0.37	<0.001	0.37	<0.001	0.42	<0.001	0.43	<0.001	0.22	<0.001	0.35	<0.001	0.39	0.003

B—non-standardized regression coefficient, obtained by generalized linear models using year, total number of entities, and type of entities responsible for guideline development as independent variables; WGs—working groups. \* Values for B are shown only for significant association; † association significant at 0.1 level.

**Table 2.** Findings of the focus group discussions, organized by topic, with representative quotes by participants.

Topic	Examples of Quotes
<b>Guideline development framework—Who should lead GD?</b>	
Bottom-up approach	<ul style="list-style-type: none"> <li>• Medical societies, “It should be the medical societies.”</li> <li>• Applications to a tender, “The Ministry of Health should issue a tender for the development of guidelines . . . Trendsetters and medical societies will apply . . . ”</li> <li>• Reference centers (judged as unfeasible in Croatian system), “There are too few reference centers for us to cover all areas of care . . . keep in mind that a reference center is made up of three, five, eight experts, and we have to think of all the other experts who work in other institutions.”</li> </ul>
Top-down approach	<ul style="list-style-type: none"> <li>• “[national] agency dedicated to guideline development.”</li> </ul>
<b>Guideline development framework—criteria for prioritization of GD</b>	
Disease prevalence	<ul style="list-style-type: none"> <li>• “If we start from patients who stand to benefit the most, we may end up developing guideline for only five patients.”</li> </ul>
Expected costs and complexity of development	<ul style="list-style-type: none"> <li>• “Those are small guidelines . . . So, guidelines we can afford.”</li> </ul>
<b>Guideline development framework—composition of a working group</b>	
Professional medical societies and medical experts	<ul style="list-style-type: none"> <li>• “Working group should consist of medical societies and experts”, while noting that “experts for certain diseases are easy to find.”</li> </ul>
Diverse teams of various medical specialists and a representative of the insurance company	<ul style="list-style-type: none"> <li>• “Teams are important and a team developing clinical guidelines should consist of, for example, a diagnostician, a radiologist, an oncologist and possibly an insurance company representative. In my opinion, it would be good if they came from different centres.”</li> </ul>
Experts and patients	<ul style="list-style-type: none"> <li>• “Maybe for this reason [in relation to the discussion about patients’ rights in Croatia, author’s note], patients should also be involved in the development of guidelines, so that they also understand the situation and take some of the responsibility.”</li> </ul>
Other health professionals as part of a working group	<ul style="list-style-type: none"> <li>• [talking about making recommendations], “If it has elements to include some resources that are needed—then other professions participate in drawing conclusions too.”</li> </ul>

Table 2. Cont.

Topic	Examples of Quotes
<b>Guideline development methodology—methodology</b>	
Methodology is inadequate	<ul style="list-style-type: none"> <li>“There are only a small number of guidelines that have been published so far in our country that have followed the correct methodological steps that should be followed in developing guidelines. . . . I would like to reiterate that most of our guidelines are methodologically inadequate.”</li> </ul>
Description of a typical GD process in Croatia reveals that CPGs are mainly expert opinion-based	<ul style="list-style-type: none"> <li>“When guidelines are developed, a group of people are brought together who are involved in a particular medical field or who work in some way with medical professionals in that field. For example, when guidelines were developed for the treatment of intraepithelial lesions of the cervix, gynecologists, pathologists, and cytologists were represented in this group, not only from Zagreb, but also from other centers. And then we met and discussed, each from the side of his profession, what to do in a particular situation. Of course, we did not make recommendations out of thin air, one part was based on our experience, another part was based on foreign experience that we had gathered at some meetings, congresses and so on. And that actually took quite a while before we as a working group published these guidelines.”</li> </ul>
<b>Guideline development methodology—knowledge on GD methodology *</b>	
Developers lack methodological expertise	<ul style="list-style-type: none"> <li>[developer], “Developers are not well-versed in the methodology of guideline development.”</li> </ul>
Definition of GD process *	<ul style="list-style-type: none"> <li>EBM and alignment with resources [user], “Guidelines should first be based on evidence-based medicine and then we should see how they can be aligned with our resources.”</li> <li>Evidence synthesis, pharmacoeconomic evaluation, strength of evidence, and consensus * [developer], “To develop a guideline, you need a meta-analysis, a systematic review for each diagnostic or therapeutic procedure, then you should include a pharmacoeconomic evaluation and then you should focus on strength of evidence and consensus. This process should be repeated several times until we agree on the final guidelines.”</li> <li>Clinical knowledge, resources, and preferences/values * [developer], “This implies that when we develop a guideline, we first identify which process doesn’t ‘breathe’ well for us, we identify the critical points in it, and then we determine if we have a problem at a critical point with: what is clinical knowledge, what some resources are, and what values or preferences we have. And then we get the take-home message that the methodology by which the guideline is created is actually to determine one of these elements that is dominant.”</li> </ul>

Table 2. Cont.

Topic	Examples of Quotes
Overconfidence on methodological knowledge **	<ul style="list-style-type: none"> <li data-bbox="810 367 1485 770">• [developer] Judgement about the methodological quality, made based on the recognition of a few topic-related elements by a developer of low-quality, expert opinion-based CPGs **, “I couldn’t physically attend all of these meetings [referring to working group meetings], but people who were more involved than me, in the sense that they knew what they were doing methodologically—I got the impression that they did a very good job and that they knew S1 and S2 and S3, so about the types of guidelines, when certain guidelines are issued, how much weight they have, and so on. So, while I didn’t have the task to do something methodologically, I got the impression that these people had done their homework thoroughly and knew what they were talking about.”</li> <li data-bbox="810 770 1485 860">• [user] Underestimated complexity of the GD methodology **, “If someone asks you to develop a guideline, you will find a way because everything is available today.”</li> </ul>
<b>How to improve GD process in Croatia</b>	
Increase the level of knowledge of GD methodology for members of WGs	<ul style="list-style-type: none"> <li data-bbox="810 949 1485 1061">• “The people involved in the development of guidelines are not sufficiently familiar with the methodology of this process. I think that they need the help of an expert to explain this methodology to them to apply it correctly.”</li> </ul>
Increase motivation of clinicians involved, possibly through financial incentives	<ul style="list-style-type: none"> <li data-bbox="810 1106 1485 1256">• “The GD process itself is very demanding. When we talk about the clinicians involved in it, they are not . . . how shall I say . . . ? I am not going to say they should be paid for it; that may sound harsh. However, they are not motivated enough for such an action.”</li> </ul>
Determine who should lead GD	<ul style="list-style-type: none"> <li data-bbox="810 1301 1485 1352">• “What is very important to say is who is in charge of writing these guidelines . . . ”</li> <li data-bbox="810 1352 1485 1442">• “Medical societies should be supported and motivated to lead GD in Croatia . . . GD should not be an obligation for professional societies—it would not work.”</li> </ul>
Map all guidelines currently in use	<ul style="list-style-type: none"> <li data-bbox="810 1487 1485 1637">• “Before prioritizing the development of new guidelines, all guidelines currently in use in Croatia should be identified. This would require screening all hospitals and interviewing the experts in those hospitals to find out which international guidelines they already use.”</li> </ul>
Allocate resources	<ul style="list-style-type: none"> <li data-bbox="810 1682 1417 1711">• “Resources—time and money should be provided.”</li> </ul>

Table 2. Cont.

