HHS Public Access

Author manuscript

Hepatology. Author manuscript; available in PMC 2015 October 21.

Published in final edited form as:

Hepatology. 2013 December ; 58(6): 2142–2152. doi:10.1002/hep.26578.

The AASLD Clinical Practice Guidelines: A Critical Review of Scientific Evidence and Evolving Recommendations

Christopher Koh[1], **Xiongce Zhao**[2], **Niharika Samala**[1], **Sasan Sakiani**[1], **T. Jake Liang**[1], and **Jayant A. Talwalkar**[3]

[1]Liver Diseases Branch, National Institute of Diabetes & Digestive & Kidney Diseases, National Institutes of Health

[2]Office of the Director, National Institute of Diabetes & Digestive & Kidney Diseases, National Institutes of Health

[3]Division of Gastroenterology & Hepatology, Mayo Clinic

Abstract

Background—The American Association for the Study of Liver Diseases (AASLD) practice guidelines provide recommendations in diagnosing and managing patients with liver disease from available scientific evidence in combination with expert consensus opinions.

Aim—To systematically review the evolution of recommendations from AASLD guidelines and identify gaps limiting the evidence-based foundations of these guidelines.

Methods—Initial and current AASLD guidelines published from January 1998 to August 2012 were reviewed. The AGREE II instrument was used to evaluate rigour and transparency of guideline development. The number of recommendations, distribution of grades (strength or certainty), classes (benefit versus risk) and types of recommendations were evaluated. Whenever possible, multiple versions were evaluated for evolving scientific evidence.

Results—A total of 991 recommendations from 28 guidelines on 17 topics were evaluated. From initial to current guidelines, the total number of recommendations increased by 36%(512 to 699). The largest increases were from chronic hepatitis B (HBV) (+71), liver transplantation (+53) and autoimmune hepatitis (AIH) (+27). Most current recommendations are grade II (44%) and less than 20% are grade I. The AGREE II evaluation showed global improvement in guideline quality. Both HBV and chronic hepatitis C guidelines had greatest increases in grade I recommendations (+383% and +67%, respectively). The greatest increases in treatment recommendations were from HBV (grade I, +1150%), liver transplantation (grade II, +112%) and AIH (grade III, +105%).

Conclusions—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.

Address Correspondence to: Christopher Koh, MD MHSc, Liver Diseases Branch, National Institute of Diabetes & Digestive & Kidney Diseases, National Institutes of Health, Bldg. 10, Room 9B-16, 10 Center Drive, MSC 1800, Bethesda, MD 20892-1800, Telephone: 301-496-1721, Fax: 301-402-0491, Christopher.koh@nih.gov.

Keywords

Management; Development; Evidence-based; Medicine

Introduction

Clinical practice guidelines are systematically developed statements that attempt to synthesize large amounts of available scientific information for providing best practices to healthcare providers.[1] These statements often represent the official opinion of single or multiple professional societies and are developed by individuals recognized for their expertise and contributions to the field. Topics often covered include conditions (diseases, signs and symptoms) and technologies (diagnostic tests and therapeutic procedures) where recommendations about preferred approaches for patient management are provided. The creation of recommendations is often based on a formal review and analysis of the published literature along with weighing the strength of the available scientific evidence. In situations where the data is inconclusive or absent, recommendations are often based on consensus expert opinion.

Internationally, more than 3,700 clinical practice guidelines from 39 countries are identified within the Guidelines International Network database.[2] In the U.S., there are over 2,300 guidelines registered within the National Guidelines Clearinghouse which is supported by the Agency for Healthcare Research and Quality (AHRQ).[3] Given the variability in terms of breadth and depth from available clinical practice guidelines, the U.S. Congress has identified the importance of establishing rigorous processes for developing trustworthy, consistent, and scientifically valid documents. In turn, the Institute of Medicine (IOM) released eight standards for the development of clinical practice guidelines in March 2011.[4] Within the framework of the IOM's recommendations, there has been little systematic review of the body of clinical practice guidelines put forth by various medical societies. Recently, clinical practice guideline catalogs from the American College of Cardiology (ACC)/American Heart Association (AHA) and all endocrinology guidelines published in North America from 2007–2010 have been examined.[5,6]

The field of hepatology has experienced significant growth in the production of relevant scientific literature over the past few decades. However, the question of whether clinical practice guidelines have truly evolved with more evidence-based recommendations has not been systematically investigated. Thus, we performed a systematic review of the American Association for the Study of Liver Diseases (AASLD) clinical practice guidelines issued from January 1998 to August 2012 with the aim of evaluating the evolution of recommendations that have been issued over time. The ultimate goal was to evaluate methodological rigour and quality of reporting of AASLD guidelines, elucidate possible gaps that limit the use of evidence-based medicine to support certain recommendations within the AASLD guidelines and to highlight potential opportunities for improvement.

Methods

Guideline Selection

All initial published versions of the AASLD practice guidelines for a given topic issued from January 1998 to August 1, 2012 were abstracted for data.[7–23] If available, the current updated versions for each topic was also evaluated.[18,24–34] Current AASLD guidelines are defined as the most recently published document on a specific topic which is posted on the AASLD website as of August 1, 2012 (http://www.aasld.org). For this investigation, only complete clinical practice guidelines and position papers were evaluated, thus focused updates were not included.

