Systematic Review Protocol

Overview – *basic study information*	
Title of the review	Harmonizing Healthcare: The Art and Statistics of Consensus Building
Project title *(If different from review title)*	N/A
First reviewer	Joshua J. Cook, M.S. DS, M.S. CRM, ACRPM-PM, CCRC (jcook3@uwf.edu)
Team of reviewers	Andrew Jimenez, M1, B.S., (amj156@georgetown.edu) Thomas Jonte, M.S., (tjonte@uwf.edu) Achraf Cohen, Ph.D. (acohen@uwf.edu)
Supervisor/project principal investigator	Achraf Cohen, Ph.D.
Organization(s)	University of West Florida, Georgetown University School of Medicine
Review method *(PRISMA, PRISMA-P, Cochrane)*	PRISMA (PRISMA statement (prisma-statement.org))
Funding mechanism	None

Delegation – *management plan*	
Protocol Development	Cook, Jimenez, Jonte, Cohen
Literature Searching	Cook, Jimenez, Jonte
Data Extraction	Cook, Jimenez
Synthesis/Analysis	Cook, Jimenez, Cohen
Writing Up	Cook, Jimenez, Jonte, Cohen

Table of Contents

1. Background to Review

1.1 Important characteristics

What are the important population and/or disease characteristics (diagnostic criteria, epidemiology, etiology, prognosis)?

In the complex realm of medicine, it's essential to acknowledge that even seasoned practitioners may require guidance, especially when faced with uncertain or risky medical scenarios where traditional research methodology is deemed not feasible or unethical. Consensus statements are crucial for identifying expert perspectives on the appropriate actions in such situations, offering a collective direction where singular expertise might be insufficient and where evidence-based methods might not be possible.[1,2] However, there are several types of consensus studies, each with their own design schema and statistical guidelines that require deep understanding to ensure the validity and reliability of study outcomes. These include the traditional Delphi method, which relies solely on individual perspectives, the Nominal Group Technique, which uses small group discussions, and the RAND/UCLA Appropriateness Method (RAM), which integrates the benefits of both of these methodologies to synthesize expert opinions.[3] The lack of standardization among studies in each category, especially in RAM studies, has made it more challenging for project managers, protocol writers, and statisticians to identify the appropriate criteria needed for implementation into future consensus study designs and statistical analysis plans.

1.2 Relevance

Does the review topic have important implications for health (individual and/or public), as well as health care, policy and research?

Consensus methodologies continue to be extremely valuable tools for healthcare professionals despite their clear distinction from evidence-based research studies.[1,2] Identifying the commonalities and differences among expert medical opinions concerning complex patient presentations can provide guidance and confidence to physicians facing similar situations, especially when existing data fails to provide clear evidence on a particular topic.[4] Since the RAND/UCLA Appropriateness Method (RAM) encompasses both individual evaluations and group discussions about a variety of interrelated patient scenarios, each expert has the ability to consider other expert viewpoints and considerations about the intricacies of an issue before making their own assessment, which enhances the information behind this guidance when compared to other consensus models.[5] Importantly, these benefits have become increasingly achievable with the rise of online video conference modalities, which have alleviated many of the logistical and geographic barriers that have traditionally limited expert assembly.[6] These innovations allow a wider array of medical experts with different professional experiences to provide insight, improving the value and generalizability of the final determinations. Even though there are clear advantages of the RAM, there is an evident lack of standardization among study designs, specifically regarding expert numbers, question numbers, and consensus threshold cutoffs for final recommendations.[6] These inconsistencies provide unnecessary barriers for project managers, protocol writers, and statisticians that seek to capture the benefits of the RAND/UCLA methodology in the development of healthcare research and policy.

1.3 Rationale

Does the evidence (including existing systematic reviews) fail to answer the review question, and why?

There are limited systematic reviews as of May 2024 on the general topic of consensus studies.[1,2] There are none specific to RAM.

1.4 Justification & Aim

Is the need for the review justified in the light of the potential health implications and current limitations of the evidence base?

This review will provide a succinct guide to the intricacies involved in crafting RAND/UCLA consensus statements, addressing the definition of consensus, the optimal number of expert participants, and the balance between agreement and discussion rounds. We will highlight the role of iterative feedback and the challenge of expert retention, supported by a systematic review of the existing literature and R simulations that assess the parameters that influence consensus achievement. Our aim is to serve as practical and statistical advisories for project managers, protocol writers, and statisticians, emphasizing that the process of reaching a consensus is not only iterative and collaborative but also integral to advancing medical practice and knowledge.

1.5 Specification

What are the PICO (population, intervention, comparison, outcome) components of the review question / objective?

See 3.1.

2. Specific Objectives

1. To evaluate the existing available literature regarding the design of RAND/UCLA Appropriateness Method (RAM) consensus studies, with a specific focus on the statistical measures implemented to determine a consensus and ensure power, replication, and validity.
2. To collect study design and outcome data from the included studies to build a baseline probability distribution for use in simulation studies that will be focused on evaluating the influence of various study design characteristics on reaching a consensus.

3. Inclusion/Exclusion Criteria

3.1 Inclusion Criteria

Population, participants and conditions of interest	Physicians or medical experts
Interventions or exposures	Any medical, clinical, or public health studies

Comparisons or control groups	N/A - but check to see if consensus is being compared to placebo or standard-of-care
Outcomes of interest	Number of included experts, expert background, number of rounds, number of discussions, consensus threshold, dropping of questions between survey, any statistical methodology
Setting	Any
Study designs	RAND UCLA Appropriateness Method (RAM) consensus studies (The RAND/UCLA Appropriateness Method User's Manual | RAND)

3.2 Exclusion Criteria

Any specific populations excluded, date range, language, region, full text availability, etc.