Topic	Examples of Quotes
<b>Definition of clinical guidelines</b>	
CPGs are evidence-based	<ul style="list-style-type: none"> <li>“First, our guidelines must be based on evidence-based medicine . . . ”</li> </ul>
CPGs are a set of rigid and formal rules that must be followed *	<ul style="list-style-type: none"> <li>“ . . . we believe that a guideline is set in stone.” *</li> </ul>
CPGs are standard operative procedure documents *	<ul style="list-style-type: none"> <li>“In our work, we use many clinical guidelines; only we do not call them that, we call them standard operative procedures.” *</li> </ul>
CPGs are clinical pathways *	<ul style="list-style-type: none"> <li>“You need to know the reason why you are moving a patient from a smaller hospital to a larger one.” *</li> </ul>
<b>Purpose of clinical guidelines</b>	
To establish a standard of care	<ul style="list-style-type: none"> <li>“So that we can ensure that every doctor in family medicine works at the same level, that pathologists read the findings the same way, and that rheumatologists treat the patient the same way.”</li> </ul>
To mitigate potential risks arising from a physician’s conflict of interest in treatment selection	<ul style="list-style-type: none"> <li>“Guidelines enable us to unify a system comprised of thousands of physicians with diverse personal interests.”</li> </ul>
To move some of the responsibilities for patient care decision making away from physicians and toward guidelines	<ul style="list-style-type: none"> <li>“Guidelines shift some of the responsibility for patient care decision-making from physicians to guidelines.”</li> </ul>
To educate clinicians about the optimal management of patients	<ul style="list-style-type: none"> <li>“If there were guidelines on how to treat a patient in certain circumstances, then people would be educated on how to solve the problem in that hospital . . . ”</li> </ul>
To achieve some level of control in the healthcare system *	<ul style="list-style-type: none"> <li>Quality of care control, “A guideline arises from how we can control the quality of work . . . ” *</li> <li>Better control and optimization of business processes, “The purpose of guidelines is to introduce order into a healthcare system. This is to be achieved indirectly by integrating CPGs into digital healthcare information systems, as such integration requires that business processes are defined.” *</li> <li>Better cost control, “At the end, the aim of a guideline implementation is to achieve target savings.” *</li> <li>Control over patients’ billing [while commenting, billing of patients who misused the ER services], “ . . . once we implement the guidelines, we will have some control over it, and this problem should be solved.” *</li> </ul>

A topic or a quote demonstrate: \*—misconception, or \*\*—overconfidence. For clarification please see the main text Sections 3.2.3 and 3.4.

### 3.2.2. Guideline Development—Composition of Working Groups

Concerning the composition of a WG for GD, users preferred a combination of “medical societies and medical experts”. They also identified the need for a broader range of experience and perspectives, listing as prospective WG members, various medical special-

ists and insurance company representatives, or patients. One guideline developer stressed the need to have “an expert”, with experience in the methodology of GD.

### 3.2.3. Guideline Development—Methodology

Participants from both groups identified inadequate methodology as a barrier to the development of high-quality CPGs. When developers described the steps that WG members normally follow, they revealed that, commonly, no systematic, evidence-based methodology was used, and that CPGs were based on expert opinion.

A lack of knowledge was seen as the major reason for weak methodology. While describing the GD process, two developers associated evidence only with the existence of meta-analysis and considered only the strength of evidence and pharmacoeconomic considerations as important factors for making recommendations. Furthermore—in opposition to the order recommended by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) handbook [44], and those cited in the AGREE II tool [39]—they stated that pharmacoeconomic considerations should be considered before assessing the strength of evidence, or they confused the GD with the GI steps—they described that when developing a guideline, developers first determine which “process does not ‘breathe’ well” and then identify “critical points in it”, pointing out that the methodology of guidelines is to eventually establish which “element (clinical knowledge + resources + preferences/values) is dominant”. However, it is a predefined clinical question, not a process, for which one should assess the quality of evidence and decide on the strength of the recommendation by considering all the factors: health benefits, side effects, and risks.

One developer of expert-opinion guidelines believed that their colleagues did a very good job methodologically. A user further stated that formal education or specialized training is not needed for GD, because anyone can self-educate using free (online) resources (Table 2, Knowledge on GD Methodology).

### 3.3. How to Improve the GD Process in Croatia

Developers suggested several approaches to improve GD in Croatia, including the following: increasing the level of knowledge of GD methodology among members of WGs, increasing the level of motivation of the clinicians involved, determining who should lead GD, mapping all CPGs currently in use before prioritizing the GD of new guidelines, and allocating enough resources in terms of time and money.

### 3.4. Definition and Purpose of CPGs

Participants had different definitions regarding what a CPG is. While some identified them as evidence-based documents, others defined them as being sets of rigid or formal rules, standard operating procedures (SOPs), and even clinical pathways. Users also perceived them to “limit the freedom” of practitioners, calling them “cookbook”, “traffic control”, or “just burdens that could not make a difference in their work environment”.

Participants in both groups identified establishing a standard of care as the primary purpose of CPGs, but they also perceived CPGs to mitigate physicians’ conflict of interest in treatment selections, transfer decision-making responsibilities from physicians to the guidelines, and to educate physicians on optimal patient management. Some also expressed that CPGs are intended to provide some form of quality control, to improve the control and optimization of the healthcare business processes, or to lead to a better control of costs or payment billing.

## 4. Discussion

Our analyses have shown that the first Croatian guideline was developed in 2004, 13 years after Croatia’s independence, and, since then, the number of guidelines has steadily increased, reaching 74 in 2017. As there was not, and still is not, a central body overseeing GD in Croatia, national guidelines were most often developed by medical societies, grouped under the CMA.