Evaluation of Methodological Rigour and Transparency

To evaluate the evolutionary process of guideline development and quality of reporting, the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument was utilized on all comparable guidelines and position papers.[35] The AGREE II has been widely utilized in the assessment of methodological rigour and transparency of guideline development and has been cited for its validity and reliability. Briefly, this tool that evaluates 23 items organized into six domains (scope and purpose, stakeholder involvement, rigour of development, clarity of presentation, applicability, and editorial independence) followed by 2 global rating items (overall assessment) and includes a user manual that provides guidance on rating of each item. The scope and purpose domain evaluates the specific health questions covered by the guideline, target population and the overall objective of the guideline. The stakeholder involvement domain evaluates the appropriateness of the guideline development group and its representation of the views of its intended users. The rigour of development domain evaluates the systemic methodology utilized to gather and synthesize evidence, methods of recommendation formulation and the mechanisms to update them. The clarity of presentation domain evaluates the overall structure, format and language of the guideline. The applicability domain evaluates barriers, facilitators and ease of implementation and resource implications of guideline application. Finally, the editorial independence domain evaluates the extent to which external influences or competing interests may have affected the specific guideline.

For this study, three appraisers conducted the assessment (CK, SS, NS) after utilizing the online training tools recommended by the AGREE collaboration. After guideline evaluation, domain scores were calculated (as per the AGREE II manual) by summing all individual scores in each domain and then scaling the total as a percentage of the maximum possible score for a given domain according to the formula:

$$\frac{(\text{Obtained score} - \text{Minimum Possible Score})}{(\text{Maximum Possible Score} - \text{Minimum Possible Score})} \text{X100}$$

Evaluation of Strength of Recommendations

All guideline recommendations published by the AASLD are classified by a “grade” or “level” of recommendation. The “grade” or “level” designations are synonyms and provide

an assessment of strength or certainty for a given recommendation. For the purposes of this study, the grade/level designation will be designated as “grade” hereafter.

Since 1998, the AASLD practice guideline development program has utilized three evidence classification systems to grade recommendations. These include 1) the Infectious Diseases Society of America’s Quality Standards, 2) the American College of Cardiology /American Heart Association system, and 3) the Grading of Recommendation Assessment, Development, and Evaluation (GRADE) workgroup system (Table 1).[36–39] Despite the utilization of three systems, these schemes are based on the same criteria and comparable structure. Therefore, for the purposes of this study, a composite grade system was created to represent all of the issued recommendations:

I. Data derived from multiple randomized controlled trials, or meta-analysis, involving a number of participants to be of statistical power and where further research is unlikely to change the confidence in the estimate of clinical effect.

II. Data derived from a single randomized trial or nonrandomized studies, cohort or case control analytic studies, and multiple time series where further research may change confidence in the estimate of the clinical effect.

III. Evidence based on clinical experience, descriptive studies, opinion of respected authorities where further research is very likely to impact confidence on the estimate of clinical effect.

Evaluation of Types of Recommendations

Another aim of this study was to evaluate the evolution of the type of recommendations issued by the AASLD. Recommendations provided in AASLD practice guidelines can be classified into three types:

1. Recommendations based on known features of a given liver disease which should prompt further evaluation (i.e.: “Wilson Disease must be excluded in any patient with unexplained liver disease along with neurological or neuropsychiatric disorder[33]”).
2. Recommendations on specific testing for a given liver disease (i.e.: “Liver biopsy is recommended to stage the degree of liver disease in C282Y homozygotes or compound heterozygotes if liver enzymes (ALT, AST) are elevated or if ferritin is >1000 μg/L[30]”).
3. Recommendations on specific treatment for a given liver disease (i.e.: “UDCA in a dose of 13–15 mg/kg/day orally is recommended for patients with PBC who have abnormal liver enzyme values regardless of histological stage[31]”).

Thus, all recommendations for this analysis were classified into one of three categories: 1) Feature of Disease Recommendation; 2) Diagnostic Recommendation; or 3) Treatment Recommendation.

Evaluation of Benefit versus Risk of Recommendations

As previously discussed, three different guideline classification systems have been utilized during the evolution of AASLD practice guidelines. Depending on the system utilized, certain guidelines provided information regarding benefit versus risk for a given recommendation. This information is different from the “grade” of recommendation and was designated as the “class” of recommendation. In the final part of this analysis, we evaluated the evolution of “class” recommendations provided in multiple versions of guidelines for a specific liver disease topic. However, unlike the grade systems assessing strength and certainty, the “class” systems utilized over time differed greatly and the development of a composite scoring system could not be created for comparative analysis. Therefore, the “class” analysis was only performed on guidelines that utilized the same scoring system.

Results

Historical Guideline Summary

From January 1998 to August 1, 2012, the AASLD issued 28 clinical practice guidelines on 17 topics, yielding a total of 991 recommendations. When examining the initial publication for each AASLD guideline topic, a total of 512 recommendations were issued. The three guidelines with the greatest number of recommendations include Vascular Disorders of the Liver (64), Hepatitis C (HCV) (49) and The Diagnosis and Management of Non-Alcoholic Fatty Liver Disease (NAFLD) (45). Of these 512 recommendations, 14% were grade I recommendations, 40% were grade II and 46% were grade III (Table 2). Regarding the types of recommendations, 14% were Feature of Disease recommendations, 28% were Diagnostic Recommendations, and 58% were Treatment Recommendations (Supplemental Table 1).

Current Guideline Summary

As of August 1, 2012, 17 AASLD guidelines were published or updated between 2005–2012, with a total of 699 recommendations identified. The greatest number of recommendations came from the Chronic Hepatitis B (HBV)(98), Liver Transplantation (78) and HCV (70) guidelines (Table 2).

In evaluating the grade of recommendations for current guidelines, 16% were grade I, 44% were grade II, and 40% were grade III recommendations (Table 2). Individually, grade II recommendations represented the majority of recommendations in 13 of 17 guidelines. The only guideline with a majority of grade I recommendations was the Prevention and Management of Gastrointestinal Varices and Variceal Hemorrhage in Cirrhosis guideline (40%). In contrast, the Autoimmune Hepatitis (AIH), ALF and Liver Biopsy guidelines had a majority of grade III recommendations (Table 2).

