

THE SEMMELWEIS FRAMEWORK: STRUCTURING EVIDENCE FOUNDATIONS AND EXPERT OPINION IN CLINICAL PRACTICE GUIDELINES

PhD thesis

Blin Nagavci

Semmelweis University, Doctoral College
Károly Rácz Conservative Medicine Division



Supervisor: Botond Lakatos, MD, PhD

Official reviewers: Gergely Peskó, MD, PhD

István Klemencsics, MD, PhD

Head of the Complex Examination Committee: Tamás Masszi, MD, PhD, DSc

Members of the Complex Examination Committee: Gergely Peskó, MD, PhD

Marienn Réti, MD, PhD

Budapest

2026

Introduction

Clinical practice guidelines are systematically developed statements containing recommendations intended to support informed healthcare decisions. They are normally informed by systematic reviews and formal appraisal of the best available evidence. In many areas, however, the available evidence is scarce. Guideline panels may then rely on expert opinion to interpret incomplete evidence or to provide guidance despite an evidence gap.

Expert opinion has a recurring practical role in guideline development. Previous research has shown that approximately one quarter of guidelines across medical fields contain recommendations based on expert opinion. In infectious diseases, almost half of strong recommendations have been reported to rely on low-quality evidence or expert opinion. Nevertheless, the concept is used inconsistently and is often insufficiently justified.

Two related methodological problems arise. First, there is no widely accepted definition of expert opinion. It may refer to accumulated clinical experience, interpretation of scarce evidence, or a consensus process. These meanings are not equivalent. Second, evidence-grading systems treat expert opinion differently. Traditional hierarchies often place it in their lowest evidentiary category, whereas GRADE regards expert opinion primarily as the interpretation of evidence through experts' experience and knowledge rather than as a separate type of evidence.

This heterogeneity creates practical challenges. Guideline developers may be uncertain when expert opinion is appropriate, how it should be elicited and documented, and how it should be distinguished from empirical evidence. Clinicians may find it difficult to determine what actually supports a

recommendation and how cautiously it should be interpreted. A transparent approach therefore requires more than a generic “expert opinion” label: it must clarify the evidentiary foundation of each recommendation and introduce safeguards for situations in which direct evidence is limited or unavailable.

The research presented in this thesis addressed these problems through three linked work packages.

Objectives

The overall objective was to develop a new method for utilising expert opinion in clinical practice guidelines. The project comprised three work packages, each resulting in a separate publication.

WP1 - Mapping current practice:

To construct an analytical map of how expert opinion is used in international infectious diseases guidelines, including its prevalence, terminology, application, and association with relevant methodological characteristics.

WP2 - Defining expert opinion:

To systematically review the methodological manuals of national and international guideline-developing organisations; examine the rationales and circumstances in which expert opinion is used; identify its foundations; and develop an evidence-informed definition.

WP3 - Developing the Semmelweis Framework:

To develop a new method for transparently utilising expert opinion in clinical guidelines with methodological safeguards, supporting medical decision making when evidence is scarce.

Methods

WP1: Meta-epidemiological study

Systematic searches were conducted in PubMed and Trip Medical Database in May 2023 to identify recent infectious diseases clinical guidelines. Eligible documents were clinical guidelines issued by international organisations involving at least two countries, based on primary research, written in English, and published from 1 January 2018 onwards. Records were screened in Rayyan by title and abstract and then in full text, and the selection process was documented in a PRISMA flow diagram.

A piloted extraction form captured bibliographic and organisational characteristics; methods for literature searching, risk-of-bias assessment, evidence appraisal, and study selection; and information on expert-opinion terminology, use, and the number of recommendations. Where reporting was incomplete, the organisation's website and methodological manual were consulted. Frequencies and narrative synthesis were used to describe expert-opinion use. Differences between guideline groups were assessed with chi-square and Mann–Whitney U tests, with alpha set at 0.05.

This work package corresponds to thesis-related Publication 1.

WP2: Methodological study of guideline manuals

A four-step systematic process was used. First, PubMed was searched for clinical guidelines published from 2019 to 2024 across medical fields. Second, guideline-developing and endorsing organisations were identified from the included guidelines and supplemented through previous work and hand searches. Third, official websites and a predefined Google Advanced Search were used to identify English-language methodological manuals published from

2010 onwards. Two reviewers independently assessed the retrieved sources, with disagreements resolved by discussion or a third reviewer. Fourth, information was extracted into piloted forms and verified by a second reviewer. Findings were summarised using descriptive measures and narrative synthesis. The definition of expert opinion was developed inductively, using a bottom-up analysis of the descriptions, rationales, and evidentiary foundations reported in the included manuals.

