

# Appraisal of Guidelines Research & Evaluation—Health Systems: AGREE-HS



AGREE-HS Research Team 2018

## COPYRIGHT AND REPRODUCTION

This document is the product of an international collaboration. It may be reproduced and used for educational purposes, quality assurance programmes and critical appraisal of health systems guidance. It may not be used for commercial purposes or product marketing. Offers of assistance in translation into other languages are welcome, provided they conform to the protocol set out by The AGREE Research Trust.

## DISCLAIMER

The AGREE-HS is a tool to assist health systems guidance (HSG) developers, users and policy makers to direct the development, appraisal and reporting of HSG. The authors do not take responsibility for the improper use of the AGREE-HS Instrument.

© 2018

## SUGGESTED CITATION FOR AGREE-HS PUBLICATION:

*Manuscripts related to the AGREE-HS have been submitted to peer-reviewed journals for publication. Citations will be added here when they are available.*

## SUGGESTED CITATION FOR AGREE-HS PDF VERSION:

AGREE-HS Research Team (2018). The Appraisal of Guidelines Research & Evaluation—Health Systems (AGREE-HS) [Electronic version]. Retrieved <Month, Day, Year, from <http://www.agreetrust.org>.

## FUNDING:

The development of the AGREE-HS Instrument was funded by a grant from the CIHR.



FOR FURTHER INFORMATION ABOUT THE  
AGREE-HS, PLEASE CONTACT:  
AGREE Scientific Office, [agree@mcmaster.ca](mailto:agree@mcmaster.ca)  
AGREE Enterprise Website, [www.agreetrust.org](http://www.agreetrust.org)

**Principal Investigators:**

Melissa C. Brouwers - McMaster University, Canada  
John N. Lavis - McMaster University, Canada

**Research Team Members:**

Andy Haines - London School of Hygiene & Tropical Medicine, United Kingdom  
Carmen Mihaela Dolea - World Health Organization, Switzerland  
Denis Ako-Arrey - McMaster University, Canada  
Fadi El-Jardali - American University of Beirut, Lebanon  
Francoise Cluzeau - Global Health and Development Group, Imperial College London, United Kingdom  
Govin Permanand - World Health Organization Regional Office for Europe, Denmark  
Iván D. Flórez - Universidad de Antioquia, Colombia  
Jillian Ross - Cancer Care Ontario, Canada  
Luis Gabriel Cuervo – Pan American Health Organization, United States of America  
Mike Wilson - McMaster University, Canada  
Mita Giacomini - McMaster University, Canada  
Pablo Perel - London School of Hygiene and Tropical Medicine, United Kingdom  
Padraig Warde - Cancer Care Ontario, Canada  
Pierre Ongolo-Zogo - Centre for Development of Best Practices in Health, Cameroon  
Sheila McNair - Program for Evidence-based Care, McMaster University, Canada  
Ulysses Panisset - World Health Organization, Switzerland  
Xavier Bosch-Capblanch - Swiss Tropical and Public Health Institute, Switzerland  
Yaolong Chen - Lanzhou University, China

**Research Staff:**

Karen Spithoff – McMaster University, Canada  
Marija Vukmirovic – McMaster University, Canada

**Acknowledgements:**

Saira Akram – McMaster University, Canada

# AN INTRODUCTION TO THE AGREE-HS

## I. Background

Health systems guidance or guidelines (HSG) documents are systematically developed statements to assist with decisions about appropriate options for addressing health system challenges, the implementation of these options, and the monitoring and evaluation of the implementation efforts<sup>1</sup>. For example, common challenges addressed by HSG are those related to health system arrangements and governance (e.g., policy, professional authority), financial arrangements (e.g., financing, incentives) and delivery arrangements (e.g., to whom, by whom, or where care is provided) for health care and population health services.

The **A**ppraisal of **G**uidelines for **R**esearch and **E**valuation for **H**ealth **S**ystems (AGREE-HS) was designed to:

- Provide a methodological framework for the development of HSG;
- Inform what and how information ought to be reported in HSG; and
- Assess the quality of HSG.

The AGREE-HS was designed through a formal review and analysis of the published literature and a series of studies involving leaders and international stakeholders with experience in the HSG field<sup>2-4</sup>. A full listing of research papers related to the AGREE-HS can be found on the AGREE Enterprise website ([www.agreetrust.org](http://www.agreetrust.org)).

## II. AGREE-HS Overview

### AGREE-HS description

The AGREE-HS tool comprises five items representing key HSG quality domains: topic, participants, methods, recommendations and implementability. Definitions and quality criteria are provided for each item. The tool concludes with two overall assessment statements.