Methodological Rigour and Transparency Evaluation

Of the 17 guideline topics published by the AASLD, 11 had initial published versions along with complete updates that were available for comparison utilizing the AGREE II assessment tool. In this comparison, most guideline topics experienced increases in the six domains evaluated by the AGREE II (range 0–53%) along with improvements in the overall assessment (range 0–33%) (Table 3). As a whole, the *editorial independence* domain had the

greatest percentage increases in all guideline topics (10 of 11 topics), however it was the worst scoring domain of current guidelines (range 39–64%). The HBV guidelines had the most improvement in terms of percentage change (5 of 6 domains). In evaluating the overall quality of current guidelines based on domains, the Role of Transjugular Intrahepatic Portosystemic Shunt in the Management of Portal Hypertension (TIPS) guideline had the highest domain score for stakeholder involvement whereas the HBV guideline had the highest domain scores for: scope and purpose, rigour of development, clarity of presentation, applicability and editorial independence (shared with PBC) (Table 3).

Comparison of Recommendation Grades between Current Guidelines

Current AASLD guidelines were evaluated by grade of recommendation (strength) with the type of recommendation (Feature of Disease Recommendation, Diagnostic Recommendation and Treatment Recommendation). In this evaluation, the most frequent types of recommendation were Treatment Recommendations (61%) followed by Diagnostic Recommendations (25%) and Features of Disease Recommendations (15%) (Supplemental Table 2).

The Treatment Recommendation category had a predominance of grade II recommendations (39%), followed by grade III (37%) and grade I (24%) recommendations. The guidelines that contributed the most to the overall number of Treatment Recommendations were HBV (18%) and HCV (12%) practice guidelines.

In the Diagnostic Recommendation category, grade II recommendations were most commonly observed (54%) followed by grade III (40%) and grade I (6%) (Supplemental Table 2). The greatest proportion of diagnostic recommendations came from the HBV, NAFLD and vascular disorders of the liver (12% for all) guidelines.

In the Feature of Disease Recommendation category, the majority of recommendations were grade II (52%), followed by grade III (42%) and grade I (6%) (Supplemental table 2). The greatest proportion of Feature of Disease recommendations were found in the Liver Transplantation (27%) and vascular disorders of the liver (19%) guidelines.

Comparison of Recommendation Grades Between Initial and Current Guidelines

Among 17 guideline topics, 11 documents have complete updates that were eligible for comparison. The average time elapsing from initial publication to the current version of the guidelines was 7.2 years (range, 5–11 years). In these 11 topics, the overall number of recommendations increased by 124% (from 292 to 654 recommendations). All of the guideline topics had an increase in the number of recommendations over time except for Primary Biliary Cirrhosis (PBC), which had a 47% decrease (Table 4). The three guidelines with the greatest increase in the number of recommendations were HBV (+71, 263%), Liver Transplantation (+53, 212%) and AIH (+27, 117%).

In evaluating individual guideline topics for the greatest change in number of grade I recommendations, the HBV guideline had the greatest increase (+23, 383%) followed by HCV (+6, 67%) increase (Table 4). In contrast, the Management of Adult Patients with

Ascites due to Cirrhosis guideline had a 25% decrease in the number of grade I recommendations.

For grade II recommendations, the greatest increase was observed with the Liver Transplantation guideline (+44, 4500%) followed by HBV (+25, 192%) and finally HCV (+16, 107%)(Table 4). By contrast, the guidelines covering topics such as Hepatocellular carcinoma (HCC), Hemochromatosis and TIPS guidelines had a reduced number of grade II recommendations. The greatest increase in grade III recommendations was observed with the AIH guideline (+26, 200%) followed by HBV (+23, 287.5%) and Liver Transplantation (+8, 33.3%)(Table 4). The guidelines focused on PBC, Wilson Disease, and HCV had a decrease in the number of grade III recommendations between initial and revised versions.

Comparison of Recommendation Grades Between Initial and Current Guidelines According to Type of Recommendation

In this comparison, the grade of recommendations (strength) between initial and current guidelines were evaluated based on the type of recommendation (Features of Disease Recommendation, Diagnostic Recommendation and Treatment Recommendation). In the Feature of Disease Recommendation category, the Liver Transplantation guideline had the greatest overall increase in the number of recommendations (+19, 271%), most of which consisted of grade II recommendations (Supplemental Table 3). This was followed by AIH and HCV, which also saw the greatest increases in grade II recommendations.

In the Diagnostic Recommendation category, the greatest numerical increase was again seen in the Liver Transplantation guideline (+16, 800%), followed by HBV (+11, 122%) and AIH (+4, 133%)(Supplemental Table 3). Notably, all three guidelines had the greatest increases in grade II recommendations. In contrast, the PBC and HCC guidelines had a decrease in the number of diagnostic recommendations from initial to current versions.

In the Treatment Recommendation category, the HBV guideline had the greatest increase in recommendations (+58, 387%), most notably with grade I recommendations (Supplemental Table 3). This was followed by AIH (+21, 105%) with a predominant increase in grade III recommendations, and the liver transplantation (+18, 112%), which had a notable increase in grade II recommendations.

Evaluation of Classes of Evidence

Since the introduction of evidence classes to quantify benefit (class I) versus risk (class III), a total of 12 out of 17 AASLD guideline topics have utilized the “classes of evidence” system in at least one version of the publication. In the initial publication for a given guideline topic, 10 out of 17 topics utilized this system. The initial guidelines developed between 2001–2005 did not utilize the “classes of evidence” system. Only 3 of 17 guideline topics (Management of Ascites, Hemochromatosis and PBC) with initial and recent versions, continued to use the class system. However, since different class systems were utilized on subsequent guideline revisions, a direct comparison was not possible.