The overall methodological quality and transparency of the national guidelines was poor. The standardized median overall score of AGREE II was 36% and almost all guidelines scored less than 50%, leaving much room for improvement. Tudor et al. studied seven Croatian neurological guidelines, which were published up until 2012, using the AGREE II instrument [45] and also reported that the guidelines largely lacked a structured development process and had a poor quality in the following domains: “applicability”, “editorial independence” and “stakeholder involvement”. Tokalic et al. found inadequate reporting of 24 Croatian and relevant transnational CPGs published between 2014 and 2016, using the RIGHT reporting checklist [46].

In line with other countries, the median standardized domain scores for “clarity of presentation” and “scope and purpose” were the highest, and their quality was deemed acceptable. In fact, these domains only received a low-quality rating from countries that were producing CPGs with an extremely low overall AGREE II score (<30%): China, Mexico, and Peru [20,33–35]. Other domain scores for Croatian CPGs were very low, with “applicability” ranking the lowest, followed by “rigor of development” and “editorial independence”. Similar results were observed in other countries that generally produced low-quality guidelines [20–22,33–35], regardless of the guideline framework used, indicating a common underlying problem. The largest difference between the countries producing high- and low-quality CPGs was observed in the “rigor of development” domain. Japan, the country producing high-quality guidelines, also had the “rigor of development” domain rated as high-quality (median standardized domain scores of 62%) [24], which was surpassed only by Australia, whose average rating for 15 guidelines following the GRADE methodology was 71% [25].

#### 4.1. Change in Quality over Time

Except for “stakeholder involvement” and “editorial independence”, which have grown by between 2–3% every year, the ratings for the AGREE II domains had not changed over the past 13 years. In Argentina, another country that produced poor-quality CPGs between 1994 and 2004, the methodological quality has not improved in any AGREE II domain, despite an increase in the number of guidelines over time [23]. In contrast, in China, the authors reported an increase in quality over time for the guidelines produced between 2014 and 2018, from 5% (“clarity of presentation”) to 13% (“scope and purpose”) [35]. A systematic review of the studies that aimed to evaluate the CPGs published globally between 1980 and 2007 found that their overall quality improved with time. However, only the ratings for the following three AGREE II domains have increased since 2003, when the instrument was first published: “scope and purpose”, “stakeholder involvement”, and “clarity of presentation”. However, the scores for “rigor of development”, “applicability”, and “editorial independence” remained unchanged. The authors were primarily concerned by the fact that the mean score for “rigor of development” was just 43% (95% CI 41–45), as they considered this category to be the most important indicator of the guidelines’ methodological quality. A fundamental issue, according to the authors, was a scarcity of methodological professionals and methodological expertise so they recommended that *de novo* guidelines should only be developed when there are minimal requirements for them [12]. A considerable improvement in the quality of the “rigor of development” domain in Japan has been linked to advances in social guideline infrastructure, rather than simply to the release of a GD handbook [24]. While the lack of a time trend in the quality improvement may be due to the short period in which published CPGs were observed, the absence of such a trend over longer periods in a country producing low-quality guidelines implies that activities aimed at improving the quality of national CPGs are either nonexistent or deficient.

#### 4.2. Factors Affecting the Quality of Croatian CPGs Guideline Framework

Our study found that even though medical societies led the bulk of GD, their guidelines were, on average, of a lower quality than those led by the Ministry of Health (worse on all domains) or informal WGs (worse only on the “rigor of development” domain). This is an interesting finding, as most of our focus group participants said they believe that medical societies should be leading national GD. Therefore, more education and guidance, for example, the inclusion of methodological experts, as one focus member suggested, but also of other experts, such as information specialists and economists, might be an approach to improve their quality. Another approach may be to simply have external experts evaluate the guidelines and suggest improvements before they are published.

While the findings from other countries that are generally producing low-quality guidelines show that simply choosing a national agency to oversee the GD process [20,22] over the medical/scientific societies [21,23] does not guarantee the production of high-quality CPGs, there is no doubt that the current GD framework should be improved further. Other nations’ experiences, such as those of Japan [24] and Australia [25], suggest that the simultaneous, coordinated actions of multiple organizations may increase the quality of national CPGs. In addition, as emerged during discussions, the issues that should be addressed at a framework level include the following: the prioritization of GD at a national level, which is currently suboptimal, preceded by the mapping of all the guidelines currently in use (guidelines published in the official CMA journal, and international guidelines that may or may not have been translated into Croatian, and, according to one participant, “guidelines that everyone tacitly follows but have never been formalized”); a resource distribution plan (time and finances); and greater clarity of responsibilities (“who is in charge”, “medical societies should be supported and motivated to lead . . . but GD should not be an obligation for them”).

#### 4.3. Inadequate Methodology of GD

Consistent with the low AGREE II scores and the above-mentioned framework findings, developers and users repeatedly identified inadequate methodology as a significant barrier to the development of high-quality guidelines.

That the methodology was, indeed, inadequate was also revealed by the descriptions of a typical GD process given by developers throughout the discussion. These descriptions highlighted that WGs do not commonly follow a systematic, evidence-based methodology and that recommendations are mostly dependent on expert opinion. While experts may have a thorough comprehension of the available research base, they may also express opinions that are unsupported by independent evidence evaluations. Expert opinion alone risks offering selective, out-of-date, or biased perspectives of existing evidence [47,48]. Even in cases when there are evidence gaps, expert opinion should be gathered via rigorous qualitative research, such as Delphi surveys, to support the available evidence claims [49].

The lack of methodological knowledge among the WG members is supported by the quality of the Croatian guidelines, which were developed by utilizing approved AGREE or GRADE [44] methodologies (AGREE: four developed by Ministry of Health WGs and one by a medical society; GRADE: two developed by medical societies and two by unofficial groups). We showed that the utilization of these methodologies is not a guarantee of high-quality. All the Ministry of Health guidelines were rated high, including the one that did not use methodologies, while all the CPGs of the other organizations that used them, ranked low. That the handbook alone is not enough was also identified by Seto et al. [24].

As evidenced by statements about “Developers not being well-versed in the methodology . . . ” and the misconceptions and overconfidence in methodological knowledge that we observed, a lack of methodological expertise among the developers surfaced as a likely explanation for the insufficient methodology. When the level of understanding of the topic is not adequate a fertile ground for knowledge overconfidence forms, which is commonly explained by the Dunning–Kruger phenomenon [50–52]. The phenomenon implies that



less experienced persons are not entirely aware of the actual complexity of the topic being addressed and, as a result, they are unaware of how much knowledge is required to grasp all the aspects of the topic and, thus, consider themselves fully educated and trained before they are. Such cognitive bias may be a barrier to the future methodological training of developers, since, to engage in voluntary self-improvement it is important to acknowledge the gaps in one's own knowledge and abilities, as well as to comprehend how one is seen by others [52].

Despite the participants acknowledging the lack of methodological knowledge, one of the users believed that EBM education during medical school or residency is not needed to acquire this knowledge as it could be acquired in other ways, while another believed the best way was to learn during the development and to have a methods expert in the working group. The question remains as to how this expert should be trained—perhaps a dedicated master's program or online training, rather than training for all medical students.

The developers also cited a lack of motivation among the clinicians involved in GD, which they attributed mostly to a lack of financial incentives and because they viewed GD as a resource-intensive process, and a lack of resources in the form of time and money as reasons for the inadequate methodology.

#### *4.4. Involvement of Other Experts and Stakeholders in Working Groups*

In addition to the Croatian guidelines being rated poorly in the “stakeholders involvement” domain, we found that Croatian WGs were comprised nearly entirely of physicians and seldom of any other healthcare professionals. A lack of association between the quality of the guidelines and the number of participating organizations (reflecting the number of experts) shows that the amount of medical expertise is not sufficient for the development of high-quality guidelines.

While none of the participants fully acknowledged the importance of multidisciplinary and diverse perspectives in working groups, the medical specialist-centric viewpoint appears to be slowly changing, as evidenced by the yearly increase in the AGREE II scores for the Stakeholder-Involvement domain and the comments about inclusion of patients (the reason for including patients was to “share responsibility,” not to broaden perspectives), or insurance company representatives in working groups. Still, the existing composition of WGs may have had an impact on the most critical step of GD—assembling an effective WG—by excluding stakeholders or experts, such as other healthcare professionals, other professions, patients and/or carers, and policy makers [53,54].

#### *4.5. Editorial Independence*

“Editorial independence” was also rated low by using AGREE II but has improved very slowly over time. Still, the topic did not emerge during the focus groups discussions. Other studies have found that national guidelines generally fail to reveal funding sources and conflicts of interest [12,24,55,56]. In countries with low-quality guidelines, the median standardized score for the domain ranged from 0–39%. This score was 39% in Japan, while in Switzerland, conflicts of interest statements were submitted in 44% and declarations of financial support in 29% of the Swiss CPGs published between 2008 and 2019 [12,20,23,24,55,57]. It should be examined further why the ethics of guideline development develops slower than the ethics of research publication [58].

#### *4.6. Applicability*

In Croatia there is no system for CPG implementation in health care. After publishing CPGs in the CMA's journal and eventually providing some dissemination and education activities, there was no further implementation plan. We do not know how many of the published guidelines that Croatia attempted to implement, but, based on the focus group discussions, implementation attempts were rare and unsuccessful. However, this is not an isolated issue; a similar problem has been observed in Norway [59].

#### 4.7. Methodological Knowledge, Misconceptions, and Overconfidence among Guideline Developers and Users

We also discovered several misconceptions regarding the definition and purpose of guidelines, which may have influenced both developers' and users' expectations, attitudes, and even educational interventions.

According to one user, "all EU rules are evidence-based", as opposed to the Croatian ones, which are thought to be based on expert opinion. It is crucial to note, however, that evidence-based does not always imply high quality, as this is dependent on the available data [60].

Developers and users also confused CPGs with SOPs, or even clinical pathways. SOPs are tailored to specific work processes, equipment, or conditions and explain how to safely perform a repeatable work process or procedure, as directed. They are written in a step-by-step format for the workers performing the process, similar to a cookbook. This misconception may be reflected in participants' descriptions of guidelines as "cookbooks" or "traffic regulations" that "place the decision in the hands of a physician", and may have also led to the misconception that guidelines are a set of rigid and formal rules that must be followed. Unlike SOPs, which are limited to a specific process or procedure, CPGs are broad recommendations, which set out general principles that are open to interpretation and are applicable to a wide range of situations. Other studies have also found that clinicians believe that recommendations are categorical, prescriptive, and that they restrict professional practice, and have interpreted their findings to reflect a lack of subject knowledge and awareness of how guidelines are formed [61].

When it comes to the purpose of guidelines, developers identified the primary objective as establishing a standard of care to eliminate unwarranted variability in treatment, whereas users also recognized other purposes. Surprisingly, no one suggested improving patient care as a goal.

Some participants, including developers, have mistakenly assumed that the purpose of CPGs is control—control of the quality of care, and the control and optimization of healthcare business processes, costs, or even patient billing. While CPGs have the potential to improve a healthcare system's cost-effectiveness and the maintenance of its quality standards [3–6], this is achieved indirectly, through successful GI. Such misconceptions may, however, cause high expectations regarding GI, which ultimately leads to disappointment and a loss of confidence in the concept of CPGs [62].

#### 4.8. Limitations

The study's limitation is associated with using the AGREE II tool. The tool assesses how specific items within various domains were reported, which may be biased if an item was included/performed during GD but not reported. Furthermore, we applied the commonly used standardized cut-off score of 60% [63–65], although no standard cut-off scores have been established to distinguish between high- and low-quality standards. Nonetheless, the tool has been validated in many clinical domains and by various specialists [66], and in all the guideline cases evaluated, information meant to be of high-quality was scored higher than the content designed to be of low-quality [66]. Furthermore, in our analysis, the median standardized scores for low-scoring domains clustered around low values, with a difference of approximately 20% between low- and high-scoring domains clearly distinguishing the groups.

We examined the guidelines up to 2017. Given that the guideline framework has not changed in the meantime, and that the COVID-19 pandemic temporarily impacted GD activity, we believe that the data reported accurately represent the state of the Croatian guidelines in a pandemic-free setting. Considering that a 14-year time span had little effect on the quality of CPGs in Croatia, and that the guideline framework has not changed, we do not expect the CPGs published in pre-pandemic 2018 (N = 6) and 2019 (N = 8) to have a significant impact on the results.

Another possible constraint of the study was the participants' discontent with the day-to-day challenges of organization and resource management in the Croatian healthcare system, which required the moderator to constantly divert or refocus the focus groups. In this setting, more structured questions might aid with the identification of additional barriers. Nonetheless, the discussions accurately mirrored reality and supported many of our quantitative findings.

## 5. Conclusions

Croatia, like other countries that produce, on average, low-quality guidelines, demonstrated the comparable pattern of high- and low-scoring AGREE-II domains and almost no quality improvement over time, with "rigor of development" and "applicability" as the two lowest scoring domains. The poor ratings in these countries occur regardless of who leads their national GD—medical or scientific societies, a GD agency, or a department under the Ministry of Health—indicating a more fundamental problem with GD quality than the framework itself.

The factor that was associated with the low quality of the Croatian guidelines was the development of the guidelines by medical societies, as opposed to the Ministry of Health or unofficial WGs. In the focus group discussions, additional factors behind the low scores were as follows: inadequate methodology; a lack of implementation systems in place; and a lack of awareness of the need for editorial independence, or border expertise and perspectives in working groups. We also identified a lack of methodological knowledge, misconceptions, and overconfidence among developers, which may interfere with the outcomes of their future training.

Other nations' experiences, such as Japan and Australia, suggest that efficient ways to improve the methodological quality of de novo, and, consequently, adapted and contextualized guidelines, may be achieved through the simultaneous collaborative actions of several organizations to build social infrastructure for guidelines, which would include the development of protocols for GD, methodological training, and an increase in guideline-related skills. Furthermore, the ongoing exposure of medical students to EBM and guideline topics during their studies may both raise the level of knowledge of the future users' and developers' and improve the evidence-based climate in the healthcare system.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijerph19159515/s1>, Figure S1: The increase in the number of published clinical practice guidelines in Croatia up to 2017; Figure S2: AGREE II domain scores as a mean percentage of the maximum score with 95% confidence interval (95% CI). Table S1: The distribution of overall assessment AGREE II scores (7-point scale) of Croatian CPGs; Table S2: Guideline frameworks in other countries and overall quality of the national guidelines. References [67–72] are cited in the supplementary materials.

**Author Contributions:** A.M., D.S., Ž.K. and A.J. designed the study. T.K., D.V., S.J.P. and A.J. rated the guidelines. D.S. and I.B. conducted the focus groups and transcribed the data. T.K. and A.J. performed the data analysis and drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by the Croatian Science Foundation, grant No. IP-2014-09-7672 to A.M. ("Professionalism in Health"). The founder had no role in the design or execution of the study or in the interpretation of the data.

**Institutional Review Board Statement:** The participants in the focus groups gave informed consent. Voice recordings from the focus group meetings were used only for obtaining the transcripts, and the transcripts were anonymized and used for further analysis. The ethics approval for conducting the interviews was obtained by the Ethics Committee at the University of Split, School of Medicine (Class: 003-08/11-0310005, Reg. No: 2181-198-03-04/10-11-0038). Personal data from the focus groups were fully protected according to the General Data Protection Regulation (GDPR).

**Informed Consent Statement:** All participants provided written informed consent before participating in the study, which included consent to publish anonymous quotes from individual participants.

**Data Availability Statement:** All relevant data are within the manuscript and its Supplementary Information Files.

**Acknowledgments:** We thank our colleagues who participated in focus groups. We would like to thank Mario Malički, METRICS, at Stanford University, for reviewing the manuscript and providing valuable comments and recommendations.

**Conflicts of Interest:** A.M. declares support from the Croatian Science Foundation, grant No. IP-2014-09-7672 (“Professionalism in Health”), and Z.K. declares he is the president of CMA (pro bono professional service) and provided the study materials free of charge. Other authors declare no competing interest.

## References

1. IoM (IOM). *Clinical Practice Guidelines We Can Trust*; Graham, R., Mancher, M., Wolman, D.M., Greenfield, S., Steinberg, E., Eds.; The National Academies Press (US): Washington, DC, USA, 2011.
2. Lewis, S.J.; Orland, B.I. The importance and impact of evidence-based medicine. *J. Manag. Care Pharm.* **2004**, *10* (Suppl. S5A), S3–S5. [[CrossRef](#)] [[PubMed](#)]
