Official Journal of the
Malaysian Medical Association

The Medical Journal of Malaysia

Volume: 80

Issue No: 3

May 2025

ISSN 0300-5283 PP 2121/01/2013 (031329) MITA(P) 124/1/91

MJM

Official Journal of the Malaysian Medical Association

Volume 80 Number 3 May 2025

EDITORIAL BOARD

PP 2121/01/2013 (031329) MCI (P) 124/1/91 ISSN 0300-5283

The Medical Journal of Malaysia is published six times a year.
MJM is published bimonthly ie. January, March, May, July, September and November.

Advertisement Rates:
Enquiries to be directed to the Secretariat.

Subscription Rates:
Price per copy is RM100.00 or RM360.00 per annum, for all subscribers.

Secretariat Address:
Malaysian Medical Association
4th Floor, MMA House, 124, Jalan Pahang, 53000 Kuala Lumpur.
Tel: (03) 4042 0617, 4041 8972, 4041 1375 Fax: (03) 4041 8187
E-mail: info@mma.org.my / mjm@mma.org.my
Website: www.mma.org.my

Printed by: Digital Perspective Sdn. Bhd.
42-1, Level 1, Plaza Sinar, Taman Sri Sinar, 51200 Kuala Lumpur. Tel: 03-6272 3767
Email: dpsbkl@gmail.com

The Medical Journal of Malaysia

MJM

INSTRUCTION TO AUTHORS

The *Medical Journal of Malaysia* ***(MJM)*** welcomes articles of interest on all aspects of medicine in the form of original papers, review articles, short communications, continuing medical education, case reports, commentaries and letter to Editor. Articles are accepted for publication on condition that they are contributed solely to *The Medical Journal of Malaysia*.

NOTE: MJM is published bimonthly ie. January, March, May, July, September and November.

REQUIREMENTS FOR ALL MANUSCRIPTS
Please ensure that your submission to MJM conforms to the International Committee of Medical Journal Editors Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals.

Neither the Editorial Board nor the Publishers accept responsibility for the views and statements of authors expressed in their contributions.

The Editorial Board further reserves the right to reject papers read before a society. To avoid delays in publication, authors are advised to adhere closely to the instructions given below.

MANUSCRIPTS
Manuscripts should be submitted in English (British English). Manuscripts should be submitted online through *MJM Editorial Manager*, http://www.editorialmanager.com/mjm.

Instructions for registration and submission are found on the website. Authors will be able to monitor the progress of their manuscript at all times via the *MJM Editorial Manager*. For authors and reviewers encountering problems with the system, an online Users' Guide and FAQs can be accessed via the "Help" option on the taskbar of the login screen.

MJM charges a one-time, non-refundable Article Processing Charge (APC) upon submission. Waiver of the APC applies only to members of the editorial board, and authors whose articles are invited by the editor. In addition, recipients of the MJM Reviewer Recognition Award from the previous year may enjoy a waiver of the APC for the next calendar year (e.g. recipients of MJM Reviewer Recognition Award 2022 will enjoy waiver of APC for articles submitted between January and December 2023).

MJM
Member: RM500
Non Member: RM800
Overseas: USD200

MJM Case Report
Member: RM400
Non Member: RM500

Preparing your manuscript

The MJM Article Processing Charge is a non-refundable administrative fee. Payment of the APC does not guarantee acceptance of the manuscript. Submitted articles will only be sent for reviews once the MJM APC has been successful completed.

All submissions must be accompanied by a completed **Copyright Assignment Form**, **Copyright Transfer Form and Conflict of Interest Form** duly signed by all authors. Forms can be download from MJM website at https://www.e-mjm.org/

Manuscript text should be submitted as **Microsoft Word** documents. Tables and flow-charts should be submitted as **Microsoft Word** documents. Images should be submitted as separate **JPEG files** (minimum resolution of 300 dpi).

PEER REVIEW PROCESS
All submissions must include at least two (2) names of individuals who are especially qualified to review the work. All manuscripts submitted will be reviewed by the Editor in-charge before they are send for peer review. Manuscripts that are submitted to MJM undergo a double-blinded peer review and are managed online. Proposed reviewers must not be involved in the work presented, nor affiliated with the same institution(s) as any of the authors or have any potential conflicts of interests in reviewing the manuscript. The selection of reviewers is the prerogative of the Editors of MJM.

ELIGIBILITY AS AN AUTHOR
MJM follows the recommendation of the International Committee of Medical Journal Editors (ICMJE) for eligibility to be consider as an author for submitted papers. The ICMJE recommends that authorship be based on the following four (4) criteria:

1 Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
2 Drafting the work or revising it critically for important intellectual content; AND
3 Final approval of the version to be published; AND
4 Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

TYPES OF PAPERS
Original Articles:
Original Articles are reports on findings from original unpublished research. Preference for publications will be given to high quality original research that make significant contribution to medicine. Original articles shall consist of a structured Abstract and the Main Text. The word count for the structured abstract should not exceed 500 words. The main text of the articles should not exceed 4000 words, tables/illustrations/figures/images up to five (5) and references up to 40. Manuscript describing original research should conform to the IMRAD format, more details are given below.

Original articles of cross-sectional and cohort design should follow the corresponding STROBE check-lists; clinical trials should follow the CONSORT check-list.

Review Articles:
Review Articles are solicited articles or systematic reviews. ***MJM*** solicits review articles from Malaysian experts to provide a clear, up-to-date account of a topic of interest to medical practice in Malaysia or on topics related to their area of expertise. Unsolicited reviews will also be considered, however, authors are encouraged to submit systematic reviews rather than narrative reviews. Review articles shall consist of a structured Abstract and the Main Text. The word count for the structured abstract should not exceed 500 words. Systematic Review are papers that presents exhaustive, critical assessments of the published literature on relevant topics in medicine. Systematic reviews should be prepared in strict compliance with MOOSE or PRISMA guidelines, or other relevant guidelines for systematic reviews.

Short Communications:
Shorts communication are short research articles of important preliminary observations, findings that extends previously published research, data that does not warrant publication as a full paper, small-scale clinical studies, and clinical audits. Short communications should not exceed 1,500 words and shall consist of a Summary and the Main Text. The summary should be limited to 100 words and provided immediately after the title page. The number of tables/illustrations/figures/images should be limited to three (3) and the number of references to ten (10).

Continuing Medical Education (CME) Articles:
A CME article is a critical analysis of a topic of current medical interest. The article should include the clinical question or issue and its importance for general medical practice, specialty practice, or public health. It shall consist of a Summary and the Main Text. The summary should be limited to 500 words and provided immediately after the title page Upon acceptance of selected articles, the authors will be requested to provide five multiple-choice questions, each with five true/false responses, based on the article. For guideline, please refer to: Sivalingam N, Rampal L. Writing Articles on Continuing Medical Education for Medical Journals. Med J Malaysia. 2021 Mar;76(2):119-124.

Case Reports:
Papers on case reports (one to five cases) must follow these rules: Case reports should not exceed 2,000 words; with a maximum of two (2) tables; three (3) photographs; and up to ten (10) references. It shall consist of a Summary and the Main Text. The summary should be limited to 250 words and provided immediately after the title page. Having a unique lesson in the diagnosis, pathology or management of the case is more valuable than mere finding of a rare entity. Being able to report the outcome and length of survival of a rare problem is more valuable than merely describing what treatment was rendered at the time of diagnosis. There should be no more than seven (7) authors.

Please note that all Case Reports will be published in the new MJM Case Reports Journal (www.mjmcasereports.org).

Commentaries:
Commentaries will usually be invited articles that comment on articles published in the same issue of the ***MJM***. However, unsolicited commentaries on issues relevant to medicine in Malaysia are welcomed. They should not exceed 2,000 words. They maybe unstructured but should be concise. When presenting a point of view, it should be supported with the relevant references where necessary.

Letters to Editor:
Letters to Editors are responses to items published in ***MJM*** or to communicate a very important message that is time sensitive and cannot wait for the full process of peer review. Letters that include statements of statistics, facts, research, or theories should include only up to three (3) references. Letters that are personal attacks on an author will not be considered for publication. Such correspondence must not exceed 1,500 words.

Editorials:
These are articles written by the editor or editorial team concerning the ***MJM*** or about issues relevant to the journal.

STRUCTURE OF PAPERS
Title Page:
The title page should state the brief title of the paper, full name(s) of the author(s) (with the surname or last name bolded), degrees (limited to one degree or diploma), affiliation(s), and corresponding author's address. All the authors' affiliations shall be provided after the authors' names. Indicate the affiliations with a superscript number at the end of the author's degrees and at the start of the name of the affiliation. If the author is affiliated to more than one (1) institution, a comma should be used to separate the number for the said affiliation.

Do provide preferred abbreviated author names for indexing purpose, e.g. L Rampal (for Lekhraj Rampal), BS Liew (for Liew Boon Seng), B Abdullah (for Baharudin Abdullah), Hoe VC (for Victor Hoe Chee Wai).
Please indicate the corresponding author and provide the affiliation, full postal address and email.

Articles describing Original Research should consist of the following sections (IMRAD format): Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgment and References. Each section should begin on a fresh page.
Scientific names, foreign words and Greek symbols should be in italic.

Abstract and Key Words:
A structured abstract is required for Original and Review Articles. It should be limited to 500 words and provided immediately after the title page. Below the abstract provide and identify three (3) to 10 key words or short phrases that will assist indexers in cross-indexing your article. Use terms from the medical subject headings (MeSH) list from Index Medicus for the key words where possible. Key words are not required for Short Communications, CME articles, Case Reports, Commentaries and Letter to Editors.

Introduction:
Clearly state the purpose of the article. Summarise the rationale for the study or observation. Give only strictly pertinent references, and do not review the subject extensively.

Materials and Methods:
Describe your selection of the observational or experimental subjects (patients or experimental animals, including controls) clearly, identify the methods, apparatus (manufacturer's name and address in parenthesis), and procedures in sufficient detail to allow other workers to reproduce the results. Give references to established methods, including statistical methods; provide references and brief descriptions of methods that have been published but are not well-known; describe new or substantially modified methods, give reasons for using them and evaluate their limitations.

Identify precisely all drugs and chemicals used, including generic name(s), dosage(s) and route(s) of administration. Do not use patients' names, initials or hospital numbers. Include numbers of observation and the statistical significance of the findings when appropriate.

When appropriate, particularly in the case of clinical trials, state clearly that the experimental design has received the approval of the relevant ethical committee.

Results:
Present your results in logical sequence in the text, tables and illustrations. Do not repeat in the text all the data in the tables or illustrations, or both: emphasise or summarise only important observations in the text.

Discussion:
Emphasise the new and important aspects of the study and conclusions that follow from them. Do not repeat in detail data given in the Results section. Include in the Discussion the implications of the findings and their limitations and relate the observations to other relevant studies.

Conclusion:
Link the conclusions with the goals of the study but avoid unqualified statements and conclusions not completely supported by your data. Avoid claiming priority and alluding to work that has not been completed. State new hypotheses when warranted, but clearly label them as such. Recommendations, when appropriate, may be included.

Acknowledgements:
Acknowledgements of general support, grants, technical assistance, etc., should be indicated. Authors are responsible for obtaining the consent of those being acknowledged.

Referencing guide:
The Medical Journal of Malaysia, follows the Vancouver numbered referencing style. Citations to someone else's work in the text, should be indicated by the use of a number. In citing more than one article in the same sentence, you will need to include the citation number for each article. A hyphen should be used to link numbers which are inclusive, and a comma used where numbers are not consecutive. The following is an example where works 1.3,4,5.have been cited in the same place in the text.

Several effective drugs are available at fairly low cost for treating patients with hypertension and reducing the risk of its sequelae.[1,3-5]

The list of all of the references that are cited in the article should be presented in a list labelled as 'References'. This reference list appears at the end of the paper. Authors are responsible for the accuracy of cited references and these should be verified by the author(s) against the original documents before the manuscript is submitted. It is important that the author should never place in the list of references a document that he or she has not seen. The Journals names should be abbreviated according to the style used in the Index Medicus. All authors when six or less should be listed; when seven or more list only the first six and add et al.

If you are citingthe author's name in your text, you must insert the citation number as well. Jewell BL (8) underlined that as focus in the SARS-CoV-2 pandemic shifts to the emergence of new variants of concern (VOC), characterising the differences between new variants and non-VOC lineages will become increasingly important for surveillance and maintaining the effectiveness of both public health and vaccination programme. If you are citing more than one author's name in your text and you want to cite author names in your text, use 'et al.' after the first author. Example: Rampal et al. (9) highlighted that the. disregard of the manuscript guidelines and instruction to authors of the journal you submit, is one of the common reasons for 'Rejection' of the article.

Example references Journals:
Standard Journal Article
Rampal L and Liew BS. Coronavirus disease (COVID-19) pandemic. Med J Malaysia 2020; 75(2): 95-7.
Rampal L, Liew BS, Choolani M, Ganasegeran K, Pramanick A, Vallibhakara SA, et al. Battling COVID-19 pandemic waves in six South-East Asian countries: A real-time consensus review. Med J Malaysia 2020; 75(6): 613-25.
NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021; 11; 398(10304): 957-80.

Books and Other Monographs:
Personal Author(s)
Goodman NW, Edwards MB. 2014. Medical Writing: A Prescription for Clarity. 4 th Edition. Cambridge University Press.

Chapter in Book
McFarland D, Holland JC. Distress, adjustments, and anxiety disorders. In: Watson M, Kissane D, Editors. Management of clinical depression and anxiety. Oxford University Press; 2017: 1-22.

Corporate Author
World Health Organization, Geneva. 2019. WHO Study Group on Tobacco Product Regulation. Report on the scientific basis of tobacco product regulation: seventh report of a WHO study group. WHO Technical Report Series, No. 1015.

NCD Risk Factor Collaboration (NCD-RisC). Rising rural body-mass index is the main driver of the global obesity epidemic in adults. Nature 2019; 569: 260–64.

World Health Organization. Novel Coronavirus (2019-nCoV) Situation Report 85, April 14, 2020. [cited April 2020] Accessed from: https://www.who.int/docs/defaultsource/coronaviruse/situationreports/20200414-sitrep-85-covid-19.

Online articles
Webpage: Webpage are referenced with their URL and access date, and as much other information as is available. Cited date is important as webpage can be updated and URLs change. The "cited" should contain the month and year accessed.

Ministry of Health Malaysia. Press Release: Status of preparedness and response by the ministry of health in and event of outbreak of Ebola in Malaysia 2014 [cited Dec 2014]. Available from: http://www.moh.gov.my/english.php/database_stores/store_view_page/21/437.