### Which documents can be appraised with the AGREE-HS?

The AGREE-HS can be applied to HSG documents produced at a global (e.g., World Health Organization), regional (e.g., Pan American Health Organization), national (e.g., National Institute for Health and Care Excellence, UK), or provincial/state/local (e.g., Cancer Care Ontario, Ontario, Canada) level. HSG documents provide recommendations related to improving health system governance, or financial or delivery arrangements.

Relevant documents might not be labeled as “HSG”, as a wide range of terms is used to describe guidance documents. Some of these terms include *manual*, *framework*, *review*, *toolkit* and *implementation plan*. As a general point of reference, if a document addresses a health systems challenge and provides recommendations or statements of action, it can be appraised with the AGREE-HS.

The following references provide **examples of HSG documents**:

- **Global:** WHO recommendations: optimizing health worker roles to improve access to key maternal and newborn health interventions through task shifting. World Health Organization 2012; 1-98.
- **Regional:** Technical guidelines for vaccination against the pandemic influenza virus. Pan American Health Organization 2009; 1-68.
- **National:** Home care: delivering personal care and practical support to older people living in their own homes. National Institute for Health and Care Excellence 2015; 1-29.
- **Provincial/state/local:** Murphy J, Varela N, Elit L, Lytwyn A, Wu V, Yudin M, Shier M, El-Khatib S, and the Cervical Cancer Screening Clinical Advisory Committee. The organization of colposcopy services in Ontario: recommended framework. Cancer Care Ontario 2015; 1-56.

## AGREE-HS versus AGREE II

In contrast to the AGREE-HS, that provides structured processes to help optimize the quality and usability of HSG, the AGREE II is a tool for clinical practice guidance. Specifically, the AGREE II is used to support guidelines that address clinical, rather than system, challenges, and that provide recommendations to inform clinician and patient decisions, the clinical encounter, or specific clinical policy<sup>5</sup>. Some guidance documents contain both clinical and health systems recommendations; both tools can be applied to the appropriate sections of these documents.

## Who can use the AGREE-HS?

The AGREE-HS is intended to be used by:

1. **Guidance developers:** The AGREE-HS provides a structured, systematic and standardized methodological framework for developing and reporting HSG.
2. **Policy makers and program managers:** To facilitate the process of translating knowledge into action, the AGREE-HS provides a means for policy makers and program managers to evaluate the quality of the HSG they recommend for use, or use themselves, to inform policy development and implementation, as well as to enhance policy monitoring and evaluation.
3. **Stakeholders:** Patients/consumers, health professionals, researchers, educators and other stakeholders interested in the development and uptake of HSG, can use the AGREE-HS as a tool to enhance their skills in the development, reporting and critical appraisal of HSG.

## AGREE-HS training materials and resources

Training is essential for successful application of the AGREE-HS. Therefore, individuals new to the tool or new to HSG are encouraged to orient themselves before using the tool. Access to AGREE-HS training materials, HSG resources, and the AGREE-HS project will become available on the AGREE Enterprise website ([www.agreetrust.org](http://www.agreetrust.org)) as they are developed.

## References

1. Bosch-Capblanch X, Lavis JN, Lewin S, et al. Guidance for evidence-informed decisions about health systems: rationale for and challenges of guidance development. *PLoS Med* 2012;9:e1001000.
2. Ako-Arrey DE. Better guidance for better health systems: Designing a knowledge translation tool for the development, appraisal and reporting of health systems guidance (doctoral thesis). McMaster University, Canada; 2015. Available at: <http://hdl.handle.net/11375/18034>.
3. Ako-Arrey DE, Brouwers MC, Lavis JN, Giacomini MK. Health systems guidance appraisal—a critical interpretive synthesis. *Implement Sci* 2016;11:9.
4. Ako-Arrey DE, Brouwers MC, Lavis JN, Giacomini MK. Health system guidance appraisal—concept evaluation and usability testing. *Implement Sci* 2016;11:3.
5. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: Advancing guideline development, reporting and evaluation in healthcare. *CMAJ* 2010;182:E839-42.

# AGREE-HS: USER MANUAL

## I. Review of HSG and Accompanying Documents

Before applying the instrument, users should read the HSG in full. Additionally, every effort should be made to identify any supporting documents and resources (e.g., organizations' statements, appendices, data supplements) and all other pertinent information related to the HSG. This supplementary information may be found in the same document as the HSG or it may be summarized in a separate technical report, methodological manual, or guidance developer policy statement. It may have been published elsewhere or made available publicly on web sites.