WP3: Development of the Semmelweis Framework

An international multidisciplinary team of researchers and guideline methodologists developed the Framework. Findings from WP1 defined the practical problems, while WP2 supplied the conceptual foundations. The categories identified in WP2 were used to construct an initial four-class system covering direct comparative evidence, indirect comparative evidence, non-comparative group-level evidence, and expert judgement without relevant group-level evidence.

The preliminary Framework was refined through qualitative application to 30 recommendations selected from 15 recent guidelines issued by different organisations. Each recommendation was independently assessed by three reviewers. Reviewers examined the original rationale and supporting study designs, assigned a class, reformulated the recommendation using the proposed wording and notation, and documented uncertainties or practical difficulties. Nine calibration meetings were held, and the flowchart, class descriptions, terminology, and implementation guidance were revised iteratively. Formal agreement statistics were not calculated because the instrument changed during development.

A Preconditions Checklist was developed in parallel. Problems identified during recommendation assessment were used to refine the checklist until no major conceptual omissions remained.

Results

WP1: Utilisation of expert opinion in infectious diseases guidelines

The searches identified 2,912 records, of which 178 were assessed in full text. Sixty-six guidelines containing 2,296 recommendations met the eligibility criteria. Of these, 32 guidelines allowed or used expert opinion, and contained 1,465 recommendations. 448 (30%) of recommendations were classified as based on expert opinion or supported by an evidence category that included expert opinion. Sixteen guidelines placed expert opinion in the same category as low- or very-low-quality evidence, whereas 14 separated it from other evidence or used systems in which it was not incorporated into the evidence hierarchy. Two guidelines did not use a grading system.

Guidelines allowing expert opinion issued more recommendations than those not allowing it (mean 48.82 versus 19.13; $p<0.001$). They were less likely to present a screening flow diagram (25% versus 65%; $p=0.002$) or report risk-of-bias assessment (19% versus 78%; $p<0.001$). Seven terms were identified, including “expert opinion”, “expert judgement”, “consensus of expert opinion”, “consensus recommendation”, “good practice principle”, and “in our practice statement”. These findings demonstrated substantial variability in terminology, classification, and methodological transparency.

WP2: Definition and foundations of expert opinion

After deduplication, 473 guideline-developing organisations were identified and searched for methodological manuals. Ninety-eight manuals were included: 61 handbooks, 29 journal articles, and 8 websites. Expert opinion was mentioned in 65 manuals, not mentioned in 30, and described too unclearly to determine its use in three.

Two principal rationales were identified. Forty-seven of the 65 organisations mentioning expert opinion used it to fill evidence gaps, allowing guidance when reliable direct data were unavailable. Five organisations used it primarily to interpret or contextualise existing evidence. In 13 manuals, the intended function could not be determined because insufficient detail was provided.

Five recurrent foundations of expert opinion emerged: clinical experience and professional judgement; indirect evidence; low- or very-low-quality evidence, including case reports and case series; mechanism-based reasoning or theoretical rationale; and systematically elicited expert evidence or expert witnesses. Most organisations relied on more than one foundation.

Based on the findings of this WP, a new comprehensive definition of expert opinion was synthesised:

“The concept of expert opinion in clinical guidelines refers to the synthesis of guidance derived from clinical experience, expert judgement, indirect evidence, very low-quality evidence or mechanism-based reasoning, aimed at supporting clinical decision-making in the absence of evidence or when inconclusive evidence requires contextualisation and interpretation.”

WP3: The Semmelweis Framework

The main output of WP3 was the Semmelweis Framework for structuring evidence foundations and expert opinion in clinical practice guidelines. The Framework classifies individual clinical recommendations according to the evidentiary foundation that supports them. It is intended to complement, rather than replace, established guideline methods such as GRADE. While GRADE assesses certainty of evidence and recommendation strength, the Semmelweis Framework makes explicit whether a recommendation is mainly based on direct evidence, indirect evidence, non-comparative evidence, or expert judgement. The Framework defines four recommendation classes.

1. Class I includes recommendations based on comparative evidence that directly addresses the guideline question. These recommendations may be strong or conditional and use the verbs “recommend” or “suggest.” The notation should specify that the recommendation is based on direct evidence.
2. Class II includes recommendations based mainly on comparative evidence that differs from the guideline question in one or more important PICO elements, such as the population, intervention, comparator, outcome, or setting. These recommendations are usually conditional and should be labelled as based on indirect evidence. Expert judgement has a greater role than in Class I because the panel must decide whether extrapolation to the target context is justified.
3. Class III includes recommendations based mainly on non-comparative group-level evidence, such as uncontrolled cohorts, single-arm trials, uncontrolled before-and-after studies, or non-comparative registries. These statements are generally expressed as considerations, using “consider” or “consider avoiding,” and must be labelled as based on non-comparative evidence. Conditional recommendations may be issued only in exceptional situations, while strong recommendations are not permitted.
4. Class IV includes recommendations based primarily on clinical experience, mechanism-based reasoning, collective panel judgement,

case reports, or case series, in the absence of relevant group-level evidence. These statements are also expressed as considerations and must be labelled “based on expert judgement.” Strong and conditional recommendations are not permitted in this class.

The Framework is applied after the conventional evidence-synthesis process has been completed, including systematic searching, study selection, risk-of-bias assessment, evidence synthesis, and certainty-of-evidence assessment where applicable. For each recommendation, the panel identifies the evidentiary foundation that had the greatest influence on the final judgement. This dominant foundation determines the recommendation class (Figure 1).

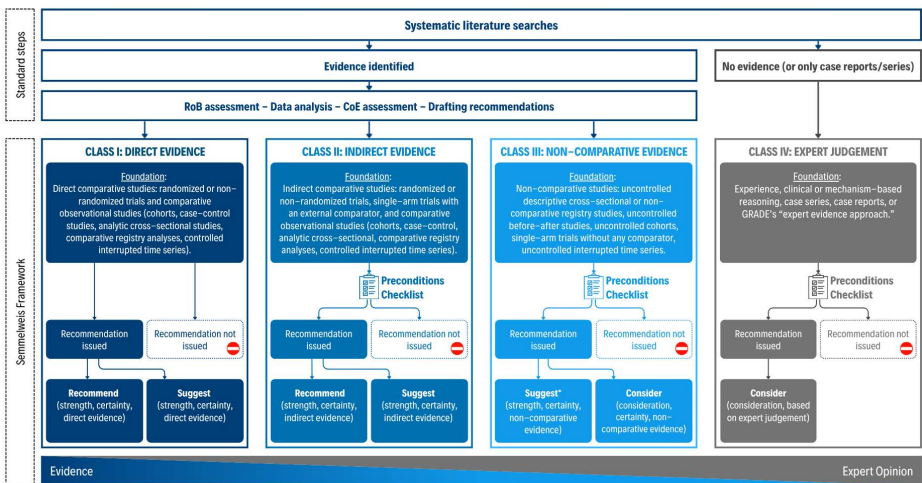


Figure 1. Algorithm for applying the Semmelweis Framework.

Reproduced unchanged from Nagavci B, Andrews MR, Vikse J, Stuenkel P, Stamm T, Lakatos B. The Semmelweis Framework: structuring evidence foundations and expert opinion in clinical practice guidelines. BMC Med Res Methodol. 2026. Used under the Creative Commons Attribution (4.0) license.

RoB: Risk of Bias, CoE: Certainty of Evidence.

The classification is applied separately to each recommendation rather than to the clinical question. Recommendations arising from the same PICO question may therefore receive different classes when they rely on different foundations.

Before issuing recommendations in Classes II, III, or IV, the panel must complete the Preconditions Checklist. The Checklist contains 11 questions organised into four domains: Evidence, Necessity, Patient and Research Impact, and Panel Integrity.

The Checklist assesses whether systematic searches were conducted, why sufficient direct evidence is unavailable, whether issuing guidance is clinically necessary, and whether potential benefits, harms, uncertainties, and assumptions were considered. It also addresses patient values and preferences, subgroup effects, evidence gaps, research priorities, panel expertise, balanced participation, and management of conflicts of interest.

All preconditions should be fulfilled before issuing a Class II, III, or IV recommendation. When they are not met, the panel should avoid issuing a recommendation, explain why no recommendation could be made, and identify future research priorities where appropriate.

The Framework is supported by four implementation resources: the classification algorithm, the recommendation-classification table, the Preconditions Checklist, and a set of frequently asked questions. Together, these resources provide guideline panels with a structured method for identifying the foundation of each recommendation, selecting appropriate wording, and communicating the role of expert judgement transparently.