Other Articles:
Newspaper Article
Panirchellvum V. 'No outdoor activities if weather too hot'. the Sun. 2016; March 18: 9(col. 1-3).

Magazine Article
Rampal L.World No Tobacco Day 2021 -Tobacco Control in Malaysia. Berita MMA. 2021; May: 21-22.

Tables:
All tables and figures should have a concise title and should not occupy more than one printed page. The title should concisely and clearly explain the content of the table or figure. They should be numbered consecutively with Roman numerals (e.g Table I) and figures with Arabic numerals (e.g. Figure 1), and placed after the sections of the manuscript which they reflect, particularly the results which they describe on separate pages. Cite tables in the text in consecutive order. Indicate table footnotes with lower-case letters in superscript font. Place the information for the footnote beneath the body of the table. If a table will be submitted as a separate document, the filename should contain the surname of the first author and match its label in the manuscript (e.g., SMITH Table I).Vertical lines should not be used when constructing the tables. All tables and figures should also be sent in electronic format on submission of the manuscript as supplementary files through the journal management platform. Clinical Photographs should conceal the subject's identity. Tables and flow-charts should be submitted as Microsoft Word documents. Images should be submitted as separate JPEG files (minimum resolution of 300 dpi).

Photographs of Patients:
Proof of permission and/or consent from the patient or legal guardian must be submitted with the manuscript. A statement on this must be included as a footnote to the relevant photograph.

Colour reproduction:
Illustrations and diagrams are normally reproduced in black and white only. Colour reproductions can be included if so required and upon request by the authors. However, a nominal charge must be paid by the authors for this additional service; the charges to be determined as and when on a per article basis.

Abbreviations:
Use only standard abbreviations. The full-term for which an abbreviation stands should precede its first use in the abstract, article text, tables, and figures, unless it is a standard unit of measurement. Abbreviations shall not be used in the Title. Abbreviations should be kept to a minimum.

Formatting of text:
Numbers one to ten in the text are written out in words unless they are used as a unit of measurement, except in tables and figures. Use single hard-returns to separate paragraphs. Do not use tabs or indents to start a paragraph. Do not use the automated formatting of your software, such as hyphenation, endnotes, headers, or footers (especially for references). Submit the Manuscript in plain text only, removed all 'field codes' before submission. Do not include line numbers. Include only page number.

BEST PAPER AWARD
All original papers which are accepted for publication by the MJM, will be considered for the 'Best Paper Award' for the year of publication. No award will be made for any particular year if none of the submitted papers are judged to be of suitable quality.

CONTENTS Page

Editorial

Original Articles

CONTENTS Page

Eye cases that widened the vision on medical negligence

By NKS Tharmaseelan LLB

M. Ananth, Suzieta & Co, Petaling Jaya, Selangor, Malaysia

Medical negligence law has undergone a significant transformation, shifting from a paternalistic, doctor-centred model to one emphasising patient autonomy and informed consent. Two landmark eye cases changed the law on medical negligence. The Australian case of Rogers vs. Whitaker (1992) moved the pendulum from being doctor-centric to being more patient-oriented.[1] Rogers' case was adopted in Malaysia in Foo Fio Na vs. Dr. Soo Mun & Anor [2007].[2] However, the Foo Fio Na case created some ambiguity as to whether the Rogers vs. Whitaker decision applied only on the standard of care for the disclosure of risks and advice, or whether it was applicable for diagnosis and treatment only or for both standards of care as Rogers was interpreted in Naxakis (1999)[3] This uncertainty was finally laid to rest in the Dr. Hari Krishnan & Anor vs. Megat Noor Ishak bin Megat Ibrahim & Anor appeal.[4] The latter case, also referred to as the Eye Appeal, set several new benchmarks in medical negligence law.[5] These eye cases widened the vision on medical negligence. This editorial, reviews these pivotal cases that have reshaped medical negligence law in Malaysia, highlighting their implications for clinical practice and patient rights.

Rogers vs. Whitaker

Material Facts[6]
Ms. Whitaker was almost blind in her right eye since a penetrating injury at the age of nine. She had continued to lead a substantially normal life, completing her schooling, entering the workforce, marrying, and raising a family. Forty years later, following consultation with Dr. Cohen, an Ophthalmologist, she was referred to Dr. Rogers for a surgical view. Ms. Whitaker saw Dr. Rogers only after a year. Ms. Whitaker, then aged 47, was advised to undergo surgery on the right eye, as it would, in addition to cosmetic improvement will markedly restore vision in that eye. Ms. Whitaker was concerned about any risks or complications of surgery to the right eye. The trial judge noted that Ms. Whitaker had "incessantly raised questions, including the risks of unintended or accidental interference with her good, left eye." She had even requested to cover her good eye to ensure that the wrong eye would not be operated on. Ms. Whitaker did not specifically ask whether the surgery on her right eye could affect her left eye. It was only after three weeks following the consultation that she consented to surgery.

Post-surgery, Ms. Whitaker's left eye became blind due to sympathetic ophthalmia. Many experts supported Dr. Rogers' action of not informing the risk of Sympathetic Ophthalmia during the consent process, as it was very rare, an incidence rate of only 1:14,000 or 0.007% although there was also evidence that the chance of occurrence was slightly greater when, as in this case where there had been an earlier penetrating injury to the eye operated upon.[7] Disclosure of risks was in line with the then-prevalent Bolam principle,[8] which was doctor-centric, whereby the profession decides on the standard of care for disclosure of risks, provided the decision withstands logical analysis by the courts as held in Bolitho.[9] It was a single comprehensive duty encompassing both standards of care, diagnosis and treatment, and disclosure of risks. The experts further asserted that they would only inform the patient of this risk if asked specifically about this complication.[10]

Held On appeal, the appellate court held that the Bolam/Bolitho test should only be applied as the standard of care for diagnosis and treatment. In this case, there was no alleged lack of skill in performing the surgery. The Rogers case involved disclosure of risks and advice. The courts held, "It is not for the profession to decide what is to be disclosed. It is for the patient to make his/her own decisions about his/her life." The courts held that for the patient to either accept or refuse the proposed treatment after receiving the relevant appropriate information. The patient had to decide whether sufficient information was given to him to make an informed decision. This was a landmark decision, making it mandatory for doctors to warn patients of material risks inherent in a proposed treatment. The courts held that material risks are what "a reasonable person in the patient's position, if warned of the risk, would be likely to attach significance to it; or if the doctor is or should reasonably be aware that the particular patient, if warned of the risk, would be likely to attach significance to it, i.e., material risk should be taken into account and anything of a risky nature, is reasonable."

The Rogers eye case, established a separate and distinct standard of care for disclosure of risks, which was subjective and patient-centric. Blindness is a significant material risk, especially so as this patient had only one good eye. The patient was very apprehensive and anxious before the surgery. This is subjective as to the patient's expectations. This case highlights that, though the incidence rate is taken into account, it is the seriousness of the inherent risk that needs to be given primary importance.

Case 2: Dr. Hari Krishnan & Anor vs. Megat Noor Ishak bin Megat Ibrahim & Anor and another appeal [2018][11]

The patient consulted an ophthalmologist in a private hospital, who diagnosed retinal detachment and advised immediate eye surgery. Following the patient's consent, he underwent eye surgery, however, due to several complications, the ophthalmologist, the anaesthetist, and

This article was accepted: 30 April 2025
Corresponding Author: NKS Tharmaseelan
Email: nks.tharmaseelan@gmail.com

the hospital were found to be negligent, and the court awarded MYR200,000 as general damages and an unprecedented sum of MYR1,000,000 as aggravated damages. The award of damages by the High Court was reaffirmed by the Court of Appeal. The three parties then obtained leave to appeal to the Federal Court on various questions of law.

Material Facts[11]

The Patient who sustained a giant retinal tear with detachment in the right eye was referred by a general practitioner to the Ophthalmologist, who saw the patient in a private clinic on 26.8.1999. The patient was advised to undergo immediate retinal detachment operation (1st operation), which was done in an Eye Hospital. The patient was discharged two days later on 30.8.1999, and given an appointment a week later. Subsequently, "the patient's right eye became watery, his vision was blurry, and there were tears of blood when he sneezed." Alarmed, the patient rang up the Ophthalmologist immediately, who reassured him and advised him to see him on the appointment date. Nevertheless, on 4.9.1999, unable to withstand the pain, the patient went to see the Ophthalmologist at his private clinic. On visual inspection, the Ophthalmologist confirmed that there was bleeding in the eye but reassured him and advised the patient to return on the appointment date on 7.9.1999.

On 7.9.1999, when seen at his private clinic, the patient was reassured and told that recovery would be slow and fixed the next appointment at the Hospital on 14.9.1999. When seen on the date, the patient still had complaints of continuous pain and strong pressure in his eye. Following a visual examination, the patient was told that the retina of his right eye had folded outward and that a second operation had to be done immediately. However, after the ophthalmologist performed a scan, the patient was informed that his earlier findings of the folded retina and the need for a second operation were incorrect. As such, there was no need for surgery, and another appointment was fixed on 21.9.1999.
On 21.9.1999, following an examination, the patient was informed that the retina in his right eye had folded or partially detached, and proposed an operation (2nd operation) to be done immediately. The patient requested a scan to confirm the findings because he felt that his vision had improved. However, he was informed that a scan was unnecessary as the ophthalmologist was able to confirm the findings on visual inspection, and that the improved vision was only temporary, which may subsequently worsen.

The patient consented for the 2nd operation and requested for the same anaesthesiologist who gave him anaesthesia for the 1st operation, but was told that an equally competent anaesthesiologist was on duty, would administer anaesthesia. Following surgery, the next day on 22.9.1999, the ophthalmologist informed the patient, that "some problems had occurred during the 2nd Operation." The patient had apparently regained consciousness during the operation and bucked while the ophthalmologist was strengthening the retina using a laser. As a result, "the patient suffered Supra-Choroidal Haemorrhage (SCH) with profuse bleeding in his right eye." Nevertheless, the patient was reassured he "would regain eyesight provided that the retina remained intact after the bleeding in the eye subsides." He was, however, not informed of the possibility of blindness in the right eye.

Subsequently, the patient experienced severe pain, continuous bleeding and a total loss of vision in his right eye. The patient was advised "to stay in the Hospital for seven days, and to sit in an upright position at all times so that the blood in his eye could subside." The patient was discharged on 26.9.1999. The next day, a referral letter was given to see another ophthalmologist (No. 2) of the Hospital for a second opinion on the status of the right eye. Only on reading the referral letter, the patient realised that the lens in his right eye had been removed during the 2nd Operation.

The ophthalmologist (No. 2) informed the patient that "retina in his right eye was badly uprooted with a lot of internal blood clotting." He was of the opinion that the suggestion to wash the front part of the eyes by Ophthalmologist (No. 1) would be futile. On 1.10.1999, the patient returned to consult the Ophthalmologist (No. 1). The patient was told that there was still bleeding in his eye and that a procedure needed to be performed and referred the him to another Ophthalmologist (No. 3) who upon examining the patient was of the opinion that the right eye was beyond salvaging. On the advice of Ophthalmologist (No. 2) the patient consulted an Ophthalmologist (No. 4) in Singapore who informed the patient that "right eye was badly damaged, having been drenched in blood for more than 25 days." On 15.10.1999, on the advice of ophthalmologist (No. 4) the patient underwent surgery, which included the patching of the retina and the removal of blood clots, in an attempt to salvage his vision. The efforts proved futile.

In a medical report dated 24.11.1999, the Ophthalmologist (No 1) affirmed that the patient's right eye being permanently blind due to retinal detachment, and that his left eye needs prolonged follow-up treatment. The Patient filed a medical negligence suit against the Ophthalmologist, the Anaesthetist and the Hospital alleging that the injuries and loss of vision in his right eye were caused by the negligence of all three defendants.

Decision of High Court

The High Court allowed the Patient's claim and held all three defendants liable. The courts found the Ophthalmologist and the Anaesthetist negligent in "failing to warn the patient of the risks of bucking and blindness, and in the care and management of the patient" On the issue of vicarious liability, the Court found the Hospital liable for the negligence of the Ophthalmologist and the Anaesthetist. The courts held that the "internal arrangements between Ophthalmologist and the Anaesthetist with the Hospital were exclusively within their knowledge, and that the Hospital had allowed the former two to hold themselves out as the Hospital's agents, servants or employees." Accordingly, the courts awarded damages as follows to the patient: general damages of MYR200,000.00, aggravated damages MYR1,000,000.00 and special damages MYR8,014.00

Decision of the Court of Appeal (COA)
The COA affirmed the decision of the High Court. The courts held that the Ophthalmologist was negligent in his care and management of the patient in the 2nd operation. The COA found no evidence that either the Ophthalmologist or the Anaesthetist had explained the risk of bucking to the patient. The Ophthalmologist had, "wrongly advised the patient to undergo the 2nd operation, and thereby subjected him to unnecessary risks, including the instance of bucking which led to the blindness." Further, "the operation adopted by the Ophthalmologist for the haemorrhaging, was found to be against all textbook and established clinical teachings." The COA concluded that the Anaesthetist had failed to explain the risk of bucking as he had never met the patient prior to the administration of anaesthesia. Further, "he failed in his responsibility to keep the patient anaesthetised completely, relaxed, and pain-free throughout the operation." Expert witnesses asserted that bucking could have been avoided and controlled by additional drugs. The COA considered the fact that "the muscle relaxant drug wore off as a clear indication of negligence, and held that there was clear mistiming of the top-up dose."

On the matter of vicarious liability, the COA, held that "patients present themselves at a hospital to seek treatment from that hospital". In view of the "inextricable relationship between hospitals and doctors, the Hospital's liability for the negligence of the doctors, are not absolved by pure internal arrangements." The COA elaborated, "In our view in the admission of a patient, a hospital must be regarded as giving an undertaking that it would take reasonable care to provide for his medical needs. There is an overriding and continuing duty upon a hospital as an organisation to provide services to its patients. The hospital cannot be a mere custodial institution to provide a place where medical personnel meet and treat patients. (see Ellis vs. Wallsend District Hospital [1989] 17 NSWLR 553)." The COA, held out that "the Ophthalmologist was a doctor of the Hospital, the patient paid the required fees to the Hospital. He did not have a choice as to the Anaesthetist. Further, the Hospital provided all the facilities, drugs and nurses for the operation." Based on these factors, the COA affirmed the High Court's finding of vicarious liability on the part of the Hospital.

On the issue of damages, the COA agreed with the award of damages by the High Court which had taken into account the patient's severe pain, loss of vision, nervous shock and distress, embarrassment and humiliation, deprivation of ordinary life experience, and lost promotion prospects.