## II. HSG Appraisers

It is recommended that at least two, and preferably four, appraisers assess each HSG to increase the reliability of the assessment. Ideally, a team of appraisers evaluating the same HSG should have various areas of expertise, to ensure that different perspectives on the HSG topic are represented. Further, some should be familiar with the AGREE-HS, HSG documents, and/or the process of critical appraisal.

First-time users should read this User Manual thoroughly and become familiar with each item, its definition and the criteria associated with it. They should consider applying the tool to a few "practice" HSG documents to ensure consistency in their application and interpretation of items. If reviewing with another appraiser, they should confirm that the items are being applied and interpreted similarly and make adjustments as required, including more specific operationalization of concepts that are appropriate for the context.

## III. Rating Scale

Each of the five AGREE-HS items is rated on the following 7-point scale:

1 Lowest Quality	2	3	4	5	6	7 Highest Quality
---------------------	---	---	---	---	---	----------------------

**Score of 1 (Lowest Quality):** A score of 1 should be given if there is no information that is relevant to the AGREE-HS item, if the criteria are very poorly reported in the HSG document, or if the authors explicitly state that it was not done.

**Score of 7 (Highest Quality):** A score of 7 should be given if the information related to the AGREE-HS item was exceptionally well reported, all criteria related to the item have been considered during the development of the guidance, and the information related to the item is applicable in its context.

**Scores between 2 and 6:** A score between 2 and 6 is assigned when not all criteria of the AGREE-HS item have been met. A score is assigned depending on the completeness and quality of reporting. Scores increase as more criteria are met.

**Comment boxes:** The boxes provided on each item page can be used by appraisers to reference relevant pages in the HSG document and/or to rationalize their item scores. Written comments can help to facilitate discussion with other appraisers of the same HSG, to reach consensus regarding item scores (see Section IV below), and to guide decisions about whether to use or implement the HSG.

**Additional considerations:**

- When assessing whether each item’s criteria have been met, appraisers should take into consideration whether the item-related content is well written, easy to find, and easy to understand.
- When rating the AGREE-HS items, judgments by the appraisers are required. The criteria are provided to guide, not to replace, these judgments. Thus, explicit expectations for each of the 7 points on the rating scale are not provided. Appraisers are encouraged to use the comment boxes to provide justification for their scores. Teams of appraisers may wish to create rules or thresholds for each item appropriate to their context and purpose.
- On occasion, some item criteria may not be applicable to the particular HSG document under review because, for example, they are out of scope. If appropriate, appraisers may rate the item(s) based only on the criteria that are deemed applicable. If any criteria are not applicable, decisions about how to rate the item(s) should be made in advance, described in an explicit manner, and applied consistently by all reviewers of that HSG document. As a general principle, excluding criteria in the appraisal process is discouraged.

## IV. Methods of Scoring HSG Documents Using the AGREE-HS

Two methods can be used to calculate final item scores and an overall score when multiple appraisers assess the same HSG document:

1. Calculating final item scores and an overall score using individual appraisers’ scores.
2. Determining final item scores by consensus and then calculating an overall score.



## i) Scoring Method 1: Using Individual Appraisers' Scores

### Final Item Scores

For this method, each appraiser reviews the HSG document and independently assigns item scores. Final item scores are then calculated by averaging the independently assigned scores (i.e., calculating the mean).

#### Example:

Four appraisers give the following scores for the *Topic* item. The final item score, or average (mean) of the four appraisers' scores, is 5.25.

Item	Appraiser 1	Appraiser 2	Appraiser 3	Appraiser 4	Final item score
Topic	6	5	4	6	5.25

### Overall Score

The overall score is calculated by summing up all the scores of the individual items (the sum of the item scores is referred to as the "obtained score" in the formula below) and by scaling this total as a percentage of the maximum possible score. The overall score is represented by a percentage, calculated as:

$$\frac{\text{Obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} = \text{Overall score}$$

#### Example:

Four appraisers give the following AGREE-HS scores:

Item	Appraiser 1	Appraiser 2	Appraiser 3	Appraiser 4	Total
Topic	6	5	7	6	24
Participants	2	1	2	1	6
Methods	3	3	5	5	16
Recommendations	5	5	5	5	20
Implementability	4	5	4	5	18
<b>Total</b>	<b>20</b>	<b>19</b>	<b>23</b>	<b>22</b>	<b>84</b>

The maximum and minimum possible scores are calculated based on the 7-point scale, the number of items and the number of appraisers.