Of the current guidelines that utilized the classes of evidence system in their recommendations, 9 of the 12 guideline topics utilized the ACC/AHA system while the

other 3 (Hemochromatosis, Primary Sclerosing Cholangitis (PSC), and NAFLD) utilized the GRADE system (Table 5). In the ACC/AHA system, 327 recommendations were issued with 214 (65.4%) designated as class I recommendations suggesting evidence and/or general agreement that a given diagnostic evaluation, procedure or treatment is beneficial, useful, and effective (Table 5). In evaluating the classes of evidence system based on types of recommendations, 64% were treatment recommendations, 23% were diagnostic recommendations and 13% were features of disease recommendations (Supplemental Table 4). In the GRADE classes of evidence system, a total of 98 recommendations were provided and 89% of the recommendations were designated as class I recommendations (Supplemental Table 4).

Discussion

The AASLD clinical practice guidelines provide a set of recommendations for guidance in managing patients with acute and chronic liver disease. Since 1998, these guidelines have provided an additional 36% increase in the overall number of recommendations from the initial development of specific guidelines. However, despite this substantial increase, less than 15% of all recommendations are categorized as grade I, suggesting that the evidence used to develop most recommendations does not come from randomized controlled trials for a variety of reasons. Therefore, areas with insufficient evidence where randomized trials can be conducted to improve the evidence base should be identified for development.

In utilizing the AGREE II guideline assessment tool for assessing methodological rigour and transparency, we identified both global and domain specific improvements in guideline quality from documents created from 1998 to 2012. The current AASLD guidelines appear either comparable or superior by AGREE II evaluation with other medical specialties both nationally and globally that have undergone similar evaluation.[40–42] This assessment demonstrates the AASLD's commitment on continued review of its recommendations along with improving the overall quality of its published guidelines for clinical use.

On AGREE II evaluation, the greatest percentage of improvement in the six different domains was found in *editorial independence*, although its performance was the least impressive among domains assessed by this evaluation. This domain relates to the formulation of recommendations not being unduly biased with competing interests. This measure exemplifies how conflict of interest has become a major issue in the development of practice guidelines, especially when 40% of recommendations within the current AASLD guidelines require input from expert clinicians (as shown by the number of grade III recommendations). Thus, in accordance with the findings of the IOM's recommendations[4], the AASLD has developed and revised a detailed policy for assessing conflict of interest in identifying writing group members for current guidelines being developed and revised which has reduced the potential effects of bias in these documents. However, there will continue to be room for improvement with future guidelines.

In this analysis, the greatest increases in the overall number of recommendations were from practice guidelines related to HBV, liver transplantation and AIH. Given that there is an estimated 350 million persons worldwide infected with HBV where the risk for cirrhosis and

hepatocellular carcinoma is measurable, it is reasonable to expect that a large volume of research is performed in this area.[27] Extensive research of HBV has resulted in a wide array of tools at the clinician's disposal: diagnostic tests for evaluation and monitoring of disease, vaccination to decrease future prevalence of disease, and multiple treatment modalities including interferon and nucleos(t)ide analogues. These observations coincide temporally with current HBV practice guidelines containing the greatest increases in grade I recommendations overall and the greatest increase in the number of treatment recommendations. Similarly, the second largest increase in grade I recommendations was observed within the HCV guideline in association with the approval of direct-acting antiviral agents (DAA) and genetic host factors such as IL-28B. The next complete revision of the HCV guidelines is expected to have even greater increases in both the overall number and grade I recommendations based on continued advances in HCV research.

It is also not surprising that the AASLD guidelines on liver transplantation had a large increase in the number of recommendations from initial to updated publication. Prior to the era of liver transplantation, patients with advanced liver disease usually died within months to years.[34] Now, many patients have the opportunity for extended survival with excellent quality of life after liver transplantation. Interestingly, the increased number of recommendations were dominated by grade II statements and no increases in grade I recommendations.

The third greatest increase in the number of recommendations between guidelines occurred within the topic of AIH. Since the initial 2002 guidelines, additional work in this field such as a modification of the original scoring system of the International Autoimmune Hepatitis Group, enhanced diagnostic serologic testing, and new data leading to multiple recommendations on therapy including the management of refractory disease. Despite the large increase in number of recommendations on this topic, the majority are still grade III in nature. A number of these recommendations will not likely undergo evaluation by randomized clinical trials (i.e. those related diagnosis), but additional randomized trials for therapies including those used for refractory disease would be most welcomed.

Although most guidelines have evolved with increased numbers of recommendations, the PBC and Management of Adult Patients with Ascites in Cirrhosis guidelines had a decrease in grade I recommendations. In the PBC guideline, the overall decrease of recommendations can be attributed to a >70% decrease in grade III recommendations with only minor increases in grade I and II recommendations. In the Management of Adult Patients with Ascites in Cirrhosis guideline, there was a 25% decrease in grade I recommendations because of the withdrawal of a recommendation in the management of tense ascites and a separate recommendation on serial therapeutic paracentesis where the strength of available evidence was demoted in the current version of the guideline. Both of these changes are examples of where recommendations are eliminated over time when evidence and/or practices do not support prior recommendations.

In evaluating the classes of evidence (risk versus benefit), a direct comparison between initial topic guidelines and current guidelines was not possible. To improve their utility for clinicians and facilitate future comparisons, subsequent guideline revisions should consider

moving to a simplified class system that could be applied to all liver disease topics. Such a standardized method of assessing risk and benefit for each individual recommendation would aid clinicians in the delivery of optimal patient care. The current GRADE system may satisfy many of these attributes.