The Federal Court (FC) made a few landmark decisions in the Eye appeal.
a. Preliminary objection – request for a retrial due to a non-speaking judgment
Objections were raised by the doctors and the hospital, as the lower court had given a non-speaking judgment, wherein the judge merely makes a finding without assigning reasons or clarifying why he was influenced to do so. They appealed for a retrial. The FC whilst disapproving such practices of giving non-speaking judgments, held that this does not automatically warrant a retrial, "because the party seeking the retrial has the burden of proving that there was some substantial wrong or miscarriage of justice by the trial court before such relief can be granted." The FC took cognisance of the fact that the adverse events occurred in 1999, the trial only began eight years later in 2007, and concluded in 2010. The trial lasted 23 days and involved 10 witnesses. The FC held that to order a re-trial after two decades would be unfair and would be unduly prejudicial to the party bearing the burden of proof. In such cases, where a non-speaking judgment is given, the courts held that "the appellate courts have a duty to make their own findings of fact based on the on the evidence available in the records of appeal."

b. The standard of care in medical negligence
The first question raised, was whether "it is the Bolam test or the test in the Australian case of Rogers v Whittaker [1993] 4 Med LR 79 which should be applied to the standard of care in medical negligence, following, after decision of Federal Court in Foo Fio Na vs. Dr Soo Fook Mun & Anor [2007] 1 MLJ 593, conflicting decisions of the Court of Appeal of Malaysia, conflicting decisions of the High Court in Malaysia, and the legislative changes in Australia, including the re-introduction there of a modified Bolam test." The Bolam test is effectively a paternalistic "doctors know best" test whereby the courts must accept the views of a responsible body of men skilled in the particular discipline, even if there exists another responsible body of men with a different view. The reasoning for this test was "the courts, not being medically trained, are not equipped to resolve genuine differences of opinion on matters that are beyond their expertise." The Bolam test8 was subsequently qualified in Bolitho vs. City & Hackney Health Authority,[9] which asserted that the expert opinion must be capable of withstanding logical analysis.

The eye appeal was heard along with Zulhasnimar Hasan Basri & Anor vs. Dr Kuppu Velumani P & Ors.[12] The FC clarified the position in Malaysian law, reiterating that "a distinction is to be made between diagnosis and treatment in medicine, and the duty to advise the patient of risks. The former is not within the expertise of the courts and thus cannot be resolved by the courts, whereas the latter is an issue of fact that the courts are able to determine." Thus the Bolam test, qualified in Bolitho still applies to the standard of care in medical diagnosis and treatment, while the Rogers test as adopted in in Foo Fio Na vs. Dr. Soo Fook Mun applies only to the duty of disclosure of risks associated with a procedure.[13] In the eye appeal the doctors were found to be negligent on both accounts, the standard of care for diagnosis and treatment, in addition to the standard of care for disclosure of risks and advice.

c. Aggravated damages in Medical Negligence
The second question of law posed by the three parties to the to the Federal Court was "Whether aggravating factors should be compensated for as general damages, therefore rendering a separate award of aggravated damages unnecessary, as decided by the English Court of Appeal in Richardson vs. Howie [2004][14] and explained in Michael Jones' Medical Negligence textbook."[15] In Malaysia, aggravated damages have previously been awarded as a separate head of damage in Mohd Ridzwan bin Abdul Razak vs. Asmah bt Hj Mohd Nor, in a sexual harassment case.16 The FC thus held that there was no reason to exclude this kind of damages from

being awarded in medical negligence cases which involve real injury to a person's body. Further, in defamation cases aggravated damages are lumped along with general damages and not awarded as a separate category as seen in Lim Guan Eng vs. New Straits Times Press (M) Bhd,[17] Ling Wah Press (M) Sdn Bhd & Ors vs. Tan Sri Dato Vincent Tan Chee Yioun,[18] Chin Choon v Chua Jui Meng [2005].[19] Hence, the FC concurred with the award of an unprecedented sum of MYR1,000,000 as aggravated damages. The aggravated awards are not compensatory but punitive and it looks ominous that the courts will not shy away from imposing aggravated damages in future, and these sums may keep increasing in tandem with other damages which have been extrapolating over the years.

d. Can a Hospital delegate its duty of care?

The third question of law posed by the Hospital was, "Where the doctors are qualified professionals in a private hospital and working as independent contractors by virtue of a contract between the private hospital and the doctors, can the private hospital be held vicariously liable for the sole negligence of the doctors?" The FC held that "the doctors were independent contractors and not agents, servants, or employees of the private hospital. As such, the hospital could not be vicariously liable for the doctors' negligence." Nevertheless, the FC held that the hospital was liable for breach of its non-delegable duty regarding the anaesthetic services provided to the patient. In the case of Dr. Kok Choong Seng & Anor vs. Soo Cheng Lin,[20] the FC held that "the doctrine of non-delegable duty of care as expounded by the English Supreme Court in Woodland vs. Swimming Teachers Association and others[21] could apply to private healthcare institutions. However, the court in Dr. Kok Choong Seng held that "the doctrine did not apply to the facts of that case, and the private hospital therein was not liable for the doctor's negligence." This provided the potential grounds to impose this non-delegable duty of care on private hospitals, in an appropriate case. Unlike in Dr. Kok Choong Seng, the FC in the Eye Appeal held that the Woodland test was fulfilled concerning the Anaesthetist's negligence but not the surgeon. As regards the surgeon's negligence, the FC found similarity of facts with those of Dr. Kok Choong Seng. In both cases the diagnosis and treatment of the patient's eye, including the surgery, was made between the patient and the surgeon, and the hospital had merely provided the facilities and services for the operation. Whereas the Anaesthetist was the only one on duty at the hospital on the day of the surgery. He was the one on duty to provide general anaesthesia for all operations at the hospital on that day. The patient was not provided a choice to select the Anaesthetist for his operation, though he had requested for the same Anaesthetist. "The patient had no control over how the hospital chose to provide anaesthetic services, whether by delegation to employees or otherwise; the hospital had delegated to the anaesthetist the responsibility to administer doses to the patient properly; the anaesthetist was negligent in the performance of the duty delegated to him by the hospital." The Eye Appeal is the first case that held that a non-delegable duty of care exists by a private hospital for the medical negligence of independent contractors. The FC took cognisance of the proviso in Woodland to impose liability "only to the extent where it is fair, just and reasonable." The question as to whether private hospitals, will be found to owe a non-delegable duty of care to their patients will continue to be addressed on a case-by-case basis. It is thus clear that private hospitals are not immune from being held liable for a non-delegable duty of care. We probably will see more such awards in future.

DISCUSSION

The evolution of medical negligence law in Malaysia has been significantly shaped by these two landmark cases involving ophthalmic surgery: Rogers vs. Whitaker and Dr Hari Krishnan & Anor vs. Megat Noor Ishak bin Megat Ibrahim & Anor (the "Eye Appeal").[1,4] Rogers v Whitaker (Australia, 1992) shifted the standard for risk disclosure from a doctor-centric to a patient-centric approach, holding that clinicians must inform patients of material risks that a reasonable person in the patient's position would consider significant, regardless of the risk's rarity.1 Malaysia adopted this principle in Foo Fio Na vs. Dr Soo Fook Mun & Anor, but ambiguity persisted as to whether this standard applied solely to risk disclosure or also to diagnosis and treatment.[2] This uncertainty was resolved in the Federal Court's "Eye Appeal" decision (Dr Hari Krishnan & Anor vs. Megat Noor Ishak bin Megat Ibrahim & Anor), which clarified that the Bolam/Bolitho test governs the standard of care in diagnosis and treatment, while the Rogers test applies to the duty of risk disclosure.[4] The Eye Appeal also set new benchmarks: it affirmed that aggravated damages can be awarded separately in cases of real injury, and that private hospitals may be liable for the negligence of independent contractors under the doctrine of non-delegable duty of care, particularly when patients have no choice in their care provider.

CONCLUSION

These cases have collectively enhanced patient autonomy, clarified legal standards for clinicians, and expanded institutional accountability. The evolving jurisprudence underscores the need for Malaysian healthcare professionals to prioritise transparent communication and informed consent, and for institutions to recognise their broader responsibilities in patient care.

REFERENCES

1. Rogers v Whitaker[1993] 4 Med LR (Aust HC).
2. Foo Fio Na v Dr Soo Fook Mun & Anor[2007] 1 MLJ 593 (FC, Malaysia).
3. Naxakis v Western General Hospital[1999] HCA 22; 197 CLR 269; 73 ALJR 782; 162 ALR 540 (Aust HC).
4. Dr Hari Krishnan & Anor v Megat Noor Ishak bin Megat Ibrahim & Anor appeal[2018] 3 CLJ 427 (FC, Malaysia).
5. Leong WH, Chan B. The Eye Appeal. Bar Council Malaysia CPD [Internet]. 2018 [cited 2025 May 2]. Available from: https://cpd.malaysianbar.org.my/the-eye-appeal
6. Rogers v Whitaker (1992) 175 CLR 479; 109 ALR 625; 67 ALJR 47; Reports 81-189 (Aust HC).
7. Rogers v Whitaker (1992) 109 ALR 625 at 628 (Aust HC).
8. Bolam v Friern Hospital Management Committee[1957] 1 WLR 582 (QB, UK).
9. Bolitho v City and Hackney Health Authority[1996] 4 All ER 771 (HL, UK).

10. Rogers v Whitaker (1992) 109 ALR 625 at 628 (Aust HC).
11. Dr Hari Krishnan & Anor v Megat Noor Ishak bin Megat Ibrahim & Anor and another appeal[2018] 3 MLJ 281;[2017] MLJU 2093;[2015] MLJU 1857 (FC, Malaysia).
12. Zulhasnimar Hasan Basri & Anor v Dr Kuppu Velumani P & Ors[2017] 8 CLJ 605 (FC, Malaysia).
13. Foo Fio Na v Dr Soo Fook Mun and Assunta Hospital[2007] 1 MLJ 593 (FC, Malaysia).
14. Richardson v Howie[2004] EWCA Civ 1127 (EWCA, UK).
15. Jones M. Medical Negligence. 4th ed. London: Sweet & Maxwell; 2008. para 12-011.
16. Mohd Ridzwan bin Abdul Razak v Asmah bt Hj Mohd Nor[2016] 4 MLJ 282 (FC, Malaysia).
17. Lim Guan Eng v New Straits Times Press (M) Bhd[2017] 9 MLJ 622 (FC, Malaysia).
18. Ling Wah Press (M) Sdn Bhd & Ors v Tan Sri Dato Vincent Tan Chee Yioun[2000] 4 MLJ 77 (FC, Malaysia).
19. Chin Choon v Chua Jui Meng[2005] 2 CLJ 569 (FC, Malaysia).
20. Kok Choong Seng & Anor v Soo Cheng Lin[2017] 10 CLJ 529 (FC, Malaysia).
21. Woodland v Swimming Teachers Association and others[2014] AC 537 (UKSC).

ORIGINAL ARTICLE

Mortality and risk factors in post-operative pancreatic fistula following pancreatoduodenectomy: A single centre experience

Ong Guang Hong, MMed (Surg), Leow Voon Meng, MMed (Surg), Manisekar K Subramaniam, FRCS

Hepatobiliary Unit, Department of General Surgery, Hospital Sultanah Bahiyah, Alor Setar, Malaysia

INTRODUCTION

Pancreatoduodenectomy (PD) is one of the most complex surgical procedures in abdominal surgical procedure. It is offered to patients with benign pathology and malignancy. The mortality has been quite significant despite it had reduced from 25%-30% in 1980s to 2%-4% in the modern era.[1] However, perioperative morbidity remained high (30%-50%).[2]

The most dreaded complication of PD is post-operative pancreatic fistula (POPF). The International Study Group of Pancreatic Surgery (ISGPS) in 2005 further the POPF into grade A, B and C.[3] In 2016, The ISGPS updated the POPF Grade A into biochemical leak, POPF Grade B and Grade C into clinically relevant POPF (CR-POPF).[4]

The reported overall POPF rate of studies in 2001 – 2019 from as 7% - 90%. The reported pooled overall POPF is 27% (CI-95%: 23%-30%). The CR-POPF reported in the same systemic review is 7% - 45%, with the pooled CR-POPF rate of 19% (CI-95%: 17%-22%) The wide variation in the reported rates of overall POPF and CR-POPF could be due to different ethnic or societal factors (such as Body Mass Index), in surgical outcomes or some form of confounding or publication bias.[5]

The complications that follow POPF consists of intra-abdominal infection, sepsis, post-pancreatectomy haemorrhage, pseudo-aneurysm and even death. Multiple strategies have been developed to prevent POPF, these are including methods of pancreatic anastomosis, external or internal pancreatic stents, somatostatins analogues, fibrin glue or topical agent to reinforce pancreatic anastomosis.[6] Thus, it is essential to predict and take action accordingly to the risk.

Risk factors that have been identified in the literature include, male sex, Body Mass Index (BMI) >25kg/m^2, pancreatic duct diameter (PDD) < 3mm, soft pancreatic texture and blood transfusion. Pancreatic adenocarcinoma, diabetes mellitus(DM), vascular resection and preoperative chemotherapy are found to be protective factors against POPF.[7] There is tremendous effort of building risk scoring system to prognosticate as evidenced by at least 10 reported POPF scoring systems by Adamu et al.[8] Among these, the scoring systems created by Callery et al. and Mungroop et al. respectively are deemed the easiest to be implemented and have the best prognosticate ability.[8–10] The Callery et al. uses the gland texture, pathology, PDD, and intraoperative blood loss in the Fistula Risk Score.[9] The Mungroop et al. created an online calculator named as Alternative Fistula Risk Score (a-FRS) which is available at pancreascalculator.com, using pancreatic texture, BMI and PDD.[10] The Ansorge et al 2012 by using pancreatic consistency (pancreatic texture) and PDD, divided the patients into high (two risk factors), intermediate group (one risk factors), and low risk group (no risk factor). [11]

An ideal score can be helpful in prognosticate and thus implementing different strategies. Total pancreatectomy can eradicate the possibility of POPF with the risk of a brittle DM.[12] External pancreatic stent was suggested to help in preventing POPF, and to be considered in the high POPF-risk patient.[13] Expensive medication such as pasireotide can be offered in the high risk patients.[2] On the other hand, in the low risk patients, abdominal drain can be removed earlier.[14] By risk stratification, individualized management can be provided.

Until now, there is no publication on incidence of POPF in Malaysia. This study provides an opportunity to assess and examine the correlation of the risk factor and POPF after PD.

