Guidelines in Gastroenterology: Careful Interpretation Is Essential

Catarina Correia\textsuperscript{a} Nuno Almeida\textsuperscript{a, b} Pedro Narra Figueiredo\textsuperscript{a, b}

\textsuperscript{a}Gastroenterology Department, Coimbra University Hospital Centre, Coimbra, Portugal; \textsuperscript{b}Faculty of Medicine, University of Coimbra, Coimbra, Portugal

Keywords
Guidelines \cdot Recommendations \cdot Gastroenterology \cdot Evidence level

Abstract

Introduction: Clinical practice guidelines (CPG) contain recommendations that aim to guide physicians in the diagnosis of and therapeutic approach toward patients affected by gastrointestinal (GI) pathologies. These CPG systematically combine scientific evidence and clinical judgment, culminating in recommendations that have been shown to improve patient care. Material and Methods: European and North American guidelines published in the area of gastroenterology in 2018 and 2019 were considered for inclusion. To standardize the results, only guidelines that used GRADE as an evidence system were included. Thus, in the end, 1,233 recommendations from 29 guidelines published between 2018 and 2019 were analyzed. Results: Of the 1,233 recommendations collected, 324 (26.3\%) had a low level of evidence and 127 (10.3\%) had a very low level of evidence, indicating little evidence or expert opinion. Of the 29 publications analyzed, 14 (48.3\%) did not present any recommendation with a high level of evidence. Regarding the 1,233 individual recommendations expressed in these 29 publications, only 336 (27.25\%) assumed a high level of evidence, with 277 (82.44\%) referring to liver pathology. Of the recommendations evaluated, 77 were from North American societies and the remaining 1,156 were European recommendations. In relation to the first group, only 3 (3.9\%) had a high level of evidence belonging to the Guidelines for Sedation and Anesthesia in GI Endoscopy. Conclusions: More than 25\% of all recommendations currently accepted to guide patients with gastroenterological disorders are based on low-quality evidence or expert opinion. Thus, these documents should guide our performance, but clinical sense and multidisciplinarity must not be overlooked in dubious cases and with weak scientific evidence. Research should focus on the development of randomized controlled trials and systematic reviews to improve the evidence supporting the guidelines that guide clinical practice.
Resumo

**Introdução:** As diretrizes que orientam a prática clínica contêm recomendações de forma a que os médicos possam determinar os cuidados mais adequados para cada paciente. Estas diretrizes combinam evidências científicas com o julgamento clínico, culminando em recomendações destinadas a otimizar a prestação de cuidados. **Material e Métodos:** Diretrizes publicadas na área de Gastroenterologia entre 2018 e 2019 foram consideradas para inclusão. Para padronizar os resultados, apenas diretrizes usando o GRADE como um sistema de evidência foram incluídas. Assim, foram analisadas 1,233 recomendações de 29 diretrizes publicadas entre 2018–2019. **Resultados:** Das 1,233 recomendações incluídas, 334 (26.3%) apresentavam um nível de evidência baixo e 127 (10.3%) um nível de evidência muito baixo, indicando pouca evidência ou mesmo opinião de especialistas. Das 29 publicações analisadas, 14 (48,3%) não apresentavam nenhuma recomendação com alto nível de evidência. Em relação às 1,233 recomendações incluídas nas 29 publicações, apenas 336 (27.25%) assumiam alto nível de evidência, sendo 277 (82.44%) referentes à patologia hepática. Das recomendações avaliadas, 77 eram de sociedades norte-americanas e as restantes 1,156 recomendações europeias. Em relação ao primeiro grupo, apenas 3 (3,9%) possuíam alto nível de evidência e pertenciam às “Diretrizes para sedação e anestesia em endoscopia gastrointestinal.” **Conclusões:** Mais de 25% de todas as recomendações atualmente aceitas para orientar pacientes com patologias gastrointestinal são baseadas em evidências de baixa qualidade ou opinião de especialistas. Estes documentos devem orientar a nossa forma de atuar, mas o senso clínico e a abordagem multidisciplinar não devem ser esquecidos em casos duvidosos e com evidência científica fraca. A investigação deve concentrar-se no desenvolvimento de ensaios clínicos randomizados e revisões sistemáticas para melhorar as evidências que apoiam as diretrizes que orientam a prática clínica.

---

Guidelines in Gastroenterology

Introduction

It is estimated that 30% of the European population is affected, at least once in their lives, by gastrointestinal (GI) diseases [1]. Among the most prevalent pathologies in the Portuguese population, we highlight *Helicobacter pylori* infection (60–84%), dyspepsia (20–40%), gastroesophageal reflux disease (35%), and irritable bowel syn-...
ologists. Therefore, the aim of this research was to review the guidelines published by some of the most relevant GI societies between 2018 and 2019. Our goal was to report the level of evidence supporting their recommendations and to identify areas where evidence can be improved with additional research.

### Materials and Methods

A list of CPG published between 2018 and 2019, in the gastroenterology area, were obtained by searching the websites from some of the international societies that follow our daily clinical practice. European CPG were obtained from the United European Gastroenterology (UEG) database [12]. American CPG data were collected from the American Gastroenterological Association (AGA), the American College of Gastroenterology (ACG), the American Association for the Study of Liver Diseases (ASSLD), and the American Society for Gastrointestinal Endoscopy (ASGE) [13–16]. A total of 89 European and North American CPG published in the gastroenterology field between 2018 and 2019 were analyzed. In order to standardize the collected sample, CPG that did not use the GRADE system to rate the quality of evidence and CPG that were still in execution or ones without levels of evidence for their recommendations were excluded. Considering this, only 29 guidelines were considered for inclusion. This proportion of CPG analyzed and included in the study is shown in Figure 1. All of the guidelines included in the research are presented in online supplementary material (for all online suppl. material, see www.karger.com/doi/10.1159/000518322).

The quality of the evidence is classified into the following 4 levels: high, moderate, low, and very low (Table 1) [7]. A total of 1,233 recommendations from the 29 guidelines were analyzed. Afterwards, all 1,233 recommendations were extracted into an SPSS spreadsheet with its associated level of evidence. We also stratified the list by the guideline in which they were published along with the year of publication.

The statistical analysis was performed using statistical software (SPSS version 23). Frequencies and proportions were calculated to describe the collected data.

### Results

A total of 29 guidelines (7 from American societies and 22 from European societies) were included in this study, representing 1,233 recommendations. Of the 1,233 recommendations collected, 324 (26.3%) were based on a low level of evidence and 127 (10.3%) were based on a very low level of evidence, indicating poor evidence or expert opinion. Four hundred forty-six (36.2%) were based on a moderate level of evidence and 336 (27.3%) were based on high levels of evidence, with 277 (82.44%) of these being related to liver disease. These results are detailed in Table 2.

Only 2 guidelines – the European Association for the Study of the Liver (EASL) Recommendations on Treatment of Hepatitis C (2018) [17] and the EASL Clinical Practice Guidelines: Management of Alcohol-Related Liver Disease [18] – had over 50% of recommendations supported by high-level evidence.

Of the 29 publications analyzed, 14 (48.3%) did not present any recommendation with a high level of evidence, 15 contained high-level recommendations, and only 2 (6.7%) did not present recommendations with a low or very low level of evidence (Fig. 2).
Of the recommendations evaluated, 77 were from North American societies and the remaining 1,156 were European recommendations. In relation to the first group, only 3 (3.9%) had a high level of evidence, belonging to the same guideline – Guidelines for Sedation and Anesthesia in GI Endoscopy. The analysis of recommendations per field showed that only the ones related to liver and bowel diseases (not IBD) had a proportion with a high level of evidence higher than 30% (34.37 and 33.3%, respectively). Table 3 reports the proportion of recommendations with high-level evidence per “field” of gastroenterology.