3. Grimshaw, J.M.; Russell, I.T. Effect of clinical guidelines on medical practice: A systematic review of rigorous evaluations. *Lancet* **1993**, *342*, 1317–1322. [[CrossRef](#)]
4. Lugtenberg, M.; Burgers, J.S.; Westert, G.P. Effects of evidence-based clinical practice guidelines on quality of care: A systematic review. *Qual. Saf. Health Care* **2009**, *18*, 385–392. [[CrossRef](#)] [[PubMed](#)]
5. Reames, B.N.; Shubeck, S.P.; Birkmeyer, J.D. Strategies for reducing regional variation in the use of surgery: A systematic review. *Ann. Surg.* **2014**, *259*, 616–627. [[CrossRef](#)] [[PubMed](#)]
6. Kamath, S.; Guyatt, G. Importance of evidence-based medicine on research and practice. *Indian J. Anaesth.* **2016**, *60*, 622–625. [[CrossRef](#)]
7. Shekelle, P.G.; Kravitz, R.L.; Beart, J.; Marger, M.; Wang, M.; Lee, M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. *Health Serv. Res.* **2000**, *34*, 1429–1448.
8. Woolf, S.H.; Grol, R.; Hutchinson, A.; Eccles, M.; Grimshaw, J. Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. *BMJ* **1999**, *318*, 527–530. [[CrossRef](#)]
9. Knai, C.; Brusamento, S.; Legido-Quigley, H.; Saliba, V.; Panteli, D.; Turk, E.; Car, J.; McKee, M.; Busse, R. Systematic review of the methodological quality of clinical guideline development for the management of chronic disease in Europe. *Health Policy* **2012**, *107*, 157–167. [[CrossRef](#)]
10. Muhlhauser, I.; Meyer, G. Evidence base in guideline generation in diabetes. *Diabetologia* **2013**, *56*, 1201–1209. [[CrossRef](#)]
11. Vanclooster, A.; Cassiman, D.; Van Steenberghe, W.; Swinkels, D.W.; Janssen, M.C.; Drenth, J.P.; Aertgeerts, B.; Wollersheim, H. The quality of hereditary haemochromatosis guidelines: A comparative analysis. *Clin. Res. Hepatol. Gastroenterol.* **2015**, *39*, 205–214. [[CrossRef](#)]
12. Alonso-Coello, P.; Irfan, A.; Solà, I.; Gich, I.; Delgado-Noguera, M.F.; Rigau, D.; Tort, S.; Bonfill, X.; Burgers, J.; Schunemann, H. The quality of clinical practice guidelines over the last two decades: A systematic review of guideline appraisal studies. *Qual. Saf. Health Care* **2010**, *19*, e58. [[CrossRef](#)] [[PubMed](#)]
13. Molino, C.G.; Romano-Lieber, N.S.; Ribeiro, E.; de Melo, D.O. Non-Communicable Disease Clinical Practice Guidelines in Brazil: A Systematic Assessment of Methodological Quality and Transparency. *PLoS ONE* **2016**, *11*, e0166367. [[CrossRef](#)]
14. Maaløe, N.; Ørtved, A.M.R.; Sørensen, J.B.; Dmello, B.S.; Akker, T.V.D.; Kujabi, M.L.; Kidanto, H.; Meguid, T.; Bygbjerg, I.C.; van Roosmalen, J.; et al. The injustice of unfit clinical practice guidelines in low-resource realities. *Lancet Glob. Health* **2021**, *9*, e875–e879. [[CrossRef](#)]
15. Feuerstein, J.D.; Castillo, N.E.; Akbari, M.; Belkin, E.; Lewandowski, J.J.; Hurley, C.M.; Lloyd, S.; Leffler, D.A.; Cheifetz, A.S. Systematic Analysis and Critical Appraisal of the Quality of the Scientific Evidence and Conflicts of Interest in Practice Guidelines (2005–2013) for Barrett’s Esophagus. *Dig. Dis. Sci.* **2016**, *61*, 2812–2822. [[CrossRef](#)] [[PubMed](#)]
16. Barriocanal, A.M.; López, A.; Monreal, M.; Montané, E. Quality assessment of peripheral artery disease clinical guidelines. *J. Vasc. Surg.* **2016**, *63*, 1091–1098. [[CrossRef](#)] [[PubMed](#)]
17. Dersch, R.; Toews, I.; Sommer, H.; Rauer, S.; Meerpohl, J.J. Methodological quality of guidelines for management of Lyme neuroborreliosis. *BMC Neurol.* **2015**, *15*, 242. [[CrossRef](#)] [[PubMed](#)]
18. Bayona, H.; Owolabi, M.; Feng, W.; Olowoyo, P.; Yaria, J.; Akinyemi, R.; Sawers, J.R.; Ovbiagele, B. A systematic comparison of key features of ischemic stroke prevention guidelines in low- and middle-income vs. high-income countries. *J. Neurol. Sci.* **2017**, *375*, 360–366. [[CrossRef](#)] [[PubMed](#)]
19. Becker, M.; Breuing, J.; Nothacker, M.; Deckert, S.; Brombach, M.; Schmitt, J.; Neugebauer, E.; Pieper, D. Guideline-based quality indicators—a systematic comparison of German and international clinical practice guidelines. *Implement. Sci.* **2019**, *14*, 71. [[CrossRef](#)] [[PubMed](#)]