MATERIALS AND METHODS

This is a single centre retrospective observational study which conducted in Hospital Sultanah Bahiyah, Alor Setar, Kedah, Malaysia. This is a government own, tertiary referral centre for Hepato-pancreato-biliary surgery cases in Northern Malaysia. The data of patients in the IT system of Hospital Sultanah Bahiyah that were classified as Pancreatoduodenectomy done during the 5 years period from 01 Jan 2017 to 31 Dec 2021 are used for analysis.

Inclusion Criteria:

1. All patients who were recorded as pancreatoduodenectomy done

Exclusion Criteria:

1. Procedure abandoned.
2. Patient had undergone total pancreatectomy
3. Patient with incomplete information of the procedure

All data collected are confidential, no personal identification data of patient will be collected. Data presented will not identify individuals.

This article was accepted: 10 March 2025
Corresponding Author: Voon Meng Leow
Email: lvmeng74@gmail.com

Table I: Baseline Demographic and Clinical Characteristics of Patient

	N (%)	Mean ± SD	Range
Demographic			
Age (years)	111 (100)	59.3 ± 12.747	19 - 82
BMI (kg/m²)*	53 (47.7)	22.61 ± 3.944	15 - 33
Gender			
Male	60 (54.1)		
Female	51 (45.9)		
Ethnic			
Malay	66 (59.5)		
Chinese	34 (30.6)		
Indian	10 (9.0)		
Other	1 (0.9)		
Diabetes Mellitus			
Yes	47 (42.3)		
No	64 (57.7)		
Comorbid			
Yes	77 (69.4)		
No	34 (30.6)		
PRE- OPERATIVE			
Pre- op Alb (g/L)	111 (100)	32.153 ± 5.3107	16 - 45
Pre- op PDD - image (cm)*	108 (97.3)	0.43 ± 0.291	0.11 - 1.5
Pre- op Biliary drainage			
No	48 (43.2)		
ERCP	51 (45.9)		
PTBD	12 (10.8)		
Neo-adjuvant therapy			
Yes	1 (0.9)		
No	110 (99.1)		
INTRA- OPERATIVE			
Operating time (hour:min:sec)	111 (100)	6:56:32 ± 1:55:23	2:06:00 - 13:15:00
Intra-op PDD (cm)*	15 (13.5)	0.36 ± 0.1298	0.2 - 0.5
Pancreas texture*			
Soft	45 (40.5)		
Hard	32 (28.9)		
Pancreatic anastomosis*			
Dunking PJ	43 (38.7)		
2 layers PJ	66 (59.5)		
3 layers PJ	1 (0.9)		
Vascular resection			
Yes	6 (5.4)		
No	105 (94.6)		
Intra-op transfusion*			
Yes	50 (45.0)		
No	55 (49.5)		
Intra-op blood loss (ml)*	45 (40.5)	889.56 ± 443.8	280 - 2500
POST- OPERATIVE			
HPE			
Pancreatic adenocarcinoma	28 (25.2)		
Pancreatitis	6 (5.4)		
Ampullary, duodenal, cystic, islet cell and other diagnosis	77 (69.4)		

*Missing data for these risk factors

Work Process

1. All patients that were classified as PD case are listed down using operating notes record.
2. The patient's information are retrieved from the IT system with data entry from.
3. The data are being analysed.
4. The workflow are as illustrated in the Diagram 1.

Variables

1. Pre-operative data: Age, ethnic, sex, BMI, PDD (on the latest preoperative imaging, measured at the location of planned pancreatic transaction by single observer with the guidance of radiologist), DM, comorbidity, preoperative biliary drainage (none, endoscopic biliary drainage, percutaneous transhepatic biliary drainage), neo-adjuvant therapy.
2. Intraoperative data: pancreatic texture (assess intraoperatively, soft or hard), vascular resection, intraoperative PDD, operation date, intraoperative blood loss, present of intra-op blood transfusion, pancreatic anastomosis (2 layers duct-to mucosa PJ [including Blumgart's and modified Blumgart's], dunking PJ, 3 layers duct-to-mucosa PJ, Pancreatogastrostomy), operating time
3. Postoperative data: Serum and drain amylase POD3 and day 5, drain removal date, date of discharge, image

Table II: Post-op drain and serum amylase, Post-op outcome, 30-day mortality and POPF result

	N (%)	Mean ± SD
D3 drain amylase (iu/L)	111 (100)	1079.33 ± 1987.46
D5 drain amylase (iu/L)	74 (66.7)	1600.12 ± 6808.7
Serum amylase D3 (iu/L)	74 (66.7)	143.66 ± 232.679
Serum amylase D5 (iu/L)	37 (33.3)	47.27 ± 35.754
	Number (N)	**Percent (%)**
Drainage day		
>21 Days	39	35.1
≤21Days	72	64.9
IR drainage post-op		
Yes	3	2.7
No	108	97.3
IR angiography post-op		
Yes	1	0.9
No	110	99.1
Re-operation*		
Yes	20	18
Re-operation for POPF	14	12.6
Re-operation for other	6	5.4
No	91	82
Organ failure post-op		
Yes	16	14.4
No	95	85.6
30 days mortality**		
Yes	10	9
Mortality due to POPF	6	5.4
Mortality for other cause	4	3.6
No	101	91
Clinically Relevant POPF		
No	68	61.3
No POPF	43	38.7
Biochemical Leak	25	22.5
Yes	43	38.7
POPF Grade B	21	18.9
POPF Grade C	22	19.8

* Six (6) cases of re-operation for non POPF leak: 1 case for drain site bleeding, 1 for strangulated incisional hernia, 1 for bleeding cystic plate, 2 for biliary leak, 1 for chyle leak.

**Three (3) cases of patient died from pneumonia, 1 case for acute coronary event.

guided management for POPF, angiographic procedures for POPF, re-operation, POPF related organ failure, 30days mortality, histopathology examination

4. POPF grading is according to ISGPS 2016.[4]

a. Biochemical leak: Amylase > 3 times upper limit institutional normal serum amylase value (300iu/ml)
b. Grade B Pancreatic Fistula: Persistent drainage > 3weeks, Percutaneous or endoscopic specific interventions for collections, angiographic procedures for POPF related bleeding
c. Grade C Pancreatic Fistula: Reoperation, organ failure and death

Surgical Procedure

The surgery is performed by HPB surgery consultants or surgeons who are under HPB surgery subspecialty training with the supervision of HPB consultants. The procedure is started with modified Makuuchi, rooftop or mid-line skin incision. The exploratory laparotomy is performed to decide on whether there is peritoneal metastasis or liver metastasis. Pancreaticoduodenectomy is then performed in a standard way, usually by Classical Kausch-Whipple Procedure. Vascular resection such as portal vein resection is performed as deemed necessary by the operating surgeon. The pancreatic anastomosis is then selected according to the surgeon preference. All patients receive pancreatic stent insertion. Somatostatin analogue is given prior to pancreas transaction and continued for 5days. Enteral feeding is started at Day 2 if there is no contraindication.

DATA ANALYSIS

The statistical analysis will be done using SPSS ™ For Windows Version 20.0. The information collected are analysed using Student T- test for continuous data, Chi-square test for categorical data. The analysis of risk factors is performed using Multivariate logistic regression. P value of less than 0.05 is considered significant.

RESULTS

There were 150 patients listed for pancreaticoduodenectomy in Hospital Sultanah Bahiyah, Alor Setar from 01/01/2017 till 31/12/2021. Among all of them, 111 (74%) patients underwent PD and completed the procedure. Among the 39 patients whom did not proceed with the PD, 15 of them had

Table III: Comparison of the CR-POPF incidence by crisk factors

	CR-POPF (N, %)			p-value
	No	Yes	Total (N, %)	
Ethnic				0.593
Malay	41 (62.0)	24 (37.9)	65 (58.5)	
Chinese	19 (55.9)	15 (44.1)	34 (30.6)	
Others	8 (72.7)	3 (27.3)	11 (9.9)	
Gender				0.924
Female	31 (60.8)	20 (39.2)	51 (45.9)	
Male	37 (61.7)	23 (38.3)	60 (54.1)	
DM			0.634	
No	38 (59.4)	26 (40.6)	64 (57.7)	
Yes	30 (63.8)	17 (36.2)	47 (42.3)	
Comorbid other than DM		*0.029		
No	26 (76.5)	8 (23.5)	34 (30.6)	
Yes	42 (54.5)	35 (45.5)	77 (69.4)	
Pre-op biliary drainage		0.634		
ERCP	33 (64.7)	18 (35.3)	51 (45.9)	
PTBD	8 (66.7)	4 (43.8)	12 (10.8)	
None	27 (56.3)	21 (43.8)	48 (43.2)	
Neo-adjuvant therapy		0.613		
No	67 (60.9)	43 (39.1)	110(99.1)	
Yes	1 (61.3)	0 (0.0)	1(0.9)	
Pancreas texture			0.618	
Soft	29 (64.4)	16 (35.6)	45 (40.5)	
Hard	20 (62.5)	12 (37.5)	32 (28.8)	
Pancreatic anastomosis		0.663		
Dunking PJ	25 (58.1)	18 (41.9)	43 (38.7)	
2 Layer PJ	41 (62.1)	25 (37.9)	66 (59.5)	
3 Layer PJ	1 (100.0)	0 (0.0)	1(0.9)	
Vascular resection			0.402	
No	63 (60.0)	42 (40.0)	105(94.6)	
Yes	5 (83.3)	1 (16.7)	6 (5.4)	
Intra-op transfusion			0.481	
No	34 (61.8)	21 (38.2)	55 (49.5)	
Yes	29 (58.0)	21 (42.0)	50 (45.0)	
HPE				0.18
PDAC and Pancreatitis	24(70.6)	10 (29.4)	34 (30.6)	
Ampullary, duodenal, cystic, islet cell and other diagnosis	44(57.1)	33 (42.9)	77 (69.4)	
	(Mean ± SD)	(Mean ± SD)		
Age (Years)	57.40 ± 13.684	62.30 ± 10.571	111 (100)	*0.048
BMI (kg/m^2)	22.48 ± 4.032	22.84 ± 3.878	53 (47.7)	0.753
Pre-op Alb (g/L)	32.794 ± 4.7176	31.140 ± 6.0537	111 (100)	0.11
PDD- image (cm)	0.47±0.297	0.37 ± 0.277	108(97.3)	0.099
Operating time (hour:min:sec)	6:53:13 ± 1:53:10	7:01:49 ± 1:59:59	111 (100)	0.704
Intra-op blood loss (ml)	851.67 ± 370.108	965.33 ± 569.484	45 (40.5)	0.424

*p value < 0.05, statistically significant.

bypass surgery, 18 patients were explored and biopsy only, 3 patients did not proceed for operation, 2 patients had radical cholecystectomy done and 1 patient had a total pancreatectomy done.

Of the 111 patients who had PD, their baseline characteristics are as in Table I. The postoperative drain and serum amylase concentration is reported in Table II. The mean of day 3 drain amylase is 1079± 1987 iu/L. The cut off value of D3 drain amylase level of 300iu/L is used to defined biochemical leak. The patients are graded according to ISGPS 2005 and 2016. The result is presented in the Table III. The number of patients with POPF Grade B is 21 (18.9%) and POPF Grade C is 22 (19.8%), make up total of 43 patients with CR-POPF (38.7%). The outcomes of operation are reported in Table II. There are 39 patients have drain >21days. Three (2.7%) patients have interventional radiological drainage, and one (0.9%) has angiographic intervention. There are 20 (18.0%) cases of re-operations, 14 (12.6%) of them are for POPF, 6 are for other causes including 2 for biliary leak, one for drain site bleeding, one for strangulated incisional hernia, one for bleeding cystic plate, and one for chyle leak. 16 (14.4%) developed organ failure. Ten (9.0%) cases of 30 days mortality, 6 (5.4%) of them due to POPF, 3 died for pneumonia, and one for acute coronary event.

The comparative statistical analysis is reported in Table III. The ethnicity Indian and others are combined into others for analysis. The ethnic, sex, DM, pre-op biliary drainage, neo-adjuvant therapy, pancreas texture, pancreatic anastomosis, vascular resection, intraoperative transfusion and HPE are not found have statistical significance. The age and

Table IV: Logistic Regression Analysis for Risk Factor

Univariate Logistic Regression						
		Coeff.	OR	95% C.I. for OR		p-value
				Lower	Upper	
Age		0.033	1.033	1.000	1.033	0.052
Gender Male		-0.037	0.964	0.448	2.073	0.924
DM	-0.188	0.828	0.381	1.800	0.624	
Comorbid		0.996	2.708	1.090	6.732	0.032*
Preop Albumin		-0.060	0.934	0.854	1.022	0.137
Pre-op biliary drainage	No					0.635
	ERCP	-0.355	0.701	0.312	1.575	0.390
	PTBD	-0.442	0.643	0.170	2.428	0.515
PDD Pre-op image		-1.229	0.293	0.067	1.282	0.293
Pancreatic Anastomosis	Duct-to-mucosa PJ**	-0.190	0.827	0.378	1.808	0.634
Intra-op transfusion		0.159	1.172	0.536	2.562	0.690
HPE (Pancreatic Ca or pancreatitis)		-0.588	0.556	0.234	1.319	0.183
		Multivariate Logistic Regression				
		Coeff.	OR	95% C.I. for OR		p-value
				Lower	Upper	
Age		0.024	1.024	0.982	1.069	0.269
Gender Male		-0.051	0.951	0.365	2.473	0.917
DM		-0.411	0.663	0.255	1.727	0.400
Comorbid		1.257	3.514	1.085	11.385	0.036*
Preop Albumin		-0.068	0.934	0.854	1.022	0.137
Pre-op biliary drainage	No					0.481
	ERCP	-0.538	0.584	0.219	1.557	0.282
	PTBD	-0.679	0.507	0.109	2.354	0.386
PDD Pre-op image		-1.312	0.269	0.048	1.503	0.135
Pancreatic Anastomosis	Duct-to-mucosa PJ**	-0.260	0.771	0.303	1.964	0.585
Intra-op transfusion		0.279	1.322	0.517	3.385	0.560
HPE (Pancreatic Ca or pancreatitis)		-0.522	0.593	0.215	1.636	0.313
Nagelkerke R Square	0.218					
Hosmer and Lemeshow Test	0.385					

* p-value < 0.05
** 2-layer and 3-layer duct-to-mucosa group are combined as duct-to-mucosa PJ group

comorbidities other than DM have statistically significant difference. Between CR-POPF and no CR-POPF group, the student t-test analysis reported no different of means in BMI, Pre-op albumin, PDD (image), operating time and intra-op blood loss volume.

