ACCSCIENCE PUBLISHING

Journal of Clinical and Translational Research

MEDICAL HYPOTHESIS

A systematic review protocol of medical and clinical research landscapes and quality in Malaysia and Indonesia (REALQUAMI)

Boon-How Chew[1]*, Shaun Wen Huey Lee[2], Lim Poh Ying[3], Soo Huat Teoh[4], Aneesa Abdul Rashid[1], Navin Kumar Devaraj[1], Adibah Hanim Ismail[1], Abdul Hadi Abdul Manap[1], Fadzilah Mohamad[1], Aaron Fernandez[5], Hanifatiyah Ali[1], Puteri Shanaz Jahn Kassim[1], Nurainul Hana Shamsuddin[1], Noraina Muhamad Zakuan[6], Akiza Roswati Abdullah[7], and Indah S. Widyahening[8]

[1]Department of Family Medicine, Faculty of Medicine and Health Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[2]School of Pharmacy, Monash University Malaysia, Bandar Sunway, Selangor, Malaysia

[3]Department of Community Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[4]Advanced Medical and Dental Institute (AMDI), Universiti Sains Malaysia, Kepala Batas, Penang, Malaysia

[5]Department of Psychiatry, Faculty of Medicine and Health Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[6]Department of Biomedical Sciences, Faculty of Medicine and Health Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[7]Medical and Health Sciences Library, Faculty of Medicine and Health Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[8]Department of Community Medicine, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia

***Corresponding author:**
Boon-How Chew
(chewboonhow@upm.edu.my)

Citation: Chew B, Lee SWH, Ying LP, *et al*. A systematic review protocol of medical and clinical research landscapes and quality in Malaysia and Indonesia (REALQUAMI). *J Clin Transl Res*. 2025;11(2):94-105.
doi: 10.36922/jctr.24.00071

Received: October 22, 2024

1st Revised: December 11, 2024

2nd Revised: December 12, 2024

Accepted: March 11, 2025

Published online: April 4, 2025

Abstract

Background: The evolving landscape of clinical and biomedical research has raised concerns about waste and quality. Poorly conducted studies mislead clinical practice and compromise patient outcomes. Reliable data from past research are essential for research quality improvement. **Aim:** The aim of the study was to characterize and assess the quality of research in Malaysia and Indonesia. **Methods**: To establish the proposed systematic review protocol, we will search PubMed, Cochrane Library, CINAHL, and PsycINFO for studies published from 1962 to 2019, supplemented by MyMedR for Malaysian research. Two reviewers will independently screen studies, extract data, and assess quality. Phase 1 will descriptively report research characteristics, including researcher profiles and journal outlets. In Phase 2, a quality screening tool will be validated across three domains: relevance, methodological credibility, and result usefulness. Associations between research characteristics and quality will be analyzed through multivariable regression and longitudinal trends will be explored. **Results**: Findings from the proposed systematic review protocol will generate baseline data for national and international comparisons, guiding stakeholders, researchers, funders, and policymakers on research evolution and quality trends. Results may inform improvement initiatives and resource allocation for understudied areas. **Conclusion:** This review aims to establish a

comprehensive baseline of research outputs and the pattern of research quality in the participating countries and discipline. The findings may underscore the presence of a valid classification method to guide future research and enhance evidence-based practice in healthcare. **Relevance for patients**: By identifying research strengths and gaps, this proposed systematic review supports the development of robust study designs that generate reliable evidence, ultimately enhancing patient care and health outcomes.

Keywords: Systematic review; Clinical research; Biomedical research; Research characteristics; Research quality; Malaysia; Indonesia

1. Introduction

There is currently an increasing number of clinical and biomedical research and publications globally, especially those originating from Asia.[1] However, the increase in quantity does not correlate with the quality of research conducted. Instead, significant research waste has been reported due to irrelevancy,[2] poor research design,[3] inaccessible research data,[4] and incomplete reporting.[5,6] Moreover, John Ioannidis, one of the co-directors at the new Meta-Research Innovation Center at Stanford, admitted that "it was very easy to make errors" when discussing the challenges encountered throughout the research process, despite the noble intentions of the researchers.[7] The clinical and biomedical research landscape in Asia has been evolving throughout the past decades, beyond those reported from a few sources, with growth seen more in terms of quantity than quality.[1] Similarly, the quality of published research in a country such as Malaysia and Indonesia has not been examined over the past few decades. Comprehensive assessments and evidence are needed to inform existing researchers, research institutes, and stakeholders in these countries about the adequacy of current efforts or the need to improvise existing research practices.