20. Colunga-Lozano, L.E.; Gerardo-Morales, V.; Pérez-Gaxiola, G.; Vázquez-Álvarez, A.O.; Torres, F.J.G.; Perales-Guerrero, L.; Yanowsky-Ortega, E.; Martínez-Tolentino, U.I.; Sánchez-Villasaca, S.J.; García-Macías, L.A.; et al. Methodological assessment of Mexican Clinical Practice Guidelines: GRADE framework adherence and critical appraisal. *J. Eval. Clin. Pract.* **2021**, *27*, 385–390. [[CrossRef](#)]
21. Dans, L.F.; Salaveria-Imperial, M.L.A.; Miguel, R.T.D.; Tan-Lim, C.S.C.; Eubanas, G.A.S.; Tolosa, M.T.S.; Ho, B.L.C.; Silvestre, M.A.A. Guidelines in Low and Middle Income Countries Paper 3: Appraisal of Philippine Clinical Practice Guidelines using Appraisal of Guidelines for Research and Evaluation II: Improvement needed for rigor, applicability, and editorial independence. *J. Clin. Epidemiol.* **2020**, *127*, 184–190. [[CrossRef](#)]
22. Loezar, C.; Pérez-Bracchiglione, J.; Arancibia, M.; Meza, N.; Vargas, M.; Papuzinski, C.; Rada, G.; Ondarza, C.; Jahr, C.; Cadena, C.; et al. Guidelines in Low and Middle Income Countries Paper 2: Quality assessment of Chilean guidelines: Need for improvement in rigor, applicability, updating, and patients' inclusion. *J. Clin. Epidemiol.* **2020**, *127*, 177–183. [[CrossRef](#)] [[PubMed](#)]
23. Esandi, M.E.; Ortiz, Z.; Chapman, E.; Dieguez, M.G.; Mejia, R.; Bernztein, R. Production and quality of clinical practice guidelines in Argentina (1994–2004): A cross-sectional study. *Implement. Sci.* **2008**, *3*, 43. [[CrossRef](#)] [[PubMed](#)]
24. Seto, K.; Matsumoto, K.; Fujita, S.; Kitazawa, T.; Amin, R.; Hatakeyama, Y.; Hasegawa, T. Quality assessment of clinical practice guidelines using the AGREE instrument in Japan: A time trend analysis. *PLoS ONE* **2019**, *14*, e0216346. [[CrossRef](#)]
25. Barker, T.H.; Dias, M.; Stern, C.; Porritt, K.; Wiechula, R.; Aromataris, E.; Brennan, S.; Schünemann, H.J.; Munn, Z. Guidelines rarely used GRADE and applied methods inconsistently: A methodological study of Australian guidelines. *J. Clin. Epidemiol.* **2021**, *130*, 125–134. [[CrossRef](#)]
26. Sundberg, L.R.; Garvare, R.; Nystrom, M.E. Reaching beyond the review of research evidence: A qualitative study of decision making during the development of clinical practice guidelines for disease prevention in healthcare. *BMC Health Serv. Res.* **2017**, *17*, 344.
27. Wiercioch, W.; Akl, E.A.; Santesso, N.; Zhang, Y.; Morgan, R.L.; Yepes-Nuñez, J.J.; Kowalski, S.; Baldeh, T.; Mustafa, R.A.; Laisaar, K.-T.; et al. Assessing the process and outcome of the development of practice guidelines and recommendations: PANELVIEW instrument development. *Can. Med. Assoc. J.* **2020**, *192*, E1138–E1145. [[CrossRef](#)]
28. Bhaumik, S.; Jagadesh, S.; Ellatar, M.; Kohli, N.; Riedha, M.; Moi, M. Clinical practice guidelines in India: Quality appraisal and the use of evidence in their development. *J. Evid.-Based Med.* **2018**, *11*, 26–39. [[CrossRef](#)]
29. Burgers, J.S.; Cluzeau, F.A.; Hanna, S.E.; Hunt, C.; Grol, R. Characteristics of high-quality guidelines: Evaluation of 86 clinical guidelines developed in ten European countries and Canada. *Int. J. Technol. Assess. Health Care* **2003**, *19*, 148–157. [[CrossRef](#)] [[PubMed](#)]
30. Armstrong, J.J.; Goldfarb, A.M.; Instrum, R.S.; MacDermid, J.C. Improvement evident but still necessary in clinical practice guideline quality: A systematic review. *J. Clin. Epidemiol.* **2017**, *81*, 13–21. [[CrossRef](#)] [[PubMed](#)]
31. Talagala, I.A.; Samarakoon, Y.; Senanayake, S.; Abeysena, C. Sri Lankan clinical practice guidelines: A methodological quality assessment utilizing the AGREE II instrument. *J. Eval. Clin. Pract.* **2019**, *25*, 630–636. [[CrossRef](#)]
32. Choi, T.Y.; Choi, J.; Lee, J.A.; Jun, J.H.; Park, B.; Lee, M.S. The quality of clinical practice guidelines in traditional medicine in Korea: Appraisal using the AGREE II instrument. *Implement. Sci.* **2015**, *10*, 104. [[CrossRef](#)] [[PubMed](#)]
33. Timana, R.; Alva-Diaz, C.; Suarez, V.; Pimentel, P.; Dongo, V. Characteristics and quality of the of clinical practice guidelines in the Social Security of Health of Peru. *Semergen* **2018**, *44*, 549–556. [[PubMed](#)]
34. Canelo-Aybar, C.; Balbin, G.; Perez-Gomez, A.; Florez, I.D. Clinical practice guidelines in Peru: Evaluation of its quality using the AGREE II instrument. *Rev. Peru. Med. Exp. Salud Publica* **2016**, *33*, 732–738. [[CrossRef](#)]
35. Zhou, Q.; Wang, Z.; Shi, Q.; Zhao, S.; Xun, Y.; Liu, H.; Zhang, H.; Liu, X.; Wang, X.; Yao, L.; et al. Clinical Epidemiology in China series. Paper 4: The reporting and methodological quality of Chinese clinical practice guidelines published between 2014 and 2018: A systematic review. *J. Clin. Epidemiol.* **2021**, *140*, 189–199. [[CrossRef](#)]
36. Kataoka, Y.; Anan, K.; Taito, S.; Tsujimoto, Y.; Kurata, Y.; Wada, Y.; Maruta, M.; Kanaoka, K.; Oide, S.; Takahashi, S.; et al. Quality of clinical practice guidelines in Japan remains low: A cross-sectional meta-epidemiological study. *J. Clin. Epidemiol.* **2021**, *138*, 22–31. [[CrossRef](#)]
37. Morciano, C.; Laricchiuta, P.; Taruscio, D.; Schunemann, H. European Reference Networks and Guideline Development and Use: Challenges and Opportunities. *Public Health Genom.* **2015**, *18*, 318–320. [[CrossRef](#)] [[PubMed](#)]
38. Liječnički Vjesnik—Glasilo Hrvatskoga Liječničkog Zbora. Available online: <https://lijecnicki-vjesnik.hlz.hr/> (accessed on 3 March 2022).
39. AGREE Next Steps Consortium. *Appraisal of Guidelines for Research & Evaluation II (AGREE II) Instrument*; Canadian Institutes of Health Research: Ottawa, ON, Canada, 2013.
40. Grimmer, K.; Dizon, J.M.; Milanese, S.; King, E.; Beaton, K.; Thorpe, O.; Lizarondo, L.; Luker, J.; Machotka, Z.; Kumar, S. Efficient clinical evaluation of guideline quality: Development and testing of a new tool. *BMC Med. Res. Methodol.* **2014**, *14*, 63. [[CrossRef](#)]
41. Van Eck, N.J.; Waltman, L. VOSviewer Manual 1.6.11. Manual. 2016, pp. 1–28. Available online: [http://www.vosviewer.com/documentation/Manual\\_VOSviewer\\_1.5.4.pdf](http://www.vosviewer.com/documentation/Manual_VOSviewer_1.5.4.pdf) (accessed on 24 February 2020).
42. Strauss, A.; Corbin, J. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*; Sage: Los Angeles, CA, USA, 1990.
43. HLZ. Hrvatski Liječnički Zbor. Available online: [https://www.hlz.hr/strucna-drustva-hlz/?wpv\\_view\\_count=3943&wpv\\_post\\_search=&wpv\\_paged=1](https://www.hlz.hr/strucna-drustva-hlz/?wpv_view_count=3943&wpv_post_search=&wpv_paged=1) (accessed on 10 March 2022).