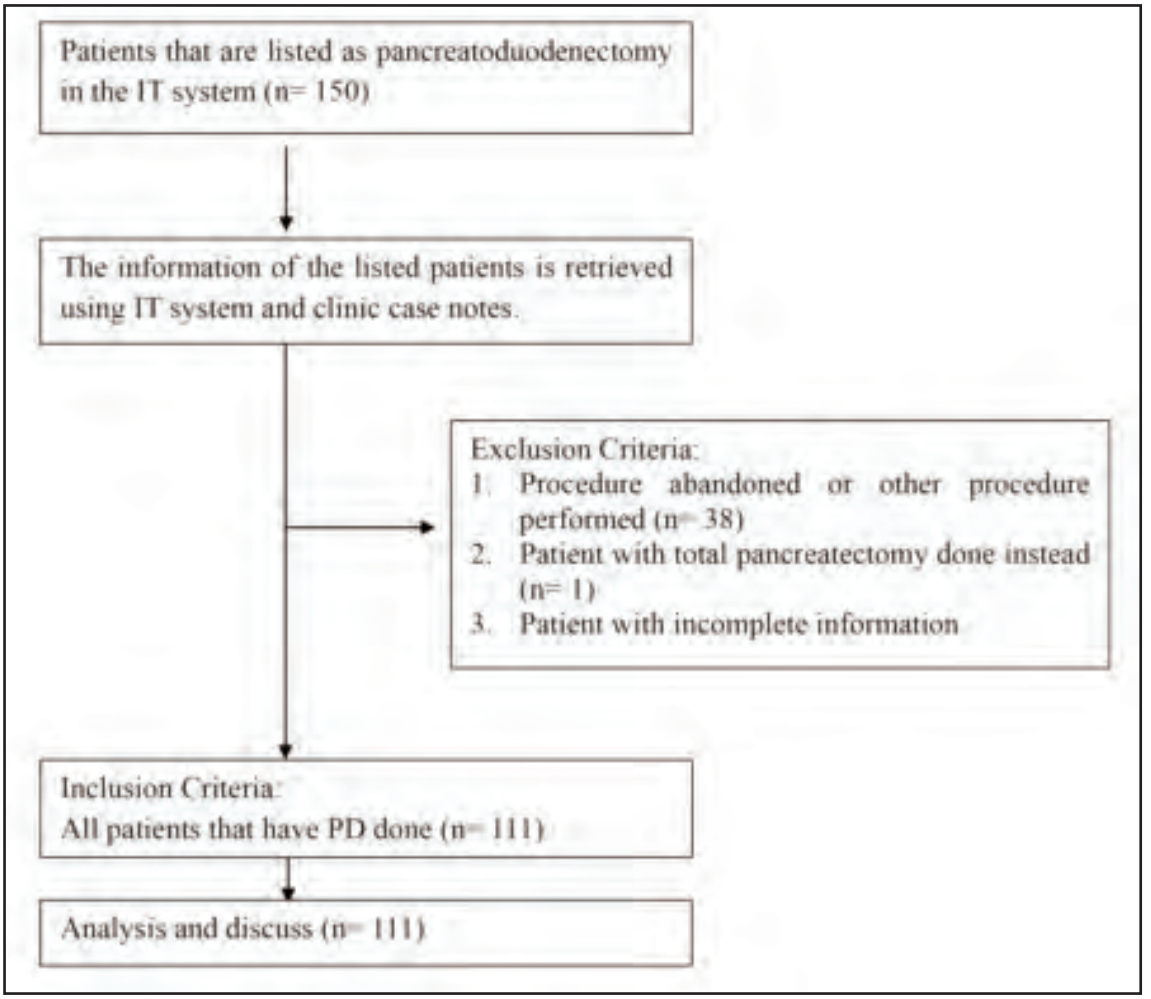
The Table IV shows the results of logistic regression analysis. The 2-layers and 3-layers duct-to-mucosa PJ are combined as duct-to-mucosa group in pancreatic anastomosis variable analysis. The Univariate Logistic Regression are conducted, and comorbid other than DM is found to be significant predicting factor. The OR is 2.708 (CI 95%: 1.090-6.732, p-value 0.032) All the other factors are not statistically significant. Multivariate analysis is performed, and comorbidity other than DM found to be a significant risk factor with OR 3.529 (CI 95%: 1.092-11.404, p-value = 0.035). The Nagelkerke pseudo R^2 = 0.223, Hosmer and Lemeshow test p-value =0.228. This means that the model is not strong, but it is consider fit as p-value of the Hosmer & Lemeshow test is >0.05.

DISCUSSION

The overall CR-POPF incidence in this study is 38.7%. The reported pooled CR-POPF rate in systematic review by Sivesh K. Kamarajah et al. is 19% (CI 95%: 17-22).5 The CR-POPF for all the studies included in the systemic review is between 7% - 45%, the reported wide range may be explained by the different population and social economy factor of the centres.[5] Our rate of CR-POPF of the local centre is within the reported range. The 30days mortality related to POPF is 5.4%. The hospital death reported by an Amsterdam paper to be 1% - 16%.[15] The 30days mortality of our centre is within the range.

The patients who have none DM morbidity experience a significant higher CR-POPF rate. This association is further supported by logistic regression analysis. Our study did not explore the specific co-morbidities as the FRS and a-FRS does not include pre-morbid disease. A 2020 systematic review by Sivesh et al. suggested a link between cardiovascular disease and CR-POPF, where pre-operative chronic pulmonary obstructive pulmonary disease and pre-operative hypertension were correlated with all POPF which including biochemical leak.[5] Age was found to be significantly different. The mean age of the group of patients who developed CR-POPF is older. The similar finding is found in the Zhao Z. et al. in 2023, where group older than 62.5-year-old has higher POPF rate (14.9%) compare to the younger group (6.9%).[16] It was proposed that senile patients carried more comorbidity, and suboptimal fitness which cause poor healing of the anastomosis. The univariate and multivariate logistic

METHODOLOGY FLOW CHART



regression found that only none DM comorbidity can predict the CR-POPF, with the OR 3.514 (CI 95%: 1.085-11.385, p-value = 0.035). The age is not a predictor. More efforts should be invested into studying the comorbidity in the population.

Most of the studies show that male sex is a risk factor for POPF.[5,7] But, Zhang JY et al. shows the different result which is similar to us.[17] PDD on the pre-operative imaging is a commonly studied factor for PD. Many scoring systems use PDD as one of the predicting factors.[5,7,9,10] Small duct size is associated with normal pancreas with a preserved endocrine function and usually represents difficult anastomosis. Small duct size is also correlated with more post-operative acute pancreatitis.[18,19] In our study, PDD is slightly smaller in CR-POPF group but it was not significant statistically (0.37cm vs 0.47cm, p = 0.099).

CR-POPF rate does not differ in the DM group and non-diabetic group in this study, even though the previous study found that DM is protective for POPF.[20] Pancreatic texture is a well-established predicting factor for POPF. This is confirmed by the multiple studies.[5,7–10,17,20] Histological research explains that patients with POPF usually has more pancreatic fat with less pancreatic fibrosis and low blood vessel density on the specimen.[21] Scoring system a-FRS included the BMI in the mathematic model.[10] Moreover, systematic review by Kamarajah et al. in year 2020 shows that BMI (> 25kg/m^2) is a significant risk factor.[5] However, our data cannot draw conclusion regarding BMI. Pre-op serum albumin is known to be correlated with serious complications follow PD, as reported in Narongsak et al. 2019.[22] But, Kamarajah et al. and Zhang B et al. showed the different conclusion.[5,7] Pre-operative biliary drainage (PBD) was reported to reduce CR-POPF or overall postoperative complications, although it is not shown in our result.[23,24]

Intra-operative blood loss was reported to increase risk of POPF, as reported by Callery et al.[9] It was hypothesized that sudden blood loss may lead to ischaemia and poor healing of the pancreatic anastomosis. This maybe compromised by the tissue oedema resulted from the aggressive fluid resuscitation in response to the blood loss. Intraoperative blood transfusion is related to large volume blood loss and is associated with POPF.[5,7]

Pancreaticogastrostomy was reported to be superior to pancreatojejunostomy in term of POPF rate.[25] However, recent meta-analysis shows that POPF rate in PG and PJ group are almost similar.[5] In our centre, the preferred anastomotic method is PJ, where no PG was performed. Methods that are

used including dunking technique and duct-to-mucosa PJ and there is no difference of POPF. Operative time does not differ significantly in this study which is in line with the study by Ke Z. et al.[26]

Callery et al. showed that the pancreatic adenocarcinoma or pancreatitis are associated with lower POPF rate, while other pathology is higher. The maybe theorized as other pathology are commonly related with a soft pancreas.[9] Our analysis revealed a lower but not statistically significant CR-POPF rate in patients with pancreatic ductal adenocarcinoma or pancreatitis.

The pathophysiology of POPF remained complex. The formation of POPF may not be just due to mechanical disruption of the anastomosis. Potential mechanism suggests the role of postoperative pancreatitis in the development of POPF.[27] Intraoperative amylase concentration in immediate intraoperatively derived fluid is highly predictive of the POPF. The local activation of pancreatic enzyme may lead to postoperative pancreatitis, and finally POPF.[28] Thus, the strategy should include methods to strengthen the pancreatic anastomosis, and to prevent the postoperative pancreatitis. Few methods have been proposed included the use of trypsin inhibitors (ulinastatin) and intravenous hydrocortisone.[29–31]

LIMITATIONS

This paper acknowledges certain limitations. First, the retrospective nature of the investigation introduces the possibility of selection bias. Second, missing data limit the analysis of potential risk factors. Third, the design with a small sample size may be insufficiently power to detect significant result. Finally, the variability in surgeon competency and experience could potentially biased the complications rate.

CONCLUSION

The CR-POPF rate is 38.7% and 30-days mortality associated with CR-POPF is 5.4%. Age and percentage of non-DM comorbidity are significantly higher in CR-POPF group patient compared to no CR-POPF group, where logistic regression identify none DM comorbidity as the sole predicting factor.

ACKNOWLEDGEMENT

We would like to thank the Director General of Health Malaysia for his permission to publish this article.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest to disclose.

REFERENCES

1. Narayanan S, Martin AN, Turrentine FE, Bauer TW, Adams RB, Zaydfudim VM. Mortality after pancreaticoduodenectomy: assessing early and late causes of patient death. Journal of Surgical Research 2018; 231: 304-8.
2. Allen PJ, Gönen M, Brennan MF, Bucknor AA, Robinson LM, Pappas MM, et al. Pasireotide for Postoperative Pancreatic Fistula. New England Journal of Medicine 2014; 370(21): 2014-22.
3. Bassi C, Dervenis C, Butturini G, Fingerhut A, Yeo C, Izbicki J, et al. Postoperative pancreatic fistula: An international study group (ISGPF) definition. Surgery 2005; 138(1): 8-13.
4. Bassi C, Marchegiani G, Dervenis C, Sarr M, Abu Hilal M, Adham M, et al. The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 Years After. Vol. 161, Surgery (United States). Elsevier Inc; 2017: 584-91.
5. Kamarajah SK, Bundred JR, Lin A, Halle-Smith J, Pande R, Sutcliffe R, et al. Systematic review and meta-analysis of factors associated with post-operative pancreatic fistula following pancreatoduodenectomy. Vol. 91, ANZ Journal of Surgery. Blackwell Publishing; 2021: 810-21.
6. Pedrazzoli S. Pancreatoduodenectomy (PD) and postoperative pancreatic fistula (POPF). Vol. 96, Medicine (United States). Lippincott Williams and Wilkins; 2017.
7. Zhang B, Yuan Q, Li S, Xu Z, Chen X, Li L, et al. Risk factors of clinically relevant postoperative pancreatic fistula after pancreaticoduodenectomy: A systematic review and meta-analysis. Vol. 101, Medicine (United States). Lippincott Williams and Wilkins; 2022. p. E29757.
8. Adamu M, Plodeck V, Adam C, Roehnert A, Welsch T, Weitz J, et al. Predicting postoperative pancreatic fistula in pancreatic head resections: which score fits all? Langenbecks Arch Surg 2022; 407(1): 175–88.
9. Callery MP, Pratt WB, Kent TS, Chaikof EL, Vollmer CM. A prospectively validated clinical risk score accurately predicts pancreatic fistula after pancreatoduodenectomy. J Am Coll Surg 2013; 216(1): 1–14.
10. Mungroop TH, Van Rijssen LB, Van Klaveren D, Smits FJ, Van Woerden V, Linnemann RJ, et al. Alternative Fistula Risk Score for Pancreatoduodenectomy (α-FRS): Design and International External Validation. Ann Surg. 2019; 269(5): 937-43.
11. Ansorge C, Strömmer L, Andrén-Sandberg Å, Lundell L, Herrington MK, Segersvärd R. Structured intraoperative assessment of pancreatic gland characteristics in predicting complications after pancreaticoduodenectomy. British Journal of Surgery 2012; 99(8): 1076-82.
12. Yang DJ, Xiong JJ, Liu XT, Li J, Siriwardena KMDL, Hu WM. Total pancreatectomy compared with pancreaticoduodenectomy: A systematic review and meta-analysis. Vol. 11, Cancer Management and Research. Dove Medical Press Ltd; 2019. p. 3899-908.
13. Guo C, Xie B, Guo D. Does pancreatic duct stent placement lead to decreased postoperative pancreatic fistula rates after pancreaticoduodenectomy? A meta-analysis. Vol. 103, International Journal of Surgery. Elsevier Ltd; 2022.
14. Wang YC, Szatmary P, Zhu JQ, Xiong JJ, Huang W, Gomatos I, et al. Prophylactic intra-peritoneal drain placement following pancreaticoduodenectomy: A systematic review and metaanalysis. Vol. 21, World Journal of Gastroenterology. WJG Press; 2015: 2510-21.
15. Gouma DJ, Van Geenen RCI, Van Gulik TM, De Haan RJ, De Wit LT, Busch ORC, et al. Rates of Complications and Death After Pancreaticoduodenectomy: Risk Factors and the Impact of Hospital Volume. Vol. 232, ANNALS OF SURGERY.
16. Zhao Z, Zhou S, Tang Y, Zhou L, Ji H, Tang Z, et al. Impact of age on short-term outcomes after pancreaticoduodenectomy: A retrospective case-control study of 260 patients. Front Surg 2023; 10.
17. Zhang JY, Huang J, Zhao SY, Liu X, Xiong ZC, Yang ZY. Risk factors and a new prediction model for pancreatic fistula after pancreaticoduodenectomy. Risk Manag Healthc Policy 2021; 14: 1897-906.