An estimated 200 tools were previously available for evaluating research quality or biases in randomized and non-randomized studies.[8-10] A recent literature review up to April 2022 identified 417 appraisal tools for non-randomized studies of interventions.[11] These tools serve critical functions in ensuring that studies adhere to robust methodological standards, minimizing bias, and enhancing the reliability of results. Deeks *et al.*[8] evaluated non-randomized intervention studies, highlighting challenges like biases and confounding factors, and emphasized the need for tailored tools to address these complexities. Their work reviewed tools designed to standardize assessments and improve comparability, providing a framework for systematic quality evaluations. MacLehose *et al.*[9] compared effect sizes from randomized and non-randomized studies, identifying significant discrepancies due to biases in study designs, and underscored the importance of quality assessment tools for detecting and addressing these differences to ensure reliable comparisons and synthesis across diverse research designs. Zeng *et al.*[10] reviewed methodological quality assessment tools across preclinical, clinical, and systematic review domains, highlighting their strengths, gaps, and domain-specific challenges. Nevertheless, most tools available for assessing non-randomized studies are generally of poor methodological quality, making it consistently difficult or even impossible to assess the methodological quality and risk of bias across primary studies[12] Many different tools exist for different study designs, such as the Cochrane Risk of Bias tool for randomized trials,[13] the A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies 2 tool[14] for diagnostic test accuracy studies, the Assessment of Multiple Systematic Reviews[15] and Risk of Bias in Systematic Reviews tools[16] for systematic reviews, and the Risk of Bias In Non-Randomized Studies of Interventions[17] for non-randomized studies of the effects of interventions. In addition, there are a few web-based tools and checklists for different study designs: the National Institutes of Health Study Quality Assessment Tool for controlled intervention studies, systematic reviews and meta-analyses, observational cohort and cross-sectional studies, case–control, pre-post, and case series studies (https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools); the Critical Appraisal Skills Program checklists by Oxford-based Better Value Healthcare Ltd. (https://casp-uk.net/casp-tools-checklists/); a web application Critical Appraisal Tools (FLC 2.0) developed by OSTEBA Spain to guide the critical appraisal process (http://www.lecturacritica.com/es/acerca.php).

Among some of the more widely used and recommended tools are the Newcastle-Ottawa scale (NOS),[18] the Downs and Black instrument,[19] and the latter RTI item bank (RTI-IB).[20] The NOS, which has been used to illustrate issues in data extraction from primary non-randomized

studies, has only eight items and is simpler to apply.[18] However, the items may require customization to the review question of interest. The Downs and Black instrument[19] has been modified for use in a methodological systematic review.[20] The reviewers found that some of the 29 items were difficult to apply to case–control studies. In addition, the instrument requires considerable epidemiological expertise and is time-consuming and difficult to operate.[21-24] The median observed inter-rater agreement for the RTI-IB was 75% (25th percentile [p25] = 61%; p75 = 89%). The median first-order agreement coefficient statistic was 0.64 (p25 = 0.51; p75 = 0.86). Although the RTI-IB facilitates a more complete quality assessment than the NOS, it is also more burdensome. In addition, there are different meanings for epidemiological terminologies across different countries. For example, the term "selection bias" describes what others may call "applicability" or "generalizability." Consequently, applying these tools presents challenges, as they require significant epidemiological expertise, are time-consuming to use, and often yield inconsistent results due to low inter-rater reliability. Thus, comprehensive manuals are required to accompany these tools to provide instructions for standardized interpretation by different users. However, this may pose a real challenge to users, as few tools have such comprehensive manuals. To the best of our knowledge, no existing tool serves as an all-rounded tool for assessing all types of study designs,[10] nor is there a recommended tool suitable for quickly screening the quality of published research.

In light of this, our research group has assimilated quality indicators used in existing tools, drawing from a series of user guides to the medical literature by the Evidence-Based Medicine Working Group,[25,26] systematic reviews,[27,28] and principles of clinical epidemiology,[29] and consequently, developed a comprehensive assessment tool based on the findings. In future work, the group will systematically assess the characteristics in research publications by researchers in Malaysia and Indonesia and subsequently identify quality indicators of these research publications.

2. Research design

The overall research design consists of two analytical phases:

(i) Phase 1 (research landscapes): Descriptive reporting of the research demographics and characteristics in each country to date.
(ii) Phase 2 (research quality): Quality assessment of research based on the published reports in journals.

2.1. Inclusion criteria and search strategy

Before data analysis, all clinical and biomedical research conducted in Malaysia or Indonesia, from January 1962 to December 2019, will be identified from various databases: PubMed, Cochrane Library, CINAHL, and PsycINFO (Figure 1); including all published peer-reviewed health and biomedical research papers from each country (Malaysia or Indonesia), or authored by citizens of each country (Malaysian or Indonesian) with an affiliation to an institution in either country. Additional literature will be extracted from the MyMedR (Malaysian Medical Repository) (http://mymedr.afpm.org.my/) database, as it specifically compiles published papers in health and biomedical research conducted in Malaysia or by authors who have a Malaysian affiliation. MyMedR also draws from MyJurnal, an online system used by the Malaysia Citation Centre of the Ministry of Higher Education Malaysia to collect and index all Malaysian journals. Grey literature will not be included, and there will be no exclusion based on language. Search results will be compiled into Endnote reference management software, where duplicates will be removed. If necessary, authors and institutions will be contacted. A medical librarian and a scientific officer at the Faculty of Medicine and Health Sciences, Universiti Putra Malaysia will assist with these tasks. The screening process for publications from both countries will be completed by two separate teams, each based in Malaysia and Indonesia, respectively.