Maximum possible score = 7 (strongly agree) x 5 (# of items) x 4 (# of appraisers) = 140

Minimum possible score = 1 (strongly disagree) x 5 (# of items) x 4 (# of appraisers) = 20

Therefore, using Scoring Method 1, the overall score for this HSG document would be:

$$\frac{\text{Obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} = \frac{84-20}{140-20} = 53\%$$

## ii) Scoring Method 2: Reaching Consensus

### Final Item Scores

Another option is for the appraisal team to *reach consensus* on final item scores, rather than averaging them across multiple raters. The final item scores are used to calculate an overall score for the HSG document.

### Overall Score

The overall score can be calculated by summing up all the final item scores of the individual items (the sum of the final item scores is referred to as the “obtained score” in the formula below) and by scaling this total as a percentage of the maximum possible score. The overall score will be represented by a percentage, calculated as:

$$\frac{\text{Obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} = \text{Overall score}$$

### Example:

The team of appraisers collectively assigns the final item scores:

Item	Appraisers' final Item scores
Topic	5
Participants	3
Methods	6
Recommendations	4
Implementability	5
<b>Obtained score</b>	<b>23</b>

The maximum and minimum possible scores are calculated based on the 7-point scale and the number of items.

Maximum possible score = 7 (strongly agree) x 5 (# of items) = 35

Minimum possible score = 1 (strongly disagree) x 5 (# of items) = 5

Therefore, using Scoring Method 2, the overall score for this HSG document would be:

$$\frac{\text{Obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} = \frac{23-5}{35-5} = \mathbf{60\%}$$

### iii) Interpreting AGREE-HS Scores

Stakeholders can use final item and overall scores to compare HSG documents, to identify limitations of the guidance being considered, or to select high quality HSG to implement. This manual does not weigh final item scores, nor does it define a threshold overall score to differentiate between high, moderate and low quality HSG. This is because there is no empirical basis to define thresholds.

While more research is required to create empirically derived thresholds to define high quality and low quality HSG, examples of how the data can be interpreted are offered:

- Users could perform a tertile split of the overall scores of the HSG documents being considered and classify documents as being higher quality, medium quality, or lower quality.
- Users may determine threshold scores through consensus. For example, HSG documents with overall scores >70% may be defined as high quality, those with overall scores <30% as low quality, and those between 30% and 70% as moderate quality.
- Users might value one item over the others for their decision-making purposes and create thresholds based on that priority item score.

Decisions about how to define minimum thresholds for quality or applicability should ideally be made by a panel of all relevant stakeholders. Decisions should be guided by the context in which the HSG is to be used and by evaluating the importance of the different items and criteria in that context.

**AGREE-HS:  
APPRAISAL TOOL**

### Item 1: Topic

*This item addresses the description of the health system challenge, the causes of the challenge and the priority accorded to it, and relevance of the guidance.*

#### Criteria:

Item content includes the following criteria:

- The **health system challenge** is clearly described (i.e., the nature of the challenge; the magnitude, frequency or intensity of the challenge; the populations affected).
- The **causes** of the health system challenge are clearly described.
- The health system challenge is described in terms of its **level of priority** in the targeted health system and the affected population; arguments to support the priority classification are provided.
- The guidance is **relevant** to (i.e., timely in relation to when decisions will be made), and **appropriate** for, the health system challenge, the system or sub-system needs, the target population(s), and the setting(s) in which they will operate.

#### Evaluation:

Rate the overall quality of this item.

1 Lowest quality	2	3	4	5	6	7 Highest quality
---------------------	---	---	---	---	---	----------------------

*Comments*

## Item 2: Participants

*This item addresses the composition of the health systems guidance development team and the management of competing interests and funder influence.*

### Criteria:

Item content includes the following criteria:

- The health systems guidance development team includes members who have an **interest or stake** in the recommendations (e.g., decision makers, program managers, operational leaders, consumers and members of the public).
- The health systems guidance development team is **multidisciplinary** (e.g., political scientists, economists, epidemiologists, methodologists).
- The health systems guidance development team is **multi-sectoral** (e.g., primary care, public health and, if appropriate to the challenge, finance and housing).
- **Competing interests** of the health systems guidance development team members (e.g., financial, professional), and the strategies used to identify and manage them, are clearly described.
- Precautions have been taken to avoid or to minimize the **influence of a funding agency**.

### Evaluation:

Rate the overall quality of this item.