18. Schuh F, Mihaljevic AL, Probst P, Trudeau MT, Müller PC, Marchegiani G, et al. A Simple Classification of Pancreatic Duct Size and Texture Predicts Postoperative Pancreatic Fistula: A classification of the International Study Group of Pancreatic Surgery. Ann Surg 2023; 277(3): E597-608.
19. Loos M, Strobel O, Dietrich M, Mehrabi A, Ramouz A, Al-Saeedi M, et al. Hyperamylasemia and acute pancreatitis after pancreatoduodenectomy: Two different entities. Surgery (United States). 2021; 169(2): 369–76.
20. Lin JW, Cameron JL, Yeo CJ, Riall TS, Lillemoe KD. Risk factors and outcomes in postpancreaticoduodenectomy pancreaticocutaneous fistula. Journal of Gastrointestinal Surgery 2004; 8(8): 951-9.
21. Mathur A, Pitt HA, Marine M, Saxena R, Schmidt CM, Howard TJ, et al. Fatty pancreas: A factor in postoperative pancreatic fistula. Ann Surg 2007; 246(6): 1058–64.
22. Rungsakulkij N, Tangtawee P, Suragul W, Muangkaew P, Mingphruedhi S, Aeesoa S. Correlation of serum albumin and prognostic nutritional index with outcomes following pancreaticoduodenectomy. World J Clin Cases 2019; 7(1): 28-38.
23. Zhu L, Yang Y, Cheng H, Cai Z, Tang N, Mao L, et al. The role of preoperative biliary drainage on postoperative outcome after pancreaticoduodenectomy in patients with obstructive jaundice. Gland Surg 2023; 12(5): 593-608.
24. Shen Z, Zhang J, Zhao S, Zhou Y, Wang W, Shen B. Preoperative biliary drainage of severely obstructive jaundiced patients decreases overall postoperative complications after pancreaticoduodenectomy: A retrospective and propensity score-matched analysis. Pancreatology 2020; 20(3): 529-36.
25. Qin H, Luo L, Zhu Z, Huang J. Pancreaticogastrostomy has advantages over pancreaticojejunostomy on pancreatic fistula after pancreaticoduodenectomy. A meta-analysis of randomized controlled trials. Vol. 36, International Journal of Surgery. Elsevier Ltd; 2016: 18-24.
26. Ke Z, Cui J, Hu N, Yang Z, Chen H, Hu J, et al. Risk factors for postoperative pancreatic fistula: Analysis of 170 consecutive cases of pancreaticoduodenectomy based on the updated ISGPS classification and grading system. Medicine (United States). 2018; 97(35).
27. Nahm CB, Connor SJ, Samra JS, Mittal A. Postoperative pancreatic fistula: A review of traditional and emerging concepts. Vol. 11, Clinical and Experimental Gastroenterology. Dove Medical Press Ltd; 2018: 105-18.
28. de Reuver PR, Gundara J, Hugh TJ, Samra JS, Mittal A. Intra-operative amylase in peri-pancreatic fluid independently predicts for pancreatic fistula post pancreaticoduodectomy. HPB 2016; 18(7): 608-14.
29. Uemura K, Murakami Y, Hayashidani Y, Sudo T, Hashimoto Y, Ohge H, et al. Randomized clinical trial to assess the efficacy of ulinastatin for postoperative pancreatitis following pancreaticoduodenectomy. J Surg Oncol 2008; 98(5): 309-13.
30. Zhang H, Tan C, Wang X, Kang D, Chen Y, Xiong J, et al. Preventive effects of ulinastatin on complications related to pancreaticoduodenectomy A Consort-prospective, randomized, double-blind, placebo-controlled trial. Medicine (United States). 2016; 95(24).
31. Laaninen M, Sand J, Nordback I, Vasama K, Laukkarinen J. Perioperative hydrocortisone reduces major complications after Pancreaticoduodenectomy a randomized controlled trial. Ann Surg 2016; 264(5): 696-702.

ORIGINAL ARTICLE

Factors associated with depression among the community-dwelling elderly in Kudat, Malaysia

Nurliyana Azhar, MD[1], Azulaikha Alias, MD[1], Wan Hafizu Nazrin Wan Mohamad Lotfi, MD[1], Azizan Omar, DrPH[1], Khalid Mokti, DrPH[1], Syed Sharizman Syed Abdul Rahim, DrPH[1], Mohd Saffree Jeffree, MSc[1], Mohd Yusof Ibrahim, MPH[1], Mazlinda Musa, MSN[2], Mohd Rohaizat Hassan, PhD[3], Yugo Shobugawa, PhD[4]

[1]Department of Public Health Medicine, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia, [2]Department of Nursing, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia, [3]Department of Community Health, Faculty of Medicine, Universiti Kebangsaan Malaysia, Cheras, Kuala Lumpur, Malaysia, [4]Department of Active Ageing, Niigata University Graduate, School of Medical and Dental Sciences, Niigata 951-8510, Japan

ABSTRACT

Introduction: The burden of elderly depression is rising with the growing ageing population, particularly in rural areas with limited healthcare access. In Malaysia, 27.8% of the elderly experience depression, with 16.5% of depressive symptoms reported among community-dwelling older adults. This study aimed to determine the prevalence of depression and its associated factors among the elderly in Kudat, a rural area in Sabah.

Materials and Methods: A cross-sectional study was conducted using the Malay version of the Geriatric Depression Scale (M-GDS-14). Logistic regression analyses were used to analyse the relationships between elderly depression and sociodemographic, socioeconomic, physical health, lifestyle, and psychosocial factors.

Results: A total of 310 participants were involved, with a mean age of 69.4 years. Most were of Rungus ethnicity (78.7%) and married (73.9%). About 72% had a household income below RM 1,000; half were retirees, while 43.2% were still working. The prevalence of depression was high at 73.2% (95% CI: 70.7, 75.7). Comorbidities and moderate-to-poor self-rated health were significantly associated with higher odds of depression, with adjusted odds ratios of 1.99 and 2.09, respectively.

Conclusion: The findings highlight the high level of depression among the elderly in Kudat and the significant association with comorbidities and self-rated health status. Public health programs should focus on managing comorbidities and promoting positive self-perceived health to reduce depression in this population.

KEYWORDS:
Elderly, ageing, older adults, depression, factors of depression, community, rural, Sabah, Malaysia

INTRODUCTION

Malaysia is undergoing a demographic shift, with individuals aged 65 and above increasing from 7.0% in 2021 to 7.3% in 2022.[1] As the elderly population grows, so does the prevalence of mental health issues, particularly depression. Elderly individuals often experience overlapping symptoms of depression and dementia, complicating diagnosis and treatment. Although less common than in younger adults, depression in the elderly has severe consequences, including increased suicide risk, impaired functioning, and a more significant physical illness burden.[2]

In Asia, the prevalence of depressive symptoms among the community-dwelling elderly ranges from 3.7% to 36.7%, with many undiagnosed cases.[3] Urbanisation in Malaysia has surged, with urban populations rising from 50.8% in 1990 to 78% in 2021, affecting healthcare resource distribution. The rural population, which has been ageing since 2020, shows a higher depression prevalence (3.6%) compared to urban areas (1.9%).[4]

Rural Malaysia, home to 7.3 million people, faces significant challenges in accessing mental health care, especially in remote regions like Sabah and Sarawak, where almost half of the rural population (3.1 million) resides.[5] Enhancing mental health treatment in these areas is critical. Studies show varying rural-urban differences in the elderly depression, highlighting the need for improved detection and treatment in primary care settings. Healthcare services will be under strain due to the increasing elderly population, as depression results in a 73% increase in costs compared to those without depression.[6] The importance of primary care screening for needs and providing early therapy must be emphasised. It is the first entry point into the healthcare system and offers the best opportunity to detect illness and initiate care.

This study aims to understand the determinants of depressive symptoms in community-dwelling elderly in Kudat, Sabah. While multiple studies have examined elderly depression risk factors in Malaysia, more research is needed in Sabah's rural context due to its unique socioeconomic and cultural characteristics. Kudat, one of Malaysia's poorest districts, provides a critical context for this research, aiming at tailoring targeted public health measures for this vulnerable population. The study aimed to determine the prevalence of depressive symptoms and major depression among the elderly in Kudat. Specific objectives include identifying significant predictors such as comorbidities, lifestyle habits, and self-rated health status.

This article was accepted: 18 March 2025
Corresponding Author: Nurliyana Azhar
Email: drliyana91@gmail.com

The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) defines a major depressive episode (MDE) as having five or more symptoms, including either a depressed mood or anhedonia, present within two weeks.[7] The Geriatric Depression Scale (GDS) is a widely used tool for screening depression in the elderly, with the GDS-15 being particularly popular due to its simplicity and effectiveness.[8] Despite its advantages, the GDS-15 has limitations, particularly in individuals with severe cognitive impairment whose criterion validity is poor.[9]

The Malay version of the GDS-14 (M-GDS-14) was developed to address the effect of cultural and linguistic differences. This version removes the question "Do you prefer staying at home over going out?" due to its non-discriminatory value in local settings. This yielded optimal cut-off points of 7/8 for major depression and 5/6 for all clinically significant depression. The tool demonstrated a Cronbach's alpha score of 0.84 and a test-retest validity of 0.84.[10]

MATERIALS AND METHODS

Research Design and Subjects

This cross-sectional study is part of the Healthy Ageing in Sabah study by Universiti Malaysia Sabah. Data collection occurred at community settings in Kudat, ensuring accessibility for rural participants. Participants included individuals aged 60 and above who consented to participate, understood Malay and passed cognitive screening using the Abbreviated Mental Test (AMT-10) with score of 7 and above. The Kish Formula was used to calculate the adequate sample size for this study based on the prevalence of depression among the Malaysian elderly, which was 27.8% (3), resulting in a calculated sample size of 310.[3]

Data Collection

Data were collected by trained personnel either by interviewer-assisted method or self-administered questionnaires.

Instruments Used

The study employed the Abbreviated Mental Test (AMT) to screen participants for cognitive function prior to inclusion. The AMT-10 is a validated and widely used tool to assess cognitive impairment, consisting of ten questions that evaluate orientation, memory, and attention. Participants were required to achieve a score of 7 or higher to ensure sufficient cognitive capacity for understanding and responding accurately to the study instruments.[11]

Additionally, the study used the Malay-validated GDS-14 to assess depressive symptoms, alongside other instruments evaluating sociodemographic, lifestyle, and health-related factors. The M-GDS-14 utilised a cut-off score of 6 or above to detect clinically significant depression and score of 8 or above to detect severe depression.[10] The study participants were instructed to indicate their feelings by answering 14 questionnaire items with a simple 'yes' or 'no' response.

The sociodemographic aspects included age, sex, religion, ethnicity, and marital status. The socioeconomic aspects encompassed income status and employment status. The lifestyle aspects were assessed by examining smoking habits, alcohol consumption, and the presence of hobbies. Morbidity was evaluated through the presence of hypertension, while self-rated health and loneliness were also measured.

Loneliness was assessed by asking the respondents the frequency with which they lacked companionship, were left out, or were isolated from others, using a 3-point Likert scale coded from 1 'hardly ever' to 3 'often.' The loneliness scale ranged from 3 to 9, with higher scores indicating greater loneliness. This study categorised loneliness as low (scores 3 to 5) or high (scores 6 to 9), consistent with previous research, which also used a cut-off of 6.[12] Additionally, the living arrangements of the respondents were recorded.

Activities of daily living (ADL) status was evaluated based on six parameters: bathing, dressing, going to the toilet, transferring, feeding, and continence. For the physical and mental examination, height, weight, and depression scores were considered. BMI was calculated using the measured height and weight values by dividing body weight (in kilograms) by height (in meters) squared. BMI was categorised per WHO standards for Asians: underweight (<18.5 kg/m²), normal (18.5–22.9 kg/m²) and overweight (≥23 kg/m²).[13]

Self-rated health status was determined based on participants' responses to a single-item question: "How would you rate your current health?" Participants could choose from a 4-point Likert scale: 1 = Very Good, 2 = Good, 3 = Moderate, and 4 = Poor. Responses of Moderate (3) and Poor (4) were grouped together as "moderate to poor" health status. This categorisation reflects participants' subjective perceptions of their health and while this approach provides valuable insights, the subjective nature of self-rated health may limit reproducibility in different contexts.

Ethical Consideration

Ethical approval was obtained from the Research and Ethics Committee of Universiti Malaysia Sabah (UMS), with the approval code of JKEtika 2/23 (6).

Informed written consent was obtained from the study participants before filling out the questionnaire and participant confidentiality was assured. Any participants who had an abnormal status of depression (scoring of 6 and above) during the survey were referred to the nearest health clinic for follow-up.

Statistical Analysis

Raw data collected was recorded into tables in Microsoft Excel 2019 before being cleaned and categorised accordingly. The data were then analysed using the IBM SPSS Version 28.0, the statistical significance of which was a p-value lower than 0.05.

Mean and standard deviation were used for continuous data, while number and percentage were used to describe the distribution of the categorical data. Simple and multiple logistic regression analyses were used to determine the association between independent and dependent variables and the factors associated with depression.

Table I: Sociodemographic data, lifestyle habits, health status and social behaviours of the respondents (n = 310)

Variables	n (%)
Age	69.4 ± 7.00a
60–69	180 (58.1%)
70–79	98 (31.6%)
≥80	32 (10.3%)
Sex	
Men	149 (48.1%)
Women	161 (51.9%)
Ethnicity	
Rungus	244 (78.7%)
Bajau/Suluk	34 (11.0%)
Others	32 (10.3%)
Marital status	
Married	229 (73.9%)
Widow & Divorce & Single	81 (26.1%)
Total Household Income Level (RM)	
0 – 999	222 (71.6%)
1000 – 1999	63 (20.3%)
>2000	25 (8.1%)
Smoking Status	
Active Smoker	51 (16.5%)
Ex-smoker	63 (20.3%)
Never Smoke	196 (63.2%)
Alcohol Status	
Active Drinker	56 (18.1%)
Ex-drinker	116 (37.4%)
Never Drink	138 (44.5%)
Comorbidity	
Yes	205 (66.1%)
No	105 (33.9%)
Body Mass Index (BMI)	
Underweight	42 (13.5%)
Normal	101 (32.6%)
Overweight	167 (53.9%)
ADL status	
independent	305 (98.4%)
dependent	5 (1.6%)
Self-rated health status	
Poor	22 (7.1%)
Moderate	127 (41.0%)
Good	161 (51.9%)
Presence of hobby	
Yes	52 (16.8%)
No	258 (83.2%)
Loneliness	
Yes	17 (5.5%)
No	293 (94.5%)
Living arrangement	
alone	25 (8.1%)
with spouse only	73 (23.5%)
with family members	212 (8.4%)

[a] Mean ± standard deviation

All variables with a p-value of less than 0.25 were included for multivariable analysis.[14] They were analysed using backward and forward methods of likelihood ratio in SPSS to get the preliminary model of predictors for depression.[15] The assumptions were subjected to further checking using the Hosmer and Lemeshow Test. The model was determined to be fit for the data if the p-value was not statistically significant.