2.2. Study selection and data extraction

The authors of the present article (i.e., physicians and medical professionals) will serve as reviewers of the identified articles and will be trained on the review protocol before commencing the paper screening. The articles will be screened by title and abstract. The full text of eligible articles will be retrieved and distributed to pairs of reviewers for screening; independently extracted information will be input into a standard data extraction template (Tables A1-4). This template has been pilot-tested on 10 articles among all the reviewers for clarity, and modification of the template was done accordingly. The final piloted template is available in Tables A1 and 3. Any discrepancy during article selection and data extraction will be solved by consensus between three or more reviewers. To ensure the inclusion of only high-quality data, B.H.C. will reassess 10 – 20% of the included articles.

In the event of duplicate publications or multiple reports of a research study, we will use the most complete data set aggregated across all known publications. Duplicate publications are defined as two or more published articles that report on the same research question.

2.3. Research landscapes

Phase 1 of the project will describe the following characteristics of the reported research (Table A1 for more details):

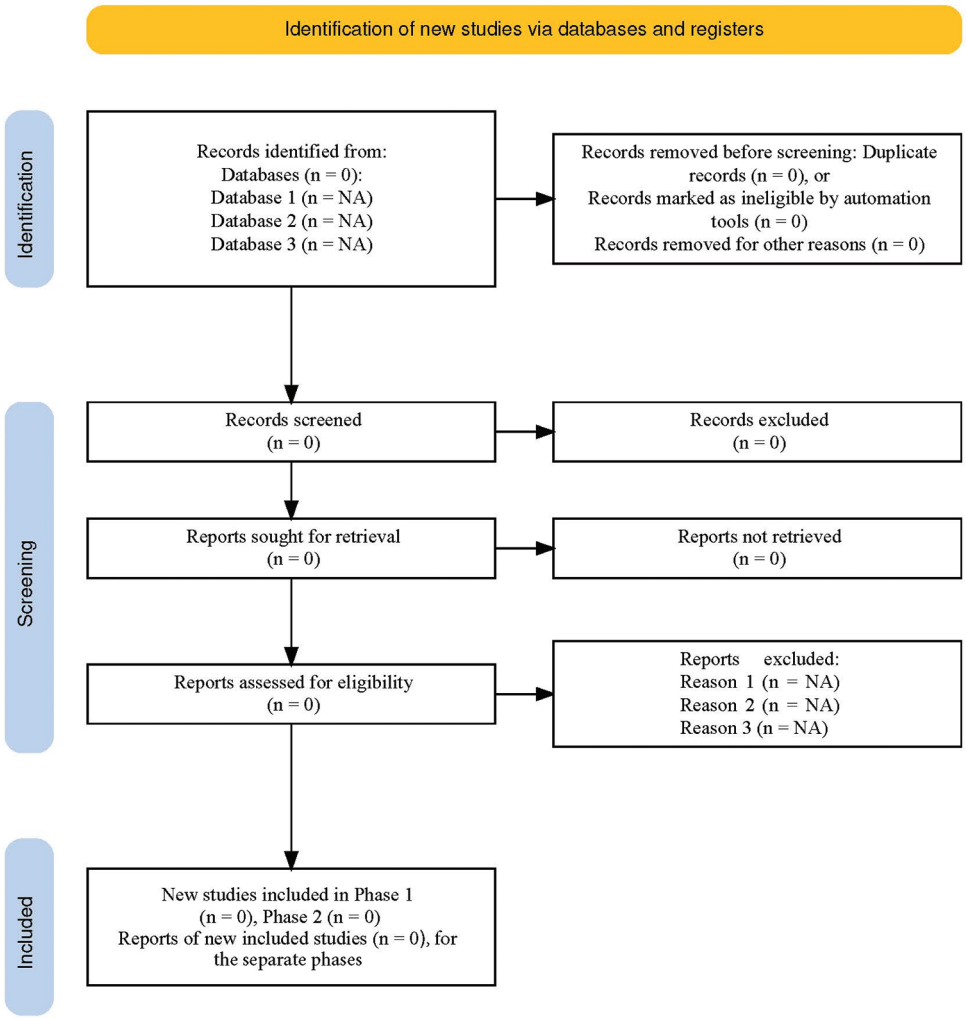

Figure 1. Outline of the flow of information through the two phases of the systematic review. Adapted from Haddaway *et al.*, (2022).[30]