- Studies beyond the date range: May 2019 – May 2024 or
- Studies outside the United States or
- Non-English articles unless full translation is available or
- No access to free full text nor institutional access or
- "Modified" versions of RAND/UCLA unless altered methodology is clearly explained

4. Search Methods

Electronic databases *(Highly suggested to include multiple, with data ranges searched for each)*	PubMed Cochrane Google Scholar MedLine JBI Web of Science
Key search terms, and method of identification of search terms *(must be sensitive and specific, can narrow down)*	litsearchr - an R package to facilitate quasi-automatic search strategy development for systematic reviews (elizagrames.github.io)
Other methods used for identifying relevant research *(i.e., experts, grey literature via CADTH, etc.)*	WHO's International Clinical Trials Registry US National Technical Information Service Pew Google Scholar
Journals hand searched *(include journals, date of search, and rationale for selection)*	N/A
Use of snowballing methods *(i.e., capture of citing papers)*	N/A

5. Methods of Review

Details of methods ***(At least 3 reviewers suggested, dealing with agreements and disagreements)***	Three main reviewers (Cook, Jimenez, Cohen) and a third (Jonte) to resolve any disagreements.
Quality assessment ***(tools or checklists used with references or URLs – CASP, AACODS)***	We anticipate that most references will come from peer-reviewed published articles. Thus, we will use the Critical Appraisal Skills Programme (CASP) tool to assess the quality of the systematic review articles (CASP Checklists - Critical Appraisal Skills Programme (casp-uk.net)).
Data extraction ***(information to be collected from each study, methods)***	Data regarding the design and outcomes of the RAM consensus studies will be collected, including number of included experts, expert background, number of rounds, number of discussions, consensus threshold, dropping of questions between surveys, whether a consensus was reached and on how many scenarios. This will be entered into Microsoft Excel. Reviewer number 1 will review first, followed by reviewer number 2 and 3, which will be done independently.
Narrative synthesis ***(what and how synthesis will be done)***	Narrative synthesis will be done alongside any meta-analysis and will be carried out using a framework which consists of four elements; 1. Developing a foundation of common design characteristics of RAM consensus studies. 2. Developing a preliminary synthesis of findings of included studies. 3. Exploring relationships within and between studies. 4. Assessing the robustness of the synthesis.
Meta-analysis ***(methods – Cochrane)***	Welcome! | Doing Meta-Analysis in R (bookdown.org)
Grading evidence ***(system used, such as GRADE)***	N/A
Bias mitigation plan ***(tools or checklists used with references or URLs – RoBiS/RoB 2)***	RoBiS will be used to mitigate bias (ROBIS: A new tool to assess risk of bias in systematic reviews was developed - PubMed (nih.gov)).

6. Processes and Resources

Background/expertise	Study design, clinical research, data analysis, manuscript writing
Computing facilities	MacBook Pro 2023
Research databases	PubMed, Cochrane, Google Scholar
Bibliographic software	Zotero
Statistical software	R/RStudio with associated packages

7. Presentation of Results

Additional material *(summary tables, flowcharts, etc. to be included in the final manuscript or associated database/GitHub)*	Systematic review protocol template; simulation code via GitHub
Outputs from review *(target conferences, journals)*	Publication – Journal of Statistical Theory and Practice (CSP Invitation; Home | Journal of Statistical Theory and Practice (springer.com))

8. Timeline for Review (*estimated*)

Protocol development & registration through PROSPERO (PROSPERO (york.ac.uk))	26JUN2024
Literature searching and study selection	28JUN2024
Study quality appraisal	03JUL2024
Data extraction	04JUL2024
Synthesis	05JUL2024
Writing up	07JUL2024
Draft manuscript for internal/peer review	07JUL2024
Submission for publication	30JUL2024

References

1. Djulbegovic B, Guyatt G. Evidence vs Consensus in Clinical Practice Guidelines. JAMA. 2019;322(8):725-726. doi:10.1001/jama.2019.9751

2. Kwong JS, Chen H, Sun X. Development of Evidence-based Recommendations: Implications for Preparing Expert Consensus Statements. Chin Med J (Engl). 2016;129(24):2998-3000. doi:10.4103/0366-6999.195475

3. Arakawa N, Bader LR. Consensus development methods: Considerations for national and global frameworks and policy development. Res Social Adm Pharm. 2022;18(1):2222-2229. doi:10.1016/j.sapharm.2021.06.024

4. Kea B, Sun BC. Consensus development for healthcare professionals. Intern Emerg Med. 2015;10(3):373-383. doi:10.1007/s11739-014-1156-6

5. Fitch K, Bernstein SJ, Aguilar MD, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND Corporation; 2001. https://www.rand.org/pubs/monograph_reports/MR1269.html. Also available in print form.

6. Sparks JB, Klamerus ML, Caverly TJ, et al. Planning and Reporting Effective Web-Based RAND/UCLA Appropriateness Method Panels: Literature Review and Preliminary Recommendations. J Med Internet Res. 2022;24(8):e33898. Published 2022 Aug 26. doi:10.2196/33898