Implications for Guideline Writing

The current AASLD format is to develop comprehensive practice guidelines focusing on assisting practitioners with the diagnosis and management of acute and chronic liver disease. It is expected to have varying degrees of strong or weak recommendations based on varying levels of evidence as few interventions have been subjected to randomized controlled trials. While the goal in theory is to optimize medical management and improve patient care, it is common in practice to follow recommendations based on lower strengths of evidence as shown by similar guidelines developed in other areas of medicine.[5,6]

The overall increase in number of recommendations is also likely due to the growing complexity in the diagnosis and treatment of liver disease. Atypical or variable presentations of disease, differential responses to therapy, and unique aspects within special populations including children and the elderly would require more definitive guidelines to aid the clinicians. Thus, with increasing evidence will come greater numbers of recommendations and perhaps stronger recommendations. However, regardless of the type of evidence, the quality of future clinical practice guidelines can be further improved, as identified by domains evaluated in the AGREE II instrument.

Study Limitations

The current analysis does not account for changes over time regarding the aims and practices of AASLD practice guideline development program, whereby the numbers of recommendations and distribution across classes may have been influenced. Given the lengthy time span, turnover of writing groups and the use of several grading systems in these guidelines, there may have been unanticipated changes in definitions, standards and thresholds in the determination of grades of recommendations that were not easily measurable. Additionally, the sporadic use of class systems and significant changes between systems prohibited a comprehensive class comparison. With the adoption of the current GRADE system for recent and future guideline updates by the AASLD, the deficiencies in assessing quality of evidence and strength of recommendations will hopefully be alleviated.

Conclusions

The evolution of the AASLD practice guidelines is featured by a substantial increase in the overall number of recommendations to assist health care providers on management of patients with liver disease. With the exception of practice guidelines focused on chronic viral hepatitis (HBV and HCV), the bulk of evidence for these recommendations still derive from observational studies or expert consensus opinions. Ideally, the basis of medical practice should be as evidence-based as possible and we should aim to perform the highest quality research to answer clinical dilemmas whenever feasible. Nonetheless, guideline development should continually strive to generate recommendations with the highest quality of evidence possible while minimizing the effect of bias from extrinsic sources.

Additionally, guideline developers should continue to strive to produce highest quality documents with compelling methodological rigor and transparency. Whenever possible, clinical practice guidelines should highlight the need for additional research agenda to fill gaps within clinical care that have the greatest impact on patient outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: This research was supported by the Intramural Research Programs of the NIDDK, NIH. None of the authors has any conflict of interest related to this research.

Abbreviations

AASLD	American Association for the Study of Liver Diseases
HCV	Hepatitis C
HBV	Hepatitis B
ALF	Acute liver failure
AIH	Autoimmune hepatitis
HCC	hepatocellular carcinoma
PBC	Primary biliary cirrhosis
PSC	Primary sclerosing cholangitis
TIPS	Transjugular intrahepatic portosystemic shunt
NAFLD	Non-alcoholic fatty liver disease
AHRQ	Agency for Healthcare Research and Quality
IOM	Institute of Medicine
ACC	American College of Cardiology
AHA	American Heart Association
GRADE	Grading of Recommendation Assessment, Development, and Evaluation

References

1. Field, M.; Lohr, K. Guidelines for Clinical Practice: From Development to Use. Washington, DC: National Academy Press; 1992.
2. Guidelines International Network. [Accessed January 31, 2012] International Guidelines Library. 2012. from http://www.g-i-n.net/library/international-guidelines-library.
3. Agency for Healthcare Research and Quality. [Accessed January 1, 2012] National Guidelines Clearinghouse. 2012. from http://www.guideline.gov/browse/index.aspx?alpha=A. (2012, at http://www.g-i-n.net/library/international-guidelines-library.)

4. Institute of Medicine of the National Academies. [Accessed January 31, 2012] Clinical Practice Guidelines We Can Trust, Consensus Report. 2011. from http://iom.edu/Reports/2011/Clinical-Practice-Guidelines-We-Can-Trust.aspx.
5. Bancos I, Cheng T, Prokop LJ, Montori VM, Murad MH. Endocrine clinical practice guidelines in North America. A systematic assessment of quality. Journal of clinical epidemiology. 2012
6. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr. Scientific evidence underlying the ACC/AHA clinical practice guidelines. Jama. 2009; 301:831–841. [PubMed: 19244190]
7. Czaja AJ, Freese DK. Diagnosis and treatment of autoimmune hepatitis. Hepatology (Baltimore, Md. 2002; 36:479–497.
8. Polson J, Lee WM. AASLD position paper: the management of acute liver failure. Hepatology (Baltimore, Md. 2005; 41:1179–1197.
9. Runyon BA. Management of adult patients with ascites caused by cirrhosis. Hepatology (Baltimore, Md. 1998; 27:264–272.
10. Lok AS, McMahon BJ. Chronic hepatitis B. Hepatology (Baltimore, Md. 2001; 34:1225–1241.
11. Bruix J, Sherman M. Management of hepatocellular carcinoma. Hepatology (Baltimore, Md. 2005; 42:1208–1236.
12. Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C. Hepatology (Baltimore, Md. 2004; 39:1147–1171.
13. Tavill AS. Diagnosis and management of hemochromatosis. Hepatology (Baltimore, Md. 2001; 33:1321–1328.
14. Heathcote EJ. Management of primary biliary cirrhosis. The American Association for the Study of Liver Diseases practice guidelines. Hepatology (Baltimore, Md. 2000; 31:1005–1013.
15. Boyer TD, Haskal ZJ. The role of transjugular intrahepatic portosystemic shunt in the management of portal hypertension. Hepatology (Baltimore, Md. 2005; 41:386–400.
16. Carithers RL Jr. Liver transplantation. American Association for the Study of Liver Diseases. Liver Transpl. 2000; 6:122–135. [PubMed: 10648593]
17. Roberts EA, Schilsky ML. A practice guideline on Wilson disease. Hepatology (Baltimore, Md. 2003; 37:1475–1492.
18. DeLeve LD, Valla DC, Garcia-Tsao G. Vascular disorders of the liver. Hepatology (Baltimore, Md. 2009; 49:1729–1764.
19. Garcia-Tsao G, Sanyal AJ, Grace ND, Carey W. Prevention and management of gastroesophageal varices and variceal hemorrhage in cirrhosis. Hepatology (Baltimore, Md. 2007; 46:922–938.
20. Chapman R, Fevery J, Kalloo A, et al. Diagnosis and management of primary sclerosing cholangitis. Hepatology (Baltimore, Md. 2010; 51:660–678.
21. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. Liver biopsy. Hepatology (Baltimore, Md. 2009; 49:1017–1044.
22. O'Shea RS, Dasarathy S, McCullough AJ. Alcoholic liver disease. Hepatology (Baltimore, Md. 2010; 51:307–328.
23. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology (Baltimore, Md. 2012; 55:2005–2023.
24. Manns MP, Czaja AJ, Gorham JD, et al. Diagnosis and management of autoimmune hepatitis. Hepatology (Baltimore, Md. 2010; 51:2193–2213.
25. Lee WM, Stravitz RT, Larson A. Introduction to the revised AASLD position paper on acute liver failure 2011. Hepatology (Baltimore, Md. 2011
26. Runyon BA. Management of adult patients with ascites due to cirrhosis: an update. Hepatology (Baltimore, Md. 2009; 49:2087–2107.
27. Lok AS, McMahon BJ. Chronic hepatitis B: update 2009. Hepatology (Baltimore, Md. 2009; 50:661–662.
28. Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology (Baltimore, Md. 2009; 49:1335–1374.

29. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology (Baltimore, Md. 2011; 53:1020–1022.
30. Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. Hepatology (Baltimore, Md. 2011; 54:328–343.
31. Lindor KD, Gershwin ME, Poupon R, Kaplan M, Bergasa NV, Heathcote EJ. Primary biliary cirrhosis. Hepatology (Baltimore, Md. 2009; 50:291–308.
32. Boyer TD, Haskal ZJ. The Role of Transjugular Intrahepatic Portosystemic Shunt (TIPS) in the Management of Portal Hypertension: update 2009. Hepatology (Baltimore, Md. 2010; 51:306.
33. Roberts EA, Schilsky ML. Diagnosis and treatment of Wilson disease: an update. Hepatology (Baltimore, Md. 2008; 47:2089–2111.
34. Murray KF, Carithers RL Jr. AASLD practice guidelines: Evaluation of the patient for liver transplantation. Hepatology (Baltimore, Md. 2005; 41:1407–1432.
35. Consortium ANS. The AGREE II Instrument [Electronic Version]. 2009
36. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008; 336:924–926. [PubMed: 18436948]
37. Gross PA, Barrett TL, Dellinger EP, et al. Purpose of quality standards for infectious diseases. Infectious Diseases Society of America. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 1994; 18:421. [PubMed: 8011826]
38. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. Annals of internal medicine. 2003; 139:493–498. [PubMed: 13679327]
39. Association. ACoCAH. Methodology Manual for ACC-AHA Guideline Writing Committees: Methodologies and Policies from the ACC/AHA Task Force on Practice Guidelines. 2006 Apr.
40. Holmer HK, Ogden LA, Burda BU, Norris SL. Quality of clinical practice guidelines for glycemic control in type 2 diabetes mellitus. PloS one. 2013; 8:e58625. [PubMed: 23577058]
41. Acuna-Izcaray A, Sanchez-Angarita E, Plaza V, et al. Quality assessment of asthma clinical practice guidelines: a systematic appraisal. Chest. 2013
42. Al-Ansary LA, Tricco AC, Adi Y, et al. A systematic review of recent clinical practice guidelines on the diagnosis, assessment and management of hypertension. PloS one. 2013; 8:e53744. [PubMed: 23349738]

Table 1

Historical summary of grades of evidence and classes of recommendations used by the AASLD

Grade of Evidence
1998–2003
I = Evidence from multiple well designated randomized controlled trials each involving a number of participant to be of sufficient statistical power.
II = Evidence from at least one large well-designed clinical trial with or without randomization, from cohort or case- control analytic studies, or well-designated meta-analysis
III = Evidence based on clinical experience, descriptive studies, or reports of expert committees.
IV = Not Rated
2004–2010
I = Randomized controlled trials
II-1 = Controlled trials without randomization
II-2 = Cohort or case-control analytic studies
II-3 = Multiple time series, dramatic uncontrolled experiments
III = Opinion of respected authorities, descriptive epidemiology
2007–2010
A = Data derived from multiple randomized clinical trials or meta-analysis
B = Data derived from a single randomized trial, or nonrandomized studies
C = Only consensus opinion of experts, case studies, or standard-of-care.
2010-Present
High (A) = Further research is unlikely to change confidence in the estimate of the clinical effect
Moderate (B) = Further research may change confidence in the estimate of the clinical effect.
Low (C) = Further research is very likely to impact confidence on the estimate of clinical effect.
Class of Recommendations
1998–2000
A = Survival Benefit
B = improved diagnosis
C = improvement in quality of life
D = Relevant pathophysiologic parameters improved
E = impacts cost of health care
2007–2010
I = Conditions for where there is evidence and/ or general agreement that a given diagnostic evaluation, procedure or treatment is beneficial, useful, and effective
II = Conditions for which there is conflicting evidence and/or divergence of opinion about the usefulness/efficacy of a diagnostic evaluation, procedure, or treatment
IIa = Weight of evidence/opinion is in favor of usefulness/efficacy
IIb = Usefulness/efficacy is less well established by evidence/opinion
III = Conditions for which there is evidence and/or general agreement that a diagnostic evaluation/procedure/treatment is not useful/effective and in some cases may be harmful
2011-Present

Grade of Evidence
Strong (1) = Factors influencing the strength of the recommendation included the quality of evidence, presumed patient-important outcomes, and cost
Weak (2) = Variability in preferences and values, or more uncertainty. Recommendation is made with less certainty, or higher cost or resource consumption

Table 2

Data summary of the first issued guidelines evaluated by grade and the current guidelines evaluated by grade.