(i) Institution and qualification of the corresponding author/s;
(ii) Number of authors, institutions, and expertise;
(iii) Numbers of international collaborating authors and institutions;
(iv) Numbers of study sites;
(v) Journal type: Local, regional, or international; open access or traditional subscription-based; general or discipline-specific;
(vi) Setting: Healthcare facility (hospital, clinic, etc.) or community;
(vii) Type of study: Audit versus research; secondary (reviews) or primary (diagnostic, prognostic, etiologic, or interventional); clinical or non-clinical (laboratory, public health, health service, etc.);
(viii) Data-collection design;
(ix) Date of study conducted, completed, and published;
(x) Health conditions studied, and/or organ systems involved;
(xi) Drugs, devices/tools, surgical, psychological, or health services. (Table A5)

2.4. Research quality

In Phase 2, the research quality will be assessed based on various criteria across three domains: relevance, credibility, and usefulness (Table 1). All reviewers will learn the principles of clinical epidemiology through a workshop and reach a consensual understanding of the terms used to represent research quality in this project. During the workshop, we will implement a training session for all reviewers, in which all reviewers will read and score the same articles. Each item will be scored 1 if present, and 0 when absent. This will be followed by a discussion on any similarities or differences in the quality assessment and scores to ensure uniformity in the understanding of the quality domains when applied to the actual screening process. We will also determine the

Table 1. Research quality domains and items to be used in the screening tool (total score: 0 – 10)

Relevance (3 items; subtotal score: 0 – 3)	Credibility (4 items; subtotal score: 0 – 4)	Usefulness (3 items; subtotal score: 0 – 3)
- Scientific relevance: Indicate this with an acceptable literature review or citing systematic reviews[a] - Societal relevance: Research area or involvement of end-users (e.g., patients) - Research team/experts: The research is led by experts in the relevant field or conducted with relevant experts	- Data-collection design: Appropriate for the research question[b]; experimental vs. non-experimental; time feature of variables considered - External validity: Representative of or generalizability to an important and relevant population; comparability between groups in randomized control trials - Internal validity: Validated instrument, measurement process, and performed by trained or blinded assessors - Precision: Appropriate sample size estimation and achievement	- Important outcomes used and reported[c] - Meaningful estimates: Practical numerical results, taking into consideration the response rate, missing data, proper statistical test, and analysis[d] - Conclusion: Accounts for study limitations[e]

Notes: [a]Set the right research priorities, clear research question/hypothesis; [b]Ethical conduct and patient safety/rights/priorities included; [c]Outcomes that truly matter to patients; [d]The study provides useful data for the intended end-users, unusual or unexpected analysis is explained and justified; [e]No over-claimed or misleading conclusion.

inter-rater reliability agreement using Cohen's kappa κ and intra-class correlation (ICC). The kappa κ is a measure of agreement between different observers beyond chance agreement.[31] The κ statistic will be computed separately for each domain's item (0 or 1). The ICC will be used to assess the three domains' subtotal (3, 4, and 3) and the total score (Table 1).

The kappa result will be interpreted as follows: Values ≤0 indicate no agreement; 0.01 – 0.20 indicate slight agreement; 0.21 – 0.40 indicate fair agreement; 0.4 – 0.60 indicate moderate agreement; 0.61 – 0.80 indicate substantial agreement; and 0.81 – 1.00 indicate almost perfect agreement.[32,33] For the ICC, values <0.40 indicate poor correlation; 0.40 – 0.59 indicate fair correlation; 0.60 – 0.74 indicate good correlation; and 0.75 – 1.0 indicate excellent correlation.[34,35] We specify that an a priori level of κ > 0.60 and ICC > 0.75 must be achieved before Phase 2 of the study begins. Retraining and reassessment of the reviewers on different articles will be conducted until the inter-rater agreement reaches the desirable levels. The expected lower bound of a 95% confidence limit for κ is no <0.60, with an assumed marginal prevalence of zero score of 30%. Using alpha and beta error rates of 0.05 and 0.2, respectively, a pair of reviewers will rate 20 papers each,[34,35] with five pairs of reviewers and 100 samples for the subtotal and total ICC estimation.[33]

2.5. Research quality domains for screening

2.5.1. Relevance

The relevance of research will be assessed from three perspectives: scientific relevance, the composition of the research team, and societal relevance. Research is considered scientifically relevant if it addresses a true and real scientific problem and provides the needed knowledge to understand an existing phenomenon. Scientific relevance also denotes that the research sets out on a justified scientific foundation and is informed by existing evidence. Thus, scientifically relevant research is usually globally relevant due to its highly generalizable topics and subjects.

Societal relevance refers to research that addresses a true and real societal problem. This relevancy may exist in smaller (e.g., a particular condition or disease in a unique population) or wider (e.g., global) populations. These two domains of scientific and societal relevance relate to having novelty in the research.

The last domain in the relevance category refers to the research team; that is, investigators and experts of relevant professional qualifications. This may include patients and the public in certain research areas when the opinions of end-users are considered important, such as in interventions or the experiences of patients or family members.

2.5.2. Credibility

This category is assessed after it is judged that the research is relevant. Four essential features are considered the minimum requirements for a research study to be credible and for its results to inform or contribute to practice change; that is, data collection design, precision, external validity, and internal validity.

The design of the data collection process has to align with the research objective or question. The approach used in data collection depends on whether the research is causal or non-causal, as well as whether the research is experimental or non-experimental, to provide more accurate data. The time feature or characteristics of the variables involved in the research should be collected in their intended phases or stages, such as a risk factor in the asymptomatic phase, or symptoms or biomarkers in the latent period.

Sampling and sample types are also important in the credibility domain. The sample of the participants should consist of the appropriate group from the population for the research, as they represent the population to which the results could later be generalized. However, in causal or experimental research, comparability between groups in research takes precedence over representativeness because confounding or prognostic factors between groups result in valid outcomes related to the exposure.

Quantitative research is essentially about measurements, measuring tools, and the measurement process. The measurement of variables is to be performed using validated tools through a standardized process, and if necessary, by trained and blinded assessors. Any query or suspicion on the methods of measurement in research will result in internal non-validity.

Credible research provides an appropriate and rational sample size estimation based on the research question and its primary objective, as well as previous research within the same discipline. Adequate sample size is required for sufficient precision in research. The achievement or non-achievement of the desired sample size should be reported or justified and discussed accordingly.

2.5.3. Usefulness

Credible research warrants attention due to the value of its results. The usefulness of research results lies in their importance, the provision of meaningful estimates, and the fair conclusion supported by the research design.

Important outcomes are of high priority and concern to the end-users. These generally refer to the hard outcomes, strong correlates, or intermediate markers to the research exposure, including condition diagnoses and surrogates (blood or serum markers).

Research results are meaningful when they are easily understood in the context of clinical practice or the daily lives of patients. Meaningful estimates are usually the direct results of the study, such as the actual numbers of occurrences, incidence and prevalence rates, and risk ratios. Indirect outcome measures, such as plasma glucose excursion, and transformed estimates, such as standardized or the logarithmic form of a unit of measurement, require a reverse transformation of the units for better clarity. Otherwise, they could complicate the translation and interpretation of results.

The conclusion of a research serves as a crucial complement to the readers' own judgment. As the final interpretation and remarks provided by the authors and investigators of the research, it is essential to present the results within the appropriate context and applicability, taking into account the study design constraints and any limitations encountered throughout the research process.

3. Data analysis

The principal investigator (B.H.C) bears the overall responsibility for the compilation, maintenance, and management of the review database. The database is stored on a password-protected computer and shared on a university server repository on completion.

Every eligible and included journal article will be assessed according to two main areas – research characteristics and quality of the research. The dataset will be checked for any missing data and errors. The data will then be reported descriptively; frequency and percentage for categorical data; mean and standard deviation (median and interquartile range) for normally distributed (and not normally distributed) continuous data. Time series plots will be used to investigate the trends and patterns of research characteristics, health conditions, and quality of research over the years. The geographic information system may also be plotted to evaluate the locations and areas of research conducted. Longitudinal trends in certain research characteristics, health conditions, or research areas across different settings and clinical or biomedical disciplines will be explored.

Associations between research characteristics and quality will be explored, and the independent effect of each determinant will be quantified using multiple linear regression analysis. In addition, research quality, as a categorical outcome, will be analyzed in tertiles. The highest tertile will be compared to the lowest tertile, and the determinants will be assessed using multiple logistic regression. Longitudinal trends in research quality will be explored, where a calculated 95% confidence interval and two-sided α of 0.05 will be used to test significance. Model checking will be conducted to get the best and most parsimonious final model that meets statistical assumptions. Estimates will be generated using PASW 25.0 (Statistical Package for the Social Sciences, United States of America) and MLwiN version 3.02 (Centre for Multilevel Modeling, University of Bristol, UK).

4. Discussion

Analytical results obtained will be informative to all stakeholders of clinical and biomedical research regarding the evolution of research conduct and performance from the past to the present. Research profiles throughout the past decades may be studied in relation to socioeconomic, political, or policy changes in specific years. The longitudinal and prospective trends in the research profiles, research quality, and the association between them could

provide suggestions for improvement initiatives or identify an institutional role model that has achieved a certain degree of success. In addition, analyzing health conditions or research areas across different settings and determining whether they are over- or under-studied may help guide future prioritization of research initiatives and resource allocation. Descriptive comparison between countries may also be possible if there are similar studies conducted in other countries. This provides meaningful benchmarking and insights into the effects of evolving historical events in clinical and biomedical research activities and quality in each country.

The research quality tool proposed in this article may be a useful screening tool for all quantitative study designs, except qualitative studies, case reports, and systematic reviews. We anticipate that it would be a useful tool for a quick critical appraisal of research quality. The sequence of Relevance-Credibility-Usefulness (acronymized as RECRUS) may enhance efficiency and empower tool users in the critical appraisal process. The main limitation of this tool would be the reporting quality of the research, including zero reporting or null publication of any completed study.[36] In addition, a relatively large number of graduate and postgraduate students' research projects that were published as thesis and not in journals[37] may not be accessible through the tool's search strategies. Reporting quality is not assessed using the research quality tool proposed in this article, as specific guides and checklists already exist for this purpose. While the quality and comprehensiveness of research reporting may not be as poor as methodological research quality, they can still affect its assessment.[38] The 10 items within the three domains of the research quality screening tool are considered fundamental minimums for most clinical and biomedical research and are expected to be available in most published articles. Missing information in the included articles may be recovered by contacting the corresponding authors via email or telephone.

5. Conclusion

This systematic review aims to provide a comprehensive overview of the clinical and biomedical research landscape in Malaysia and Indonesia from 1962 to 2019. By systematically identifying and characterizing research output over this period, the review will establish a baseline for evaluating research productivity, methodological rigor, and quality. The findings are expected to highlight both the progress made and the existing gaps in research, offering critical insights for researchers, funding bodies, and policymakers. Ultimately, this work will serve as a catalyst for future efforts to enhance research capacity, reduce waste, and promote evidence-based practices in the region. Continued monitoring and comparative evaluations over time will further support strategic planning and resource allocation to strengthen the research ecosystem in Malaysia and Indonesia.

Acknowledgments

The authors would like to thank the following colleagues for their helpful suggestions during the protocol development stage: Dr. Alfi Yasmina (Lambung Mangkurat University, Indonesia), Dr. Tan Kit-Aun, Dr. Sanjiv Rampal Lekhraj, Dr. Dhashani Sivaratnam, Dr. Maliza Mawardi, Associate Professor Dr. Cheong Ai Theng, Professor Dr. Lee Ping Yein, and Professor Dr. Ching Siew Mooi (Universiti Putra Malaysia, Malaysia).

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Boon-How Chew

Investigation: Boon-How Chew, Shaun Wen Huey Lee, Lim Poh Ying, Soo Huat Teoh, Aneesa Abdul Rashid, Navin Kumar Devaraj, Adibah Hanim Ismail, Abdul Hadi Abdul Manap, Fadzilah Mohamad, Aaron Fernandez, Hanifatiyah Ali, Puteri Shanaz Jahn Kassim, Nurainul Hana Shamsuddin,Noraina Muhamad Zakuan, Akiza Roswati Abdullah

Writing - original draft: Boon-How Chew, Shaun Wen Huey Lee

Writing - review & editing: Boon-How Chew, Shaun Wen Huey Lee, Lim Poh Ying, Soo Huat Teoh, Aneesa Abdul Rashid, Navin Kumar Devaraj, Adibah Hanim Ismail, Abdul Hadi Abdul Manap, Fadzilah Mohamad, Aaron Fernandez, Hanifatiyah Ali, Puteri Shanaz Jahn Kassim, Nurainul Hana Shamsuddin,Noraina Muhamad Zakuan, Indah S. Widyahening

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The protocol proposed in this article has been registered in PROSPERO (CRD42020152907; 2020; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020152907) and the Open Science Framework registry for Research on the Responsible Conduct of Research (https://osf.io/w85ce).

References

1. UNESCO. *UNESCO Science Report: Towards 2030*. Paris, France: United Nations Educational, Scientific and Cultural Organization; 2015. Available from: https://unesdoc.unesco.org/ark:/48223/pf0000235406 [Last accessed on 20 Mar 2025].

2. Chalmers I, Bracken MB, Djulbegovic B, *et al*. How to increase value and reduce waste when research priorities are set. *Lancet.* 2014;383(9912):156-165.

   doi: 10.1016/S0140-6736(13)62229-1

3. Ioannidis JP, Greenland S, Hlatky MA, *et al*. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet.* 2014;383(9912):166-175.

   doi: 10.1016/S0140-6736(13)62227-8

4. Chan AW, Song F, Vickers A, *et al*. Increasing value and reducing waste: Addressing inaccessible research. *Lancet.* 2014;383(9913):257-266.

   doi: 10.1016/S0140-6736(13)62296-5

5. Macleod MR, Michie S, Roberts I, *et al*. Biomedical research: Increasing value, reducing waste. *Lancet.* 2014;383(9912):101-104.

   doi: 10.1016/S0140-6736(13)62329-6

6. Glasziou P, Altman DG, Bossuyt P, *et al*. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet.* 2014;383(9913):267-276.

   doi: 10.1016/S0140-6736(13)62228-X

7. Ioannidis JPA. Why most published research findings are false. *PLoS Med.* 2005;2(8):e124.

   doi: 10.1371/journal.pmed.1004085

8. Deeks JJ, Dinnes J, D'Amico R, *et al*. Evaluating non-randomised intervention studies. *Health Technol Assess.* 2003;7(27):3-10, 1-173.

   doi: 10.3310/hta7270

9. MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AM. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technol Assess.* 2000;4(34):1-154.

10. Zeng X, Zhang Y, Kwong JS, *et al*. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: A systematic review. *J Evid Based Med.* 2015;8(1):2-10.

    doi: 10.1111/jebm.12141

11. Jiu L, Hartog M, Wang J, *et al*. Tools for assessing quality of studies investigating health interventions using real-world data: A literature review and content analysis. *BMJ Open*. 2024;14(2):e075173.

    doi: 10.1136/bmjopen-2023-075173

12. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.5*, *Cochrane*; 2024. Available from https://www.training.cochrane.org/handbook

13. Higgins JP, Altman DG, Gotzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343:d5928.

    doi: 10.1136/bmj.d5928

14. Whiting PF, Rutjes AW, Westwood ME, *et al*. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.

    doi: 10.7326/0003-4819-155-8-201110180-00009

15. Shea BJ, Grimshaw JM, Wells GA, *et al*. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7:10.

    doi: 10.1186/1471-2288-7-10

16. Whiting P, Savovic J, Higgins JP, *et al*. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* 2016;69:225-234.

    doi: 10.1016/j.jclinepi.2015.06.005

17. Sterne JA, Hernan MA, Reeves BC, *et al*. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919.

    doi: 10.1136/bmj.i4919

18. Wells GA, Shea B, O'Connell D, *et al*. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses*. Ottawa Hospital Research Institute; 2011. Available from: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [Last accessed on 2024 Mar 20].

19. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377-384.

    doi: 10.1136/jech.52.6.377

20. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. *J Clin Epidemiol.* 2012;65(2):163-178.

    doi: 10.1016/j.jclinepi.2011.05.008

21. Lo CK, Mertz D, Loeb M. Newcastle-Ottawa Scale: Comparing reviewers' to authors' assessments. *BMC Med Res Methodol.* 2014;14:45.

    doi: 10.1186/1471-2288-14-45

22. Hartling L, Milne A, Hamm MP, *et al*. Testing the newcastle ottawa scale showed low reliability between individual reviewers. *J Clin Epidemiol.* 2013;66(9):982-993.

doi: 10.1016/j.jclinepi.2013.03.003

23. Oremus M, Oremus C, Hall GB, McKinnon MC. Inter-rater and test-retest reliability of quality assessments by novice student raters using the Jadad and Newcastle-Ottawa Scales. *BMJ Open.* 2012;2(4):e001368.

doi: 10.1136/bmjopen-2012-001368

24. Margulis AV, Pladevall M, Riera-Guardia N, *et al*. Quality assessment of observational studies in a drug-safety systematic review, comparison of two tools: The Newcastle-Ottawa Scale and the RTI item bank. *Clin Epidemiol.* 2014;6:359-368.

doi: 10.2147/CLEP.S66677

25. Guyatt G, Drummond R, Meade M, Cook D. *The Evidence Based-Medicine Working Group Users' Guides to the Medical Literature.* 2nd ed. Chicago: McGraw Hill; 2008.

26. Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. *JAMA.* 1993;270(17):2093-2095.

27. Von Niederhausern B, Schandelmaier S, Mi Bonde M, *et al*. Towards the development of a comprehensive framework: Qualitative systematic survey of definitions of clinical research quality. *PLoS One.* 2017;12(7):e0180635.

doi: 10.1371/journal.pone.0180635

28. Belcher BM, Rasmussen KE, Kemshaw MR, Zornes DA. Defining and assessing research quality in a transdisciplinary context. *Res Eval.* 2015;25(1):1-17.

doi: 10.1093/reseval/rvv025

29. Grobbee DE, Hoes AW. *Clinical Epidemiology: Principles, Methods, and Applications for Clinical Research.* United States: Jones and Bartlett Learning; 2014.

30. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Syst Rev*. 2022;18:e1230.

doi: 10.1002/cl2.1230

31. McGinn T, Wyer PC, Newman TB, Keitz S, Leipzig R, for GG. Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). *CMAJ.* 2004;171(11):1369-1373.

doi: 10.1503/cmaj.1031981

32. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb).* 2012;22(3):276-282.

33. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor Quant Methods Psychol.* 2012;8(1):23-34.

doi: 10.20982/tqmp.08.1.p023

34. Bujang MA, Baharum N. Guidelines of the minimum sample size requirements for Kappa agreement test. *Epidemiol Biostatist Public Health.* 2017;14(2):1-10.

doi: 10.2427/12267

35. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther.* 2005;85(3):257-268.

36. Groves T. What makes a high quality clinical research paper? *Oral Dis.* 2010;16(4):313-315.

doi: 10.1111/j.1601-0825.2010.01663.x

37. Post RE, Mainous AG 3rd, O'Hare KE, King DE, Maffei MS. Publication of research presented at STFM and NAPCRG conferences. *Ann Fam Med.* 2013;11(3):258-261.

doi: 10.1370/afm.1503

38. Widyahening IS, Wangge G, Saldi SR, *et al*. Quality and reporting of publications by Indonesian researchers: A literature survey. *J Evid Based Med.* 2014;7(3):163-171.