Conclusions

This thesis introduced a new systematically developed definition of expert opinion. Building on this definition, the Semmelweis Framework provides a structured approach for integrating empirical evidence and expert opinion into clinical practice guidelines. It classifies recommendations according to their evidentiary foundations and distributes the roles traditionally grouped under “expert opinion” across clearly defined recommendation classes. Supported by a dedicated Preconditions Checklist, the Framework avoids using expert opinion as a generic recommendation label and instead promotes transparent and actionable guidance. In this way, the Framework offers guideline developers a practical tool for supporting clinical decision-making when evidence is limited, indirect, or absent.

Bibliography of the candidate's publications

Publications related to the thesis

1. **Nagavci B**, Schwingshackl L, Martin-Loeches I, Lakatos B. Utilization of expert opinion in infectious diseases clinical guidelines- A meta-epidemiological study. *PLoS One*. 2024;19(6):e0306098.
2. **Nagavci B**, Gáspár Z, Lakatos B. Defining expert opinion in clinical guidelines: insights from 98 scientific societies - a methodological study. *BMC Med Res Methodol*. 2025;25(1):87.
3. **Nagavci B**, Andrews MR, Vikse J, Studenic P, Stamm T, Lakatos B. The Semmelweis Framework: structuring evidence foundations and expert opinion in clinical practice guidelines. *BMC Med Res Methodol*. 2026.

Publications not related to the thesis:

4. Gáspár Z, **Nagavci B**, Szabó BG, Lakatos B. Gut Microbiome Alteration in HIV/AIDS and the Role of Antiretroviral Therapy—A Scoping Review. *Microorganisms*. 2024;12(11):2221.
5. Bartoletti M, Azap O, Barac A, Ben Selma M, Ergonul O, Gkrania-Klotsas E, Grossi PA, Krause R, **Nagavci B**, Paño-Pardo JR, Pierrotti LC, Power N, Rodríguez-Baño J, Sibani M, Slavin MA, Szabo BG, Tazza B, Yung Tsang NN, Tsiodras S, Zollner-Schwetz I, Chemaly RF. Management of COVID-19 in immunocompromised patients: an European Society of Clinical Microbiology and Infectious Diseases consensus document. *Clin Microbiol Infect*. 2025;31(10):1655-66.
6. Righi E, Yahav D, Nasim A, Akova M, Barac A, Scudeller L, **Nagavci B**, Armellini M, Zanchi C, Shirin N, Tunali V, Paño-Pardo JR, Rodríguez-Baño J, Jamrozik E. Incorporating ethics into infectious disease clinical practice guidelines. *Clin Microbiol Infect*. 2025;31(4):560-7.
7. Yahav D, Nasim A, Shirin N, Armellini M, Zanchi C, Tunali V, Mirandola M, Barac A, Scudeller L, **Nagavci B**, Paño-Pardo JR, Rodríguez-Baño J, Jamrozik E, Akova M, Righi E. Incorporating

- Ethics in Clinical Guidelines in Infectious Diseases: A Scoping Review. JAMA Netw Open. 2025;8(7):e2519826.
8. Albers B, **Nagavci B**, Joean O, Tschudin Sutter S, Tacconelli E, Paul M, Langendam M, Kostyanov T, Molina J, Stahl JP, Pano Pardo JR, Gkrania-Klotsas E, Hassoun-Kheir N, Clack L, Scudeller L, Zervos M. Guiding principles for implementing ESCMID clinical practice guidelines. Clin Microbiol Infect. 2026;32(7):1107-14.
 9. Joean O, Sermet K, Ashkenazi-Hoffnung L, Kiyamaz YC, Blumenthal K, Bonazzetti C, Chiriac AM, Gomez-Zorrilla S, Karakike E, Khatamzas E, Powell N, Trubiano JA, Wijnakker R, Sandoe J, **Nagavci B**, De Boer MGJ. ESCMID clinical guidelines on the evaluation and management of a reported antibiotic allergy. Clin Microbiol Infect. 2026;32(5):767-87.
 10. **Nagavci B**, Joean O, Tschudin-Sutter S, Tacconelli E, Paul M, Langendam M, Kostyanov T, Molina J, Stahl JP, Zervos M, Bodilsen J, Paño-Pardo JR, Scudeller L. The ESCMID framework for reaching consensus in clinical guidelines. Clin Microbiol Infect. 2026.
 11. Tunalı V, **Nagavci B**, Carmena D, Robertson LJ, Doğruman Al F, Gentekaki E, Tsaousis AD. Relationship between Blastocystis infection and clinical outcomes: A scoping review protocol. Open Res Eur. 2025;5:15.

Cumulative impact factor of publications related to the thesis: 10.0

Cumulative impact factor of publications unrelated to the thesis: 59.4

Total cumulative impact factor of the candidate's publications: 69.4