First Issued Guideline Evaluation By Grade					
Guideline	**Year**	**Number of Recommendations**	**Grade I Number**	**Grade II Number**	**Grade III Number**
AIH	2002	23	4 (17%)	6 (26%)	13 (56%)
ALF	2005	44	7 (16%)	7 (16%)	30 (68%)
Ascites	1998	22	8 (36%)	6 (27%)	8 (36%)
HBV	2001	27	6 (22%)	13 (48%)	8 (30%)
HCC	2005	21	4 (19%)	12 (57%)	5 (24%)
HCV	2004	49	9 (18%)	15 (31%)	25 (51%)
Hemochromatosis	2001	12	0 (0%)	12 (100%)	0 (0%)
PBC	2000	22	2 (9%)	1 (4%)	19 (86%)
TIPS	2005	26	2 (7%)	18 (69%)	6 (23%)
Transplantation	2000	25	0 (0%)	1 (4%)	24 (96%)
Wilson Disease	2003	21	0 (0%)	8 (38%)	13 (62%)
Vascular	2009	64	2 (3%)	30 (47%)	32 (50%)
Varices	2007	25	10 (40%)	6 (24%)	9 (36%)
PSC	2010	36	6 (17%)	22 (61%)	8 (22%)
Biopsy	2009	34	0 (0%)	7 (20%)	27 (80%)
Alcoholic Liver	2010	16	6 (37%)	6 (37%)	4 (25%)
NAFLD	2012	45	6 (13%)	34 (76%)	5 (11%)
Summary		512	72 (14%)	204 (40%)	236 (46%)
% median			15.9%	37.5%	36.4%
% quartile 1			3.13%	24.00%	23.81%
% quartile 3			19.05%	57.14%	61.90%
Current Guideline Evaluation By Grade					
Guideline	**Year**	**Number of Recommedations**	**Grade I Number**	**Grade II Number**	**Grade III Number**
AIH	2010	50	5 (10%)	6 (12%)	39 (78%)
Alcoholic Liver	2010	16	6 (37%)	6 (37%)	4 (25%)

First Issued Guideline Evaluation By Grade					
Guideline	**Year**	**Number of Recommendations**	**Grade I Number**	**Grade II Number**	**Grade III Number**
ALF	2011	48	7 (15%)	7 (15%)	34 (71%)
Ascites	2009	30	6 (20%)	13 (43%)	11 (37%)
Biopsy	2009	34	0 (0%)	7 (21%)	27 (79%)
HBV	2009	98	29 (30%)	38 (39%)	31 (32%)
HCC	2010	22	5 (23%)	11 (50%)	6 (27%)
HCV	2009	70	15 (21%)	31 (44%)	24 (34%)
Hemochromatosis	2011	17	4 (23%)	10 (59%)	3 (18%)
PBC	2009	15	5 (33%)	5 (33%)	5 (33%)
PSC	2010	36	6 (17%)	22 (61%)	8 (22%)
TIPS	2009	28	5 (18%)	17 (61%)	6 (21%)
Transplantation	2005	78	0 (0%)	46 (59%)	32 (41%)
Varices	2007	25	10 (40%)	6 (24%)	9 (36%)
Vascular	2009	64	2 (3%)	30 (47%)	32 (50%)
Wilson Disease	2008	23	1 (4%)	16 (70%)	6 (26%)
NAFLD	2012	45	6 (13%)	34 (76%)	6 (13%)
Summary		699	112 (16%)	305 (44%)	283 (40%)
% median			17.9%	44.3%	33.3%
% quartile 1			8.59%	33.33%	25.00%
% quartile 3			23.53%	58.97%	41.03%

Hepatology. Author manuscript; available in PMC 2015 October 21.

Table 3

Evolutionary Rigour and Transparency Evaluation by AGREE II

	Domain 1. Scope and Purpose %	Domain 2. Stakeholder Involvement %	Domain 3. Rigour of Development %	Domain 4. Clarity of Presentation %	Domain 5. Applicability %	Domain 6. Editorial Independence %	Overall Guideline Assessment %
AIH (2002)	74	54	58	81	58	6	56
AIH (2010)	76	59	60	85	65	39	67
Percentage Change	2	6	3	4	7	33	11
ALF (2005)	83	56	56	81	50	25	72
ALF (2011)	83	56	56	85	58	50	72
Percentage Change	0	0	1	4	8	25	0
Ascites (1998)	76	56	63	80	67	11	72
Ascites (2008)	80	57	65	85	69	42	83
Percentage Change	4	2	2	6	3	31	11
HBV (2001)	87	59	59	80	67	11	72
HBV (2009)	100	59	72	100	79	64	100
Percentage Change	13	0	13	20	13	53	28
HCC (2005)	83	56	63	85	75	50	72
HCC (2010)	85	59	63	85	78	50	89
Percentage Change	2	4	0	0	3	0	17
HCV (2004)	85	57	64	85	71	42	89
HCV (2009)	85	57	69	85	76	53	89
Percentage Change	0	0	5	0	6	11	0
Hemochromatosis (2001)	70	56	59	74	72	22	72
Hemochromatosis (2011)	80	57	67	85	72	47	83
Percentage Change	9	2	8	11	0	25	11
PBC (2000)	70	50	56	67	61	11	56
PBC (2009)	81	56	63	85	67	64	89
Percentage Change	11	6	7	19	6	53	33
TIPS (2005)	81	63	59	81	76	44	72