1 Lowest quality	2	3	4	5	6	7 Highest quality
---------------------	---	---	---	---	---	----------------------

*Comments*

### Item 3: Methods

*This item addresses the use of systematic methods and transparency in reporting; the use of the best available and up-to-date evidence; the consideration of effectiveness and cost-effectiveness of the potential options; and the weighting of benefits and harms in the guidance document.*

#### Criteria:

Item content includes the following criteria:

- **Systematic** and **transparent** methods were used to identify and review the evidence (e.g., integrated review, scoping review, review of the grey literature, systematic review).
- The **best available** and **most contextually relevant** evidence was considered.
- The evidence base is **current**.
- Evidence of **effectiveness** of the potential options is clearly described, including descriptions of the contexts in which the options were tested.
- Evidence of **cost** and **cost-effectiveness** of the potential options is described.
- The weighting of the **benefits** and **harms** of the potential options is described.
- There is a **link** between the recommendations and evidence.
- The **rationale** behind the recommendations is clear.
- **Systematic** and **transparent** methods were used to agree upon the final recommendations (e.g., informal or formal consensus, Delphi method, nominal group methods).

#### Evaluation:

Rate the overall quality of this item.

1 Lowest quality	2	3	4	5	6	7 Highest quality
---------------------	---	---	---	---	---	----------------------

*Comments*

## Item 4: Recommendations

*This item addresses the outcomes orientation and comprehensiveness of the guidance; the ethical and equity considerations drawn upon in its development; the details for its operationalization; the sociocultural and political alignment of the guidance; and the updating plan.*

### Criteria:

Item content includes the following criteria:

- The anticipated **outcomes** of implementing the recommendations are clearly described (including indicators, performance thresholds or targets, and standards to measure them).
- The recommendations are **comprehensive** and provide direction to all relevant health system levels (e.g., national, provincial/state), subsystems (e.g., cancer, mental health) and sectors (e.g., primary care, public health).
- The **ethical** principles used to develop the recommendations are described.
- The recommendations promote **equity** among the target population (e.g., in terms of age, sex, gender, culture, religion, race, sexual orientation).
- The recommendations' acceptability to, and **alignment** with, **sociocultural and political interests** were considered.
- The recommendations are easily **identifiable, clear, and succinct**.
- The recommendations are **actionable** and are sufficiently detailed to be **operationalized**.
- A **plan for updating** the recommendations is described.

### Evaluation:

Rate the overall quality of this item.

1 Lowest quality	2	3	4	5	6	7 Highest quality
---------------------	---	---	---	---	---	----------------------

*Comments*



## Item 5: Implementability

*This item addresses the barriers and enablers to implementing the recommendations; the cost and resource considerations in implementing the recommendations; the affordability of implementation and anticipated sustainability of outcomes; the flexibility and transferability of the guidance; and the strategies for disseminating the guidance, monitoring its implementation and evaluating its impact.*

### Criteria:

Item content includes the following criteria:

- **Barriers** and **enablers** to the implementation of the recommendations are described, including factors that are internal (e.g., resources, incentives, administrative structure) and external (e.g., legal system, social system, state of the economy, corruption, beliefs) to the health system. A plan to mitigate barriers and optimize enablers is included.
- **Cost** and **resource** considerations for the recommended actions are described (e.g., money, time, infrastructure, equipment, administrative capacity, supplies, staffing, and training).
- The stakeholders' **acceptability** of the recommendations is described.
- The **affordability** of the recommendations, in the context where implementation will take place, is described.
- The anticipated **sustainability** and requirements to maintain long-term outcomes is described.
- The recommendations are **flexible** and there is a description of how they can be adapted or tailored for specific contexts in which they will be implemented.
- A description of the degree to which the recommendations are **transferable** to other similar or different contexts is provided.
- Strategies for **disseminating** the health systems guidance are described.
- Strategies for **assessing** the implementation process and the impact of the recommendations are described.

### Evaluation:

Rate the overall quality of this item.

1 Lowest quality	2	3	4	5	6	7 Highest quality
---------------------	---	---	---	---	---	----------------------

Comments

## Overall Assessment

*The overall assessment requires the AGREE-HS user to make a judgment about the overall quality of the health systems guidance, taking into account the five AGREE-HS items.*

1. I would recommend this health systems guidance for use **in the appropriate context**.

- Yes
- Yes, with modifications
- No

*Comments*

2. I would recommend this health systems guidance for use **in my context** (optional).

- Yes
- Yes, with modifications
- No

*Comments*