The year 2018 had the highest percentage of recommendations supported by a high level of evidence (88.7%, 298 out of 336). Between 2018 and 2019, the percentage of recommendations supported by a low level of evidence increased from 20.8 to 37.3%.

### Table 2. Analysis of recommendations included in this study by society, field of gastroenterology, and proportion of evidence levels for each society

| Society | Recommendations, n | Recommendations per field, n | Proportion of evidence levels for each society, n (%) |
|---------|-------------------|-------------------------------|------------------------------------------------------|
|         |                   | BD | IBD | LD | ED | GD | PD | Tc | very low | low | moderate | high |
| AGA     | 22                | 6  | 16  | –  | –  | –  | –  | –  | 1 (4.5) | 11 (50) | 8 (36.4) | 2 (9.1) |
| ACG     | 30                | –  | –   | 10 | –  | 20 | –  | –  | 19 (63.3) | 9 (30) | 2 (6.7)  | –     |
| ASGE    | 16                | –  | –   | 8  | –  | –  | 8  | –  | 3 (18.8) | 5 (31.3) | 5 (31.3) | 3 (18.8) |
| AASLD   | 15                | –  | –   | 15 | –  | –  | –  | –  | 8 (53.3) | 4 (26.7) | 3 (20)   | –     |
| EASL    | 637               | –  | –   | 637| –  | –  | –  | –  | 62 (9.7) | 65 (10.2)| 242 (38) | 268 (42.1) |
| ECCO    | 29                | –  | –   | 29 | –  | –  | –  | –  | 6 (20.7) | 8 (27.6) | 14 (48.3)| 1 (3.4) |
| ESGCTP  | 82                | –  | –   | –  | –  | 82 | –  | –  | –         | 67 (81.7)| 15 (18.3) | –     |
| WGPPC   | 25                | –  | –   | –  | –  | 25 | –  | –  | –         | 11 (44) | 11 (44)  | 3 (12) |
| ESGE    | 110               | –  | –   | –  | –  | –  | –  | 110 | 15 (13.6)| 63 (57.3)| 28 (25.5)| 4 (3.6) |
| BSG     | 195               | –  | –   | 144| 31 | –  | 20 | –  | 13 (6.7) | 72 (36.9)| 80 (41) | 30 (15.4) |
| ESsCD   | 72                | 72 | –   | –  | –  | –  | –  | –  | –         | 9 (12.5) | 38 (52.8)| 25 (34.7) |
| TOTAL   | 1,233             | 78 | 45  | 806| 8  | 31 | 127| 138| 127 (10.3)| 324 (26.3)| 446 (36.2)| 336 (27.3) |

ECCO, European Crohn’s and Colitis Organization; ESGCTP, European Study Group on Cystic Tumors of the Pancreas; WGPPC, Working Group of the Polish Pancreatic Club; BSG, British Society of Gastroenterology; ESsCD, European Society for the Study of Coeliac Disease; BD, bowel disease (not IBD); LD, liver disease; ED, esophageal disease; GD, gastric disease; PD, pancreatic disease; Tc, techniques in gastrointestinal endoscopy.

### Table 3. Proportion of high-level evidence per field of gastroenterology

| Field of gastroenterology | Proportion of recommendations of high-level evidence |
|---------------------------|-----------------------------------------------------|
| Liver disease             | 277/806 (34.37)                                     |
| Bowel disease (not IBD)   | 26/78 (33.3)                                        |
| Esophageal disease        | 18/100 (18)                                         |
| Gastric disease           | 3/31 (9.68)                                         |
| Techniques in gastrointestinal endoscopy | 7/138 (5.07)                                      |
| IBD                       | 2/45 (4.44)                                         |
| Pancreatic disease        | 3/127 (2.36)                                        |
| Total                     | 336/1,233 (27.25)                                   |

Values in parentheses are percents.
Discussion

The 29 guidelines analyzed contained 1,233 recommendations. More than a quarter (26.3%) were based on a low evidence level and only 27.3% were supported by high evidence levels.

A similar study performed by the American College of Emergency Physicians found that less than 10% of their recommendations were based on high-quality evidence, and the majority were based on expert opinion [19]. Similarly, the American College of Chest Physicians found that only 0.4% of recommendations for the treatment of thromboembolism were based on high-level evidence [20]. In the field of gastroenterology, a similar study done by Meyer et al. [21], analyzing the scientific evidence underlying the American College of Gastroenterology (ACG) CPG also concluded that very few recommendations made by the AGA are supported by high levels of evidence. More than half of all recommendations made by the AGA are based on low-quality evidence or expert opinion [21]. Feuerstein et al. [22] also referred that when the gastroenterology guidelines rate the quality of evidence for their recommendations most recommendations are based on lower-quality evidence. A systematic analysis and critical appraisal of the quality of the scientific evidence in Practice Guidelines for Barrett’s Esophagus revealed that nearly 50% of the recommendations are based on expert opinion or poor-quality evidence [23]. A critical review of scientific evidence and evolving recommendations of AASLD CPG also concluded that, despite significant increases in the numbers of recommendations within AASLD practice guidelines over time, only a minority are supported by grade I evidence, highlighting the need for developing well-designed investigations to provide evidence for areas of uncertainty and improving the quality of future guidelines in hepatobiliary diseases [24].

Although the amount of high-level evidence supporting the CPG included in our study (26.3%) compares favorably to the previously mentioned papers, this is far from what would be expected. When considering the proportion of recommendations with a high level of evidence demonstrated in Tables 2 and 3, it is notable that these low percentages of recommendations with a high level of evidence will affect the overall quality of the guidelines reported. These results confirm that, although guidelines should be followed, interpretation with caution and careful clinical judgment is mandatory.

A limitation of our study was the impossibility to analyze how evidence levels can differ from recommendations related with the same topic in different guidelines. In this research we only included guidelines published in 2018 and 2019, and consequently the topics covered are pretty much all different, even within the same area (online suppl. material). Considering this limitation of our research, the only 2 guidelines which address the same topic are the ACG Clinical Guideline: Diagnosis and Management of Pancreatic Cysts and European evidence-based guidelines on pancreatic cystic neoplasms [25, 26]. In this case, there were no big discrepancies between them – they were published in the same year (2018) and addressed the same topic. In both, there were no recommendations with a high level of evidence. Regarding recommendations with a low and very low level of evidence, the American guideline and the European guideline presented a proportion of 100 and 81.7%, respectively.

RCT are the cornerstone of clinical decision making, and the field of gastroenterology has a poor history of producing influential RCT [27]. Nearly 25,000 RCT are published each year but, given that 14 of the analyzed guidelines do not contain recommendations with a high level of evidence, it appears that few RCT find their way into CPG in the field of gastroenterology [28].

The discrepancy between the total number of gastroenterology RCT and those supporting guideline recommendations may be due to 2 factors: overlap between RCT and practical barriers to conducting RCT. The potential overlap between RCT, also known as research waste, may delay the advent of treatments for patients with preventable diseases [29, 30]. For example, in our study we found that, of the 14 guidelines with no high-quality evidence, almost all were focused on either prevention, treatment, or management. While some of these individual recommendations may not be subjectable to an RCT due to ethical or practical concerns, some of these may be tested in a randomized fashion.

Koh et al. [24] reported a 36% increase in recommendation number for the AASLD since their development in 1998. Nonetheless, despite this substantial increase, less than 15% were based on high-grade evidence. Since 2003, the National Institutes of Health budget for digestive disease research reached a plateau, while corporate funding for gastroenterology research dropped by more than 60% since 2008 [31, 32]. This translates into increased competition for grant applications, which are being awarded at the lowest rate in decades [33]. RCT are among the most time-consuming and expensive research studies, despite producing the highest level of evidence. We suggest that the multiple gastroenterology organizations should encourage future research to strengthen rec-
ommendations that are currently supported by expert opinion or a low evidence level.

Some studies have estimated that 54–70% of physicians consistently use CPG in practice, so their quality is of the greatest importance [5, 34]. Therefore, the scarcity of high-quality evidence affects physicians seeking evidence-based treatment options and patients seeking evidence-based care. Recommendations that are based on low levels of evidence are important areas for research as they may expose patients to unnecessary risks and inflate health care costs [35]. CPG can give physicians a false sense of security, causing them to rely more on the guideline than on critical-thinking and updated research [28]. This shows the importance of basing guideline recommendations on high-level evidence. When creating guidelines from expert consensus they are subject to bias. Conflicts of interest are potential sources of bias in the development of CPG [36]. Another major problem in basing recommendations on expert consensus is the fact that opinions vary between experts. This is illustrated in a study where Marras et al. [37] found the highest percent of expert agreement on any recommendation was 81%. Without further evidence validating one opinion over another, physicians will use their judgment to treat patients, leading to variability in care. The number of gastroenterology recommendations supported by low-level evidence and expert opinion highlights the need for further research leading to better evidence and improved patient outcomes.

**Limitations**

A limitation of our study is the use of guideline repositories, which are disposed to more variations than public databases. The reduce number of guidelines and recommendations included, in some fields, should lead to careful interpretation of results.

Because the guidelines were published before the current year, they may not be an accurate reflection of the current levels of evidence in gastroenterology literature and therefore our study may underestimate the current research quality in the field.

**Conclusion**

More than a quarter of all recommendations are based on low-quality evidence or expert opinion. Considering the CPG included, 14 of the 29 CPG that guide our practice do not contain recommendations supported by high-level evidence. All CPG recommendations should be considered equally relevant. The recommendations contain supporting evidence ranging from high quality (RCT) to low quality (expert opinion). Research should focus on the development of RCT and systematic reviews to improve the evidence supporting the CPG that guide our daily practice.

**Statement of Ethics**

All of the procedures performed were in accordance with the ethical standards of the institutional and/or national research committee.

**Conflict of Interest Statement**

The authors have no conflict of interests to declare.

**Funding Sources**

The authors received no financial support.

**Author Contributions**

The authors contributed equally to the writing of this paper.

**References**

1 de Carvalho IM, Nunes S, de Matos L, et al. *Gastroenterologia e Hepatologia*. Rede Nacional de Especialidade Hospitalar e de Referência. 2017.
2 Bastos J, Peleteiro B, Barros R, Alves L, Severo M, de Fátima Pina M, et al. Sociodemographic determinants of prevalence and incidence of *Helicobacter pylori* infection in Portuguese adults. *Helicobacter*. 2013 Dec;18(6):413–22.
3 World Health Organization. *The Global Cancer Observatory – Portugal*. 2020 [cited 2021 Apr 4]. Available from: https://gco.iarc.fr/today/data/factsheets/populations/620-portugal-fact-sheets.pdf.
4 Graham R, Mancher M, Miller Wolman D, et al. *Clinical practice guidelines we can trust*. Washington: National Academies Press; 2011.
5 Falareti G, Legnani C, Cosmi B, Antonucci E, Erba N, Pelli D, et al.; DULCIS (D-dimer-ULtrasonography in Combination Italian Study) Investigators. Comparison between different D-Dimer cutoff values to assess the individual risk of recurrent venous thromboembolism: analysis of results obtained in the DULCIS study. *Int J Lab Hematol*. 2016 Feb;38(1):42–9.
6 Clinical Practice Guidelines. 2005. Available from: http://www.library.tmc.edu/pracguid.html.