	Domain 1. Scope and Purpose %	Domain 2. Stakeholder Involvement %	Domain 3. Rigour of Development %	Domain 4. Clarity of Presentation %	Domain 5. Applicability %	Domain 6. Editorial Independence %	Overall Guideline Assessment %
TIPS (2009)	83	63	69	85	76	56	83
Percentage Change	2	0	10	4	0	11	11
Transplant (2000)	83	44	54	70	56	17	50
Transplant (2005)	89	59	60	81	67	56	83
Percentage Change	6	15	6	11	11	39	33
Wilson (2003)	81	46	60	74	61	22	72
Wilson (2008)	85	48	63	81	64	61	83
Percentage Change	4	2	2	7	3	39	11

Table 4

Grade comparison between first and current guidelines.

Grade Comparison Between First and [Current Guidelines											
Guideline	**Year**	**Number of Recommendations**	**Grade I # (Change)**	**Grade I %**	**Grade I % change %**	**Grade II # (Change)**	**Grade II %**	**Grade II % change %**	**Grade III # (Change)**	**Grade III %**	**Grade III % change %**
AIH	2002	23	4	17%		6	26%		13	57%	
	2010	50	5 (+25%)	10%	−43%	6 (0%)	12%	−54%	39 (+200%)	78%	38%
ALF	2005	44	7	16%		7	16%		30	68%	
	2011	48	7 (0%)	15%	−8%	7 (0%)	15%	−8%	34 (+13%)	71%	4%
Ascites	1998	22	8	36%		6	27%		8	36%	
	2009	30	6 (−25%)	20%	−45%	13 (+117%)	43%	59%	11 (+37%)	37%	1%
HBV	2001	27	6	22%		13	48%		8	30%	
	2009	98	29 (+383%)	30%	33%	38 (+192%)	39%	−19%	31 (+287%)	32%	7%
HCC	2005	21	4	19%		12	57%		5	24%	
	2010	22	5 (25%)	23%	19%	11 (−8%)	50%	−13%	6 (+20%)	27%	15%
HCV	2004	49	9	18%		15	31%		25	51%	
	2009	70	15 (67%)	21%	17%	31 (+107%)	44%	45%	24 (−4%)	34%	−33%
Hemochromatosis	2001	12	0	0%		12	100%		0	0%	
	2011	17	4	24%		10 (−17%)	59%	−41%	3	18%	
PBC	2000	22	2	9%		1	5%		19	86%	
	2009	15	55 (+150%)	33%	267%	5 (+400%)	33%	633%	5 (−74%)	33%	−61%
TIPS	2005	26	2	8%		18	69%		6	23%	
	2009	28	5 (+150%)	18%	132%	17 (−6%)	61%	−12%	6 (0%)	21%	−7%
Transplantation	2000	25	0	0%		1	4%		24	96%	
	2005	78	0 (0%)	0%		46 (+4500%)	59%	1374%	32 (+33%)	41%	−57%
Wilson Disease	2003	21	0	0%		8	38%		13	62%	
	2008	23	1 (+100%)	4%		16 (+100%)	70%	83%	6 (−54%)	26%	−58%
change in #			40			101			46		
change in % median			46%		18%	53%		−8%	17%		−3%
change in % quartile 1			19%		−17%	−1%		−16%	−3%		−51%

Grade Comparison Between First and [Current Guidelines											
Guideline	**Year**	**Number of Recommendations**	**Grade I # (Change)**	**Grade I %**	**Grade I % change %**	**Grade II # (Change)**	**Grade II %**	**Grade II % change %**	**Grade III # (Change)**	**Grade III %**	**Grade III % change %**
change in % Quatile 3			150%		58%	136%		71%	36%		6%

Table 5

Class comparison of current guidelines

Class Comparison of Current Guidelines							
Class I, II, IIA, IIB, III Comparison							
Guideline	**Year**	**Number of Recommendations**	**Class I Number**	**Class II Number**	**Class IIA Number**	**Class IIB Number**	**Class III Number**
AIH	2010	50	14 (28%)	0 (0%)	31 (62%)	1 (2%)	4 (8%)
Alcoholic Liver	2010	16	13 (81%)	0 (0%)	0 (0%)	0 (0%)	3 (20%)
Ascites	2009	30	15 (50%)	0 (0%)	10 (33%)	2 (7%)	3 (10%)
Biopsy	2009	34	30 (88%)	0 (0%)	2 (6%)	2 (6%)	0 (0%)
HCV	2009	70	34 (49%)	1 (1%)	25 (36%)	5 (7%)	5 (7%)
PBC	2009	15	14 (93%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Varices	2007	25	19 (76%)	0 (0%)	3 (12%)	0 (0%)	3 (12%)
Vascular	2009	64	53 (83%)	0 (0%)	3 (5%)	2 (3%)	6 (10%)
Wilson Disease	2008	23	2 2 (96%)	0 (0%)	0 (0%)	1 (4%)	0 (0%)
summary		327	214 (65%)	1 (0.3%)	75 (23%)	13 (4%)	24 (7%)
% median			81%	0%	7%	3%	8%
% quartile 1			50%	0%	5%	0%	0%
% quartile 3			88%	0%	33%	6%	10%

Class 1–2 Comparison				
Guideline	**Year**	**Number of Recommendations**	**Class 1 Number**	**Class 2 Number**
Hemochromatosis	2011	17	15 (88%)	2 (12%)
PSC	2010	36	31 (86%)	5 (14%)
NAFLD	2012	45	39 (87%)	6 (13%)