RESULTS

The Prevalence of Depression

The prevalence of depression among older people in Kudat, as screened using M-GDS-14, was 73.2% (95% CI= 70.7-75.7) %, with 36.1% (112 participants) classified as having severe depression, defined as a GDS-14 score of 8 and above. Out of the total 310 respondents, 26.8% were classified as having no depression symptoms (95% CI= 24.3-29.3).

The Respondents' Demographics

Table I demonstrates the respondents' demographics. The mean age of the participants was 69.4 ± 7.00, ranging from 60 to 94 years. Most of the respondents are from Rungus ethnicity and are married, with a percentage of 78.7% and 73.9%, respectively. Seventy-one percent have a cumulative total household income of less than RM 1,000, and only 8% have more than RM 2,000. For employment status, half were

Table II: Simple logistic regression of sociodemographic factors, lifestyle habits, health status and social behaviours associated with depression among the elderly in Kudat

Variables	Depression (%)	Crude OR (95% CI)	p-value
Age			
60–69	126 (70.0%)	ref	
70–79	77 (78.6%)	1.57 (0.88-2.80)	0.126
≥80	24 (75.0%)	1.29 (0.54-3.04)	0.567
Sex			
Men	103 (69.1%)	ref	
Women	124 (77.0%)	1.50 (0.90-2.48)	0.118
Total Household Income Level (RM)			
1000 and above	67 (76.1%)	ref	
0 – 999	160 (72.1%)	1.00 (0.48-2.11)	0.993
Smoking Status			
Never Smoke	147 (75%)	ref	
Active Smoker	37 (72.5%)	0.88 (0.44-1.77)	0.721
Ex-smoker	43 (68.3%)	0.72 (0.39-1.33)	0.293
Alcohol Status			
Never Drink	98 (71.0%)	ref	
Active Drinker	36 (64.3%)	0.74 (0.38-1.42)	0.359
Ex-drinker	93 (80.2%)	1.65 (0.92-2.97)	0.094
Comorbidity			
No	66 (62.9%)	ref	
Yes	161 (78.5%)	2.16 (1.29-3.63)	0.003
Body Mass Index (BMI)			
Normal	66 (65.3%)	ref	
Underweight	32 (76.2%)	1.70 (0.75-3.85)	0.206
Overweight	129 (77.2%)	1.80 (1.04-3.11)	0.035
ADL status			
independent	223 (73.1%)	ref	
dependent	4 (80.0%)	1.47 (0.16-13.35)	0.732
Self-rated health status			
Good	106 (65.8%)	ref	
Moderate-Poor	121 (81.8%)	2.24 (1.33-3.79)	0.003
Presence of hobby			
Yes	36 (69.2%)	ref	
No	191 (74%)	1.27 (0.66-2.43)	0.476
Loneliness			
Yes	213 (72.7%)	ref	
No	14 (82.4%)	1.75 (0.49-6.26)	0.388
Living arrangement			
with family members	156 (73.6%)	ref	
alone	21 (84.0%)	1.89 (0.62-5.73)	0.264
with spouse only	50 (68.5%)	0.78 (0.44-1.40)	0.403

retirees, 43.2% were still working, and only 6.8% were unemployed.

Most of them denied smoking (63.2%); only 16.5% were active smokers, while the rest were ex-smokers. Forty-four percent denied drinking alcohol, 37.4% had quit drinking, and only 18.1% were active drinkers. Regarding the health status, most respondents suffer from at least one comorbid illness (66.1%). The survey found that only 5 out of the 310 respondents were ADL-dependent. Regarding BMI status, one-third of the respondents have a normal BMI; more than the majority (51.3%) reported being overweight, and only 13.5% had a BMI of less than 18. Only 7.1% reported poor health status, 41.0% were moderate, and most (51.9%) reported good health. The presence of hobbies, loneliness, and living arrangements were grouped into the social behaviours of the respondents. Most of them had no hobbies (83.2%) and denied feelings of loneliness (94.5%). Only 8% lived alone, 68.4% lived with family members, and 23.5% lived with their spouse.

A simple logistic regression was conducted to determine the sociodemographic variables associated with depression. Taking a p-value of less than 0.05, this study found no significant relationship between sociodemographic, lifestyle and elderly depression. The presence of comorbidity, overweight and moderate to poor self-rated health status are significant predictor variables related to depression. The elderly with comorbidity have more than twice the odds of having depression (crude OR: 2.16, 95% CI: 1.29-3.63). Those who are overweight have an 80% increase in odds of developing elderly depression (crude OR: 1.80, 95% CI: 1.042-3.110). Those who reported moderate to severe health status are also likely to have depression with an odd of more than twice (crude OR: 2.24, 95% CI: 1.33-3.79). For social behaviours associated with depression, the findings noted that the absence of hobbies, feelings of loneliness and living arrangements had no significant relationship with depression among the elderly.

Table III: Factors associated with depression among elderly in Kudat on multivariable logistic regression analysis

Variables	Adjusted OR (95% CI)	p-value
Comorbidity		
No	ref	
Yes	2.00 (1.18-3.38)	0.010
Self-rated health status		
Good	ref	
Moderate-Poor	2.09 (1.23-3.55)	0.007

Age, sex, employment status, alcohol status, comorbidity, BMI, and perception of health status were significant variables (when taken as a p-value of less than 0.25) to be included in the multiple logistics regression model.[14] Using forward method Likelihood Ratio statistics, the presence of comorbidity and self-rated health status were found to be significant predictor variables. The model generated had a Hasmer and Lameshow test of 0.280 and Nagelkerke R square score of 0.128. The elderly with comorbidity have almost twice the odds of having depression (aOR: 1.99, 95%: 1.17, 3.37). Moderate to poor self-rated health status is also likely to have elderly depression with an odd of 2 (aOR: 2.09, 95%: 1.23, 3.55).

DISCUSSION

The Prevalence Of Elderly Depression

The prevalence of depression among the Malaysian elderly was shown to be 27.8%, while the prevalence of depressive symptoms among community-dwelling older adults in Malaysia is 16.5%.[3,16] The findings of this study highlight a high prevalence of depressive symptoms (73.2%) among the elderly in Kudat, with 112 participants (36.1%) meeting the criteria for severe depression, defined as a GDS-14 score of 8 and above. A meta-analysis found that the pooled prevalence of depression among older adults globally was 31.74%. The study ranges from 7.7% in Malaysia and Australia to 81.1% in India.[17] The prevalence of depression among the elderly is similarly high in Asia. A study done in urban Vietnam, for example, revealed that depression is more prevalent in 66.89% of the elderly.[18]

There was significant prevalence heterogeneity between the screening tools, likely due to the different levels of sensitivity and specificity of the screening tools. This may be partly due to the questionnaire tool used in this research, which utilised a 14-item version of GDS. Other different types of scales can be used to screen for depression in the elderly population, like the Patient Health Questionnaire (PHQ), Hamilton Depression Rating Scale (HDRS), Beck Depression Inventory (BDI), Center for Epidemiologic Studies-Depression Scale (CES-D) and Cornell Scale for Depression in Dementia. Another possibility for high prevalence could be due to the high sensitivity of GDS that was used as a screening tool for depression.[10] Apart from that, the small size of this study yielded a higher prevalence of depression, as evidenced in a meta-analysis done on the prevalence of depression.[17] The proportion of depression in the elderly with Type 2 DM attending the Klinik Kesihatan Bandar, Sungai Petani, Kedah was 32.1%.[19], whereas those in Asajaya, Samarahan, Sarawak recorded a prevalence of 65.1%.[20] The prevalence of depression among the elderly in the community area of Felda Gunung Besaut 2 is the highest at 85.5%.[21] In comparison, the prevalence of elderly depression attending the outpatient clinic at Universiti Sains Malaysia Hospital was the lowest at 13.9%.[22] All of these studies used the same screening instruments, M-GDS 14.

The considerable difference between prevalence in our study compared to others, particularly in Malaysia, signifies the healthcare gap and unmet need for the rural communities in Sabah to manage health problems. The social disparity in the rural region is contributed to by low socioeconomic conditions, high unemployment, and limited access to health and digital technology. In 2022, Malaysia's projected target for the doctor-to-population ratio was 1:425. However, the ratio in Sabah is among the lowest at 1:872.[23]

More dire is the imbalance of mental health professionals density; for example, the psychiatrist-to-population ratio in Sabah in 2018 was 0.54 psychiatrists per 100,000 compared to Kuala Lumpur, which had 5.24 psychiatrists per 100,000 population.[24] The accessibility to these health professions is further limited because only 50% of them serve in the Ministry of Health, and the rest are either in other ministries (Ministry of Higher Education or Ministry of Defence) or private clinical practices and universities. The need for mental health professionals is paramount as the psychiatrist-to-population ratio was still the lowest in Sabah, despite an improvement from 0.30 to 0.54 psychiatrists per 100,000 population.[25]

Based on data released by the Malaysian Department of Statistics, the number of households in Sabah that reside less than five kilometres from public health centres increased from 74% in 2016 to 84% in 2019.[26] Though access to primary healthcare facilities in Sabah have been improving, the geographical landscape of other areas such as Kudat and its infrastructure facilities for example transportation, need to be addressed for rural populations.27

Not to mention, the health literacy rate is still worrying. Based on the National Health Morbidity Survey (NHMS) 2019, 24.3% had excellent health literacy, 40.7% possessed sufficient health literacy, and 35.0% possessed limited health literacy. Significantly, 43.2% of the Sabahans had a limited health literacy level, the highest prevalence among the other states.[28] Because of this, they are more prone to seek alternative care through religious or traditional methods, which can delay the diagnosis and proper treatment of mental disorders.

Those who are older, have a lower educational level, have a lower income and have a low unemployment status were

linked to low health literacy.[29] The digital divide in Sabah, made worse by social inequity, is more prominently affecting the elderly's adoption of technology in their daily activities. Low socioeconomic status, involving income, education, and social skills, limits resources for obtaining information beyond healthcare providers. This issue is more pronounced in rural communities due to social disparities. Limited internet access and difficulty reaching primary health care services result in poor health maintenance, delayed diagnosis, and disease progression.

Significant Factors And Elderly Depression

The findings of this study found that the presence of comorbidity and moderate to poor self-perception of health are predictive factors for depression in Kudat at a 5% level of significance. Owing to their age and degenerative body functions, they are also prone to chronic pain. If poorly managed, subsequent complications could lead to the impairment of performing daily activities and the development of depressive symptoms.[22] This study only dichotomised the presence of comorbidity into yes or no and did not consider the number or type of medical illnesses. Even with this limitation, the finding is still significant, proving that the presence of comorbidity increases the odds of having depression in the elderly, regardless of the severity.

In this study, having poor or moderate self-perception of health status is one of the risk factors for mild to severe depression. Studies conducted in Malaysia, specifically among institutionalised elderly in Sabah[30] and rural Malaysia [20,31] highlighted these significant findings. Self-reported health status is one of the elements linked with self-efficacy, a vital feature in the elderly since higher self-efficacy levels are connected with excellent perceived health and decreased depressive symptoms.[32] This association is supported by a meta-analysis done to determine the relationship between health status, including self-rated health status and chronic disease, and depression risk among the elderly. Interestingly, the study indicated that poor self-reported health status appeared to be more significantly related to depression than the occurrence of chronic disease, which is a similar finding to this study.[33]

Though this variable is subjective to respondents' perceptions, rating one's health as poor or fair, especially among those with recent experience of depressive symptoms, may differentiate a subgroup of patients with complex and difficult-to-treat forms of depression complicated by physical ill-health and social disadvantage. Some studies also note that self-reporting of poor health is a significant indicator due to the high frequency of chronic conditions and depressive symptoms in older adults.[32] Collaborative care models, such as the Program for Active and Healthy Ageing Pusat Aktiviti Warga Emas (PAWE) in Sabah, demonstrate the potential for addressing the mental health needs of older adults.[34] Community mental health services, when combined with initiatives like PAWE, offer holistic support tailored to the multifaceted needs of older adults. Beyond mental health care, PAWE also promotes social support and community engagement, which are crucial for improving the quality of life for older adults. Expanding PAWE's reach in rural areas like Kudat can significantly reduce depression symptoms and its long-term effects.

CONCLUSION

The study demonstrates that elderly depression in Kudat is relatively higher compared to national prevalence and warrants early management. As many studies emphasised the association of various adverse health outcomes with depression, screening for early detections and variation of intervention programs for the elderly population must be emphasised. This is crucial, especially in rural areas, because early detection of community diseases depends on primary health facilities. The result of this study concluded that those with comorbidities or moderate to poor self-rated health status are more likely to have elderly depression.

By understanding the specific factors contributing to depression, interventions and strategies can be designed to mitigate the risk and promote mental well-being among the elderly population. These interventions may include targeted social support programs, improved healthcare accessibility, mental health awareness campaigns, and culturally sensitive approaches. Information regarding depression and the availability of effective treatments for depression should be made available to the elderly themselves. Overall, the research on factors associated with elderly depression in Kudat contributes to the existing knowledge base on geriatric mental health and highlights the importance of addressing the unique challenges faced by the elderly population in this region.

Recommendations

Further research is needed to understand the specific mechanisms and causal relationships between these factors and elderly depression in Kudat. Longitudinal studies and more robust study designs, like extensive sampling methods, can provide additional insights into these relationships' temporal and contextual dynamics. In an effort to ascertain potential interventions that may alleviate depressive symptoms in this demographic, additional research should investigate protective factors, including family cohesion, social support, and community engagement.

Strength

To our knowledge, this is the first study to investigate the prevalence of depression and its associated factors among the elderly in the rural area of Kudat, Sabah. The multi-stage sampling used in this study represented Sabah, at least in the rural part of the state. Another strength of this study is using locally trained enumerators for data collection.

Limitation

This study has several limitations, mainly related to its cross-sectional nature. Due to this, the temporality of the association between some of the variables studied in this study cannot be established.