doi: 10.1111/jebm.12112

Appendix

Table A1. Data extraction form: Author and article characteristics

Article ID	Title	Institution of the CA	Specialty of the CA[a]	Professional qualification of the CA	Number of authors	Number of institutions	Number of overseas authors	Number of overseas institutions	Number of specialties/ disciplines of all authors	Journal name	Journal locality	Journal scope	Journal subscription
First name, year	Full title	Name, Country	(i) Family medicine (ii) Public health (iii) Medical (iv) Pediatric (v) Psychiatry (vi) Surgery (vii) O&G (viii) Orthopedic (ix) ENT (x) Eye (xi) Pharmacy (xii) Nursing (xiii) Dietetic (xiv) Biomedical	(i) PhD (clinician) (ii) PhD (non-clinician) (iii) MD (iv) Master (MMed) (v) Master (non-MMed) (vi) Bachelor (vii) Diploma (viii) Others (lower degrees) (ix) Not stated	1 to _	1 to _	0 to _	0 to _	1 to _	Full name	(i) Local-Malaysia (ii) Regional-within Asia (iii) International	General (multidisciplinary) Specific	OA Subscription

Note: [a]Up to two corresponding authors.
Abbreviations: CA: Corresponding author; ENT: Ear, nose, and throat; ID: Identity; MMed: Master's in Medicine; OA: Open-access; O&G: Obstetrics and gynecology.

Table A2. Example of a complete data extraction form: Author and article characteristics

Article ID	Title	Institution of the CA	Specialty of the CA[a]	Professional qualification of the CA	Number of authors	Number of institutions	Number of overseas authors	Number of overseas institutions	Number of specialties/ disciplines of all authors	Journal name	Journal locality	Journal scope	Journal subscription
Chan KE 1962 (2 variables)	String	String (2 variables)	Nominal	Ordinal Highest professional qualification	Scale	Scale	Scale	Scale Including local authors affiliated with overseas institutions	Scale	String	Ordinal Regional; i.e., Asian regions	Nominal General versus discipline-specific	Nominal OA versus traditional subscription-based

Note: [a]Up to two corresponding authors. Abbreviations: CA: Corresponding author; ID: Identity; OA: Open-access.

Table A3. Data extraction form: Research characteristics

Type of article	Field of study	Level of study	Class of study	Type of quantitative study	Data collection	Year when the study was conducted and completed	Number of study site	Setting[a]	Condition/ organ system studied[b]	Measures[c]	Intervention
(i) Research (original and reviews) (ii) Research protocol (iii) Audits (iv) Case report/ series (v) Commentary, quizzes, CME, etc. (vi) Letters	(i) Clinical (ii) Public health (iii) Health services (iv) Laboratory (v) Education	(i) Primary (ii) Secondary (database) (iii) Secondary (review) (iv) Opinion	(i) Quantitative (ii) Qualitative (iii) Mixed (iv) Others	(i) Prevalence (ii) Diagnostic (iii) Diagnostic test (iv) Prognostic (v) Etiologic (vi) Interventional (vii) Psychometric study	(i) Cross-sectional (ii) Cohort (iii) Case–control (iv) Nested case–control (v) RCT (vi) Quasi-RCT (voi) Cross-over (viii) Secondary data retrieval	Year started; year completed	1 to _	(i) Hospital (ii) Clinic (iii) Community (iv) Laboratory (v) University/ college	ICPC-2[d]	*Subjective*: (i) Self-report (ii) Other's report *Objective*: (i) Operator dependence (ii Non-operator-dependence (e.g., lab analyzer)	(i) No (ii) Drug (iii) Device, tool, or app (iv) Surgical procedure (v) Psychological or behavioral (vi) Socio-economic (vii) Health services

Notes: [a]Up to two corresponding authors; [b]consists of two variables, the primary and secondary settings; [c]consists of two variables, the subjective and objective outcome measures; [d]check Table A5.
Abbreviations: CME: Continuous medical education; RCT: Randomized controlled trial; ICPC-2: International classification of primary care (Second Edition).

Table A4. Example of a complete data extraction form: Research characteristics

Type of article	Field of study	Level of study	Class of study	Type of quantitative study	Data collection	Year when the study was conducted and completed	Number of study site	Setting[a]	Condition/ organ system studied[b]	Measures[c]	Intervention
Nominal (Only proceed with no. 1)	Nominal	Nominal	Nominal	Nominal	Nominal	Scale 2011, 2015 (2 variables)	Scale	Nominal	Nominal (2 variables- primary and secondary)	Nominal	Nominal

Notes: [a]Up to two corresponding authors; [b]consists of two variables, the primary and secondary settings; [c]consists of two variables, the subjective and objective outcome measures.

Table A5. International classification of primary care (ICPC) chapters and components

Chapter	Component
A	General and unspecified
B	Blood, blood-forming organs, lymphatics, spleen
D	Digestive
F	Eye
H	Ear
K	Circulatory
L	Musculoskeletal
N	Neurological
P	Psychological
R	Respiratory
S	Skin
T	Endocrine, metabolic, and nutritional
U	Urology
W	Pregnancy, childbirth, family planning
X	Female genital system and breast
Y	Male genital system
Z	Social problems