44. GRADE Handbook. Available online: <https://gdt.gradepro.org/app/handbook/handbook.html#h.33qgws879zw> (accessed on 9 March 2022).
45. Tudor, K.I.; Kozina, P.N.; Marusic, A. Methodological rigour and transparency of clinical practice guidelines developed by neurology professional societies in Croatia. *PLoS ONE* **2013**, *8*, e69877. [[CrossRef](#)] [[PubMed](#)]
46. Tokalic, R.; Vidak, M.; Buljan, I.; Marusic, A. Reporting quality of European and Croatian health practice guidelines according to the RIGHT reporting checklist. *Implement. Sci.* **2018**, *13*, 135. [[CrossRef](#)]
47. Song, F.; Parekh, S.; Hooper, L.; Loke, Y.K.; Ryder, J.; Sutton, A.J.; Hing, C.; Kwok, C.S.; Pang, C.L.; Harvey, I. Dissemination and publication of research findings: An updated review of related biases. *Health Technol. Assess.* **2010**, *14*, 1–220. [[CrossRef](#)]
48. Kane, R.L. Creating practice guidelines: The dangers of over-reliance on expert judgment. *J. Law Med. Ethics A J. Am. Soc. Law Med. Ethics* **1995**, *23*, 62–64. [[CrossRef](#)] [[PubMed](#)]
49. Scottish Intercollegiate Guidelines Network. *SIGN 50: A Guideline Developer's Handbook*; SIGN: Edinburgh, Scotland, 2014; Available online: <http://www.sign.ac.uk> (accessed on 15 March 2022).
50. Ehrlinger, J.; Johnson, K.; Banner, M.; Dunning, D.; Kruger, J. Why the Unskilled Are Unaware: Further Explorations of (Absent) Self-Insight Among the Incompetent. *Organ. Behav. Hum. Decis. Processes* **2008**, *105*, 98–121. [[CrossRef](#)] [[PubMed](#)]
51. Motta, M.; Callaghan, T.; Sylvester, S. Knowing less but presuming more: Dunning-Kruger effects and the endorsement of anti-vaccine policy attitudes. *Soc. Sci. Med.* **2018**, *211*, 274–281. [[CrossRef](#)] [[PubMed](#)]
52. Rahmani, M. Medical Trainees and the Dunning-Kruger Effect: When They Don't Know What They Don't Know. *J. Grad. Med. Educ.* **2020**, *12*, 532–534. [[CrossRef](#)] [[PubMed](#)]
53. NICE. The Guideline Development Group. In *The Guidelines Manual: Process and Methods*; National Institute for Health and Clinical Excellence: London, UK, 2012; Available online: <https://www.nice.org.uk/process/pmg6/chapter/the-guideline-development-group> (accessed on 13 October 2021).
54. ESHRE. *Manual for ESHRE Guideline Development*; European Society of Human Reproduction and Embryology: Belgium, Germany, 2019; Available online: <https://www.eshre.eu/Guidelines-and-Legal/Guidelines/Guideline-development-process> (accessed on 10 October 2021).
55. Yaşar, I.; Kahveci, R.; Artantaş, A.B.; Başer, D.A.; Cihan, F.G.; Sencan, I.; Koc, E.M.; Ozkara, A. Quality Assessment of Clinical Practice Guidelines Developed by Professional Societies in Turkey. *PLoS ONE* **2016**, *11*, e0156483. [[CrossRef](#)] [[PubMed](#)]
56. Buchan, H.A.; Currie, K.C.; Lourey, E.J.; Duggan, G.R. Australian clinical practice guidelines—A national study. *Med. J. Aust.* **2010**, *192*, 490–494. [[CrossRef](#)]
57. Bachmann, L.; Ulyte, A.; Dressel, H. Clinical practice guidelines of medical societies in Switzerland: Analysis of the current state. *Swiss Med. Wkly.* **2019**, *149*, w20134. [[CrossRef](#)]
58. Malicki, M.; Jeroncic, A.; Aalbersberg, I.J.; Bouter, L.; Ter Riet, G. Systematic review and meta-analyses of studies analysing instructions to authors from 1987 to 2017. *Nat. Commun.* **2021**, *12*, 5840. [[CrossRef](#)]
59. Lien, L.; Pedersen, M.S.; Landheim, A. Norwegian guidelines for persons with concurrent mental disorders and substance use disorders: Assessment, treatment and rehabilitation—How to bridge gaps between current practice and clinical guidelines? *BMC Health Serv Res.* **2014**, *14* (Suppl. S2), P67. [[CrossRef](#)]
60. Grol, R.; Dalhuijsen, J.; Thomas, S.; Veld, C.; Rutten, G.; Mokkink, H. Attributes of clinical guidelines that influence use of guidelines in general practice: Observational study. *BMJ* **1998**, *317*, 858–861. [[CrossRef](#)]
61. Slade, S.C.; Kent, P.; Patel, S.; Bucknall, T.; Buchbinder, R. Barriers to Primary Care Clinician Adherence to Clinical Guidelines for the Management of Low Back Pain: A Systematic Review and Metasynthesis of Qualitative Studies. *Clin. J. Pain* **2016**, *32*, 800–816. [[CrossRef](#)] [[PubMed](#)]
62. Kingdon, C.; Downe, S.; Betran, A.P. Non-clinical interventions to reduce unnecessary caesarean section targeted at organisations, facilities and systems: Systematic review of qualitative studies. *PLoS ONE* **2018**, *13*, e0203274. [[CrossRef](#)]
63. Smith, C.A.; Toupin-April, K.; Jutai, J.W.; Duffy, C.M.; Rahman, P.; Cavallo, S.; Brosseau, L. A Systematic Critical Appraisal of Clinical Practice Guidelines in Juvenile Idiopathic Arthritis Using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) Instrument. *PLoS ONE* **2015**, *10*, e0137180. [[CrossRef](#)] [[PubMed](#)]
64. Brosseau, L.; Rahman, P.; Poitras, S.; Toupin-April, K.; Paterson, G.; Smith, C.; King, J.; Casimiro, L.; De Angelis, G.; Loew, L.; et al. A systematic critical appraisal of non-pharmacological management of rheumatoid arthritis with Appraisal of Guidelines for Research and Evaluation II. *PLoS ONE* **2014**, *9*, e95369.
65. Brosseau, L.; Rahman, P.; Toupin-April, K.; Poitras, S.; King, J.; De Angelis, G.; Loew, L.; Casimiro, L.; Paterson, G.; McEwan, J. A systematic critical appraisal for non-pharmacological management of osteoarthritis using the appraisal of guidelines research and evaluation II instrument. *PLoS ONE* **2014**, *9*, e82986. [[CrossRef](#)] [[PubMed](#)]
66. Brouwers, M.C.; Kho, M.E.; Browman, G.P.; Burgers, J.S.; Cluzeau, F.; Feder, G.; Fervers, B.; Graham, I.D.; Hanna, S.E.; Makarski, J. Development of the AGREE II, part 2: Assessment of validity of items and tools to support application. *Can. Med. Assoc. J.* **2010**, *182*, E472–E478. [[CrossRef](#)]
67. Brouwers, M.C.; Makarski, J.; Durocher, L.D.; Levinson, A. E-learning interventions are comparable to user's manual in a randomized trial of training strategies for the AGREE II. *Implement. Sci.* **2011**, *6*, 81. [[CrossRef](#)]
68. Attal, N.; Cruccu, G.; Haanpää, M.; Hansson, P.; Jensen, T.S.; Nurmikko, T.; Sampaio, C.; Sindrup, S.; Wiffen, P. EFNS guidelines on pharmacological treatment of neuropathic pain. *Eur. J. Neurol.* **2006**, *13*, 1153–1169. [[CrossRef](#)] [[PubMed](#)]

69. Marinovic, D.; Hren, D.; Sambunjak, D.; Rašić, I.; Škegro, I.; Marušić, A.; Marušić, M. Transition from longitudinal to block structure of preclinical courses: Outcomes and experiences. *Croat. Med. J.* **2009**, *50*, 492–506. [[CrossRef](#)]
70. Bloom, J.D.; Hoxha, I.; Sambunjak, D.; Sondorp, E. Ethnic segregation in Kosovo's post-war health care system. *Eur. J. Public Health* **2007**, *17*, 430–436. [[CrossRef](#)]
71. Then, K.L.; Rankin, J.A.; Ali, E. Focus group research: What is it and how can it be used? *Can. J. Cardiovasc. Nurs.* **2014**, *24*, 16–22.
72. Guest, G.; Namey, E.; McKenna, K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods* **2016**, *29*, 3–22. [[CrossRef](#)]