A questionnaire was used for data collection, subjecting the study to self-report bias, leading to recall bias or social desirability bias, which can subsequently affect the accuracy and reliability of the responses. Although the study used Malay-validated GDS-14, future studies should consider using the local dialect spoken by the elderly in Kudat as the population comprises multiple ethnicities with diverse languages and cultures. As there could be some language barrier, the data were not well captured.

The study focused on hypertension and self-rated health as predictors of depression, omitting other critical factors such as diabetes, cardiovascular disease, and mental health history. These conditions are strongly associated with depression in the elderly, and their exclusion may limit the comprehensiveness of the findings. Future studies should incorporate these variables to provide a more holistic understanding of the factors influencing depression.

Other risk factors, such as stressful life events and grief should have been investigated in this study. Protective factors of depression in the elderly, like social support and family dynamics, were also not collected and should be included in future research.

DISCLOSURE OF FUNDING

The authors confirm that Niigata University and the Malaysian Association of Epidemiology contributed funds to help carry out this study project. This research was funded by Grants in aid of Scientific Research from the Japan Society for the Promotion of Science, for the project titled, "What are the older persons who live with happiness even if they are ill or have their own disability? -Follow-up surveys of Japan and Southeast Asia-" (21K18453), and the Grants in aid for Health and Labor Administration Promotion Research Project titled "Study on promotion of active and healthy aging in ASEAN" (20BA2002) and "Research project on promoting quality long term care for the older in ASEAN countries" (23BA0301).

REFERENCES

1. Current population estimates, Malaysia, 2022. 2022.
2. Fiske A, Wetherell JL, Gatz M. Depression in older adults. Annu Rev Clin Psychol. 2009; 5: 363-89.
3. Malaysian Department of Statistics [Internet]. 2022 [cited 2023 Jul 5]. Key Findings Population and Housing Census of Malaysia, 2020 Administrative District. Available from: https://www.dosm.gov.my/uploads/publications/20221018093511.pdf
4. Adnan H. Revitalising rural Malaysia [Internet]. 2019 [cited 2024 Jul 9]. Available from: https://www.nst.com.my/opinion/columnists/2019/02/463706/revitalising-rural-malaysia
5. König H, König HH, Konnopka A. The excess costs of depression: a systematic review and meta-analysis. Epidemiol Psychiatr Sci. 2019 Apr 5; 29: e30.
6. Tolentino JC, Schmidt SL. DSM-5 criteria and depression severity: Implications for clinical practice. Front Psychiatry. 2018; 9: 450.
7. Li Z, Jeon YH, Low LF, Chenoweth L, O'Connor DW, Beattie E, Brodaty H. Validity of the geriatric depression scale and the collateral source version of the geriatric depression scale in nursing homes. Int Psychogeriatr. 2015; 27(9): 1495-504.
8. Conradsson M, Rosendahl E, Littbrand H, Gustafson Y, Olofsson B, Lövheim H. Usefulness of the Geriatric Depression Scale 15-item version among very old people with and without cognitive impairment. Aging Ment Health. 2013; 17(5): 638-45.
9. Ewe Eow T, Che Ismail H. Validation Malay Geriatric Depression Scale [Internet]. 2004 [cited 2023 Feb 15]. Available from: https://priory.com/psych/MalayGDS.htm
10. Norfadila S, Taib M, Wati A, Abidin Z, Akmal H, Hamid2 A. Prevalence and Associated Factors of Depression Among Malaysian Elderly Using Machine Learning Techniques. International Journal for Studies on Children, Women, Elderly and Disabled. 2022; 17: 75-82.
11. Burleigh E, Reeves I, McAlpine C, Davie J. Can doctors predict patients' abbreviated mental test scores? Age Ageing. 2002 Jul 1; 31(4): 303-6.
12. Lee SL, Pearce E, Ajnakina O, Johnson S, Lewis G, Mann F, et al. The association between loneliness and depressive symptoms among adults aged 50 years and older: A 12-year population-based cohort study. Lancet Psychiatry. 2021; 8(1): 48-57.
13. National Coordinating Committee on Food and Nutrition (NCCFN). Malaysia Dietary Guidelines 2020. Putrajaya (MY): Ministry of Health Malaysia; 2021.
14. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. Source Code Biol Med. 2008; 3: 17.
15. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health. 2020; 8(1): e000262.
16. Vanoh D, Shahar S, Yahya HM, Hamid TA. Prevalence and determinants of depressive disorders among community-dwelling older adults: findings from the towards useful aging study. Int J Gerontol. 2016; 10(2): 81-5.
17. Zenebe Y, Akele B, W/Selassie M, Necho M. Prevalence and determinants of depression among old age: a systematic review and meta-analysis. Ann Gen Psychiatry. 2021; 20(1): 55.
18. Dao ATM, Nguyen VT, Nguyen HV, Nguyen LTK. Factors associated with depression among the elderly living in urban Vietnam. Biomed Res Int. 2018; 2018: 2370284.
19. Muhamad Radzi A, Draman N, Suhaila Mohd Yusoff S, Muhamad R. Depression and potential risk factors among the elderly with Type 2 Diabetes Mellitus in Kedah, Malaysia. Medical Journal of Malaysia. 2019; 74(2): 103-8.
20. Wan Ahmad SN. Depression and its associated factors among elderly in Asajaya District in Samarahan division, Sarawak [master's thesis on the Internet]. Kota Samarahan (MY): Universiti Malaysia Sarawak; 2010 [cited 2025 Jan 20]. Available from: https://ir.unimas.my/id/eprint/12069/
21. Wahab R, Nor Azni SN, Masuri MG. The Prevalence of Depression Among Elderly in Rural Community Area of Southern Perak. Journal of Islamic, Social, Economics and Development. 2021; 6(6): 246-54.
22. Imran A, Azidah AK. Prevalence of Depression and its Associated Factors Among Elderly Patients in Outpatient Clinic of Universiti Sains Malaysia Hospital. Med J Malaysia. 2009; 64(2): 134-9.
23. Unto R. Provide fair allocation to Sabah: Shahelmy | Daily Express Online - Sabah's Leading News Portal [Internet]. 2023 [cited 2023 Jul 3]. Available from: https://www.dailyexpress.com.my/news/208106/provide-fair-allocation-to-sabah-shahelmy/
24. Raaj S, Navanathan S, Tharmaselan M, Lally J. Mental disorders in Malaysia: an increase in lifetime prevalence. BJPsych Int. 2021; 18(4): 97-9.
25. Midin M, Zainal NZ, Lee TC, Ibrahim N. Mental Health Services in Malaysia. Taiwanese Journal of Psychiatry (Taipei). 2018; 32(3): 281-92.
26. Household Income and Basic Amenities Survey Report by State and Administrative District Sabah [Internet]. 2020 [cited 2023 Jun 25]. Available from: https://www.dosm.gov.my/v1/uploads/files/1_Articles_By_Themes/Prices/HIES/HIS-Report/HIS_Sabah.pdf
27. Naing OT, Shoesmith WD, Tan CBY, Ibrahim MY, Hussein SS. Geographic accessibility of healthcare services and health-seeking behaviours of rural communities in Kudat and Pitas areas of Sabah. Borneo Epidemiol J. 2020;1(1):46–54. doi:10.51200/bej.v1i1.2436.
28. Institute for Public Health. Key findings of NHMS 2019 [Internet]. Putrajaya (MY): Ministry of Health Malaysia; 2020 [cited 2025 Jan 20]. Available from: https://iku.moh.gov.my/images/IKU/Document/REPORT/NHMS2019/Infographic_Booklet_NHMS_2019-English.pdf

29. Jaafar N, Perialathan K, Krishnan M, Juatan N, Ahmad M, Yong TS, et al. Malaysian health literacy: scorecard performance from a national survey. Int J Environ Res Public Health. 2021; 18(11): 5813.
30. Sajali H, Rahim SSSA, Abidin A, Subramaniam P, Sazali MF, Jeffree MS, et al. Prevalence and risk factors of depressive symptoms among institutionalized elderly in Sabah, Malaysia Borneo. Malays J Public Health Med. 2021; 21(1): 253-9.
31. Abdul Manaf MR, Mustafa M, Abdul Rahman MR, Yusof KH, Abd Aziz NA. Factors influencing the prevalence of mental health problems among Malay elderly residing in a rural community: A cross-sectional study. PLoS One. 2016; 11(6): e0156937.
32. Patrão AL, Alves V. The influence of self-perceived health status, social support and depression on self-efficacy among Brazilian elderly people. Interdisciplinaria. 2019; 36(1): 23-32.
33. Chang-Quan H, Xue-Mei Z, Bi-Rong D, Zhen-Chan L, Ji-Rong Y, Qing-Xiu L. Health status and risk for depression among the elderly: a meta-analysis of published literature. Age Ageing. 2010 Jan 1;39(1):23–30. doi:10.1093/ageing/afp187.
34. Jabatan Perkhidmatan Kebajikan Am Sabah. Pusat Aktiviti Warga Emas (PAWE) [Internet]. 2025 [cited 2025 Jan 20]. Available from: https://jpka.sabah.gov.my/perkhidmatan/pembangunan-komuniti/pusat-aktiviti-warga-emas-pawe

ORIGINAL ARTICLE

Evaluation of dry eye pre and post phacoemulsification in diabetic patients

Nur Faizah Harun, MD[1,2,3], Mohd Feendi Mohd Fauzi Yap, DrOphthal[1], Siti Zakiah Md Khair, DrOphthal[1], Wan Hazabbah Wan Hitam, MSurg(Ophthal)[2,3]

[1]Department of Ophthalmology, Hospital Melaka, Jalan Mufti Haji Khalil, Melaka, Malaysia, [2]Department of Ophthalmology and Visual Science, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia, [3]Hospital Pakar Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

ABSTRACT

Introduction: Dry eye is a common condition influenced by various factors, including cataract surgery and systemic diseases like diabetes. Phacoemulsification, a widely used cataract procedure, often leads to increased postoperative dry eye symptoms due to inflammation and changes in tear film stability. Diabetic patients, already prone to dry eye, may experience further worsening after surgery. This study evaluates dry eye status in diabetic patients before and after phacoemulsification.

Materials and Methods: This study included 126 patients, divided into diabetic and non-diabetic groups from Hospital Melaka and Hospital Pakar Universiti Sains Malaysia between September 2022 and July 2024. Patient demographics and dry eye parameters, including the Ocular Surface Disease Index (OSDI), Tear Break-Up Time (TBUT), and Schirmer's test, were evaluated at baseline, one week, and three months after surgery. Only patients who underwent uncomplicated phacoemulsification were included, while those undergoing extracapsular cataract extraction, intracapsular cataract extraction, or lens aspiration were excluded. The mean values of OSDI, TBUT, and Schirmer's test across the three time points were compared among groups with diabetic retinopathy, without diabetic retinopathy and non-diabetic patients using repeated measures ANOVA.

Results: 126 patients were studied: 44 non-diabetic, 40 diabetic without retinopathy (no DR), and 42 diabetics with retinopathy (DR). The mean ± SD (standard deviation) age was 64.06 ± 5.30 years, with males comprising 54.0% of the cohort. Hypertension was the highest proportion of comorbidity (75.4%), particularly in the DR group (90.5%). Dry eye parameters showed significant temporary changes post-cataract surgery. OSDI scores improved significantly from baseline to three months in all groups, with diabetic groups showing higher scores at three months than non-diabetics ($p < 0.05$). TBUT declined significantly at one week in the diabetic groups (DR, $p = 0.028$; no DR, $p = 0.019$) but showed substantial recovery by three months, with significant improvements across all groups. In all groups, Schirmer's test values improved significantly between one week and three months ($p < 0.05$), although baseline and one-week differences were not statistically significant.

Conclusion: Three months after cataract surgery, significant improvements in OSDI scores, TBUT, and Schirmer's test values were observed, indicating a recovery in dry eye status. Diabetic patients experienced more pronounced early postoperative changes but demonstrated comparable recovery trends to non-diabetics by three months. These findings highlight the importance of monitoring dry eye parameters in diabetics, particularly during the early postoperative period, to optimise outcomes and patient satisfaction.

KEYWORDS:
Dry eye, phacoemulsification, diabetic, non-diabetic, retinopathy

INTRODUCTION

Dry eye disease is a multifactorial condition that can significantly affect ocular comfort and vision. Various factors, including systemic diseases like diabetes and surgical interventions such as cataract surgery, can exacerbate dry eye symptoms.[1,2] Phacoemulsification, a common and minimally invasive procedure for cataract extraction, is associated with an increased incidence of postoperative dry eye symptoms due to factors such as reduced tear production, changes in corneal innervation, and inflammation.[2,3]

Diabetic patients predisposed to dry eye due to neuropathy, inflammation, and metabolic dysregulation may experience a heightened risk of developing or worsening dry eye following cataract surgery.[3,4] Given the prevalence of diabetes and the growing number of cataract surgeries performed globally, understanding the impact of phacoemulsification on dry eye symptoms in diabetic patients is critical for optimising postoperative care.

This study aims to evaluate the changes in dry eye status in diabetic patients undergoing phacoemulsification by assessing pre- and postoperative dry eye parameters, including the Ocular Surface Disease Index (OSDI) questionnaire, tear break-up time (TBUT) and Schirmer's test.

MATERIALS AND METHODS

Study design and study population

This prospective cohort study was conducted between September 2022 and July 2024 at the Ophthalmology Clinics

This article was accepted: 25 March 2025
Corresponding Author: Wan Hazabbah Wan Hitam
Email: hazabbah@usm.my, hazabbah@yahoo.com

of Hospital Melaka and Hospital Pakar Universiti Sains Malaysia. Ethical approval was obtained from the Medical Research and Ethics Committee of the Ministry of Health Malaysia (NMRR ID-22-01285-XDE (IIR)) and the Human Research Ethics Committee of Universiti Sains Malaysia (USM/JEPeM/22060353). The study included 82 diabetic patients, aged 40 to 75 years, who underwent cataract surgery via phacoemulsification, along with 44 non-diabetic patients as controls.

Participants were screened based on specific inclusion and exclusion criteria. Those with underlying ocular conditions known to contribute to dry eye, such as blepharitis, ectropion, entropion, pre-existing dry eye diagnosis, glaucoma, or a history of ocular trauma, were excluded. Similarly, patients undergoing non-phacoemulsification cataract surgeries, including extracapsular cataract extraction, intracapsular cataract extraction, or lens aspiration, were not eligible.

Participants with complicated cataract surgeries or those lost to follow-up at the one-week or three-month postoperative assessments were also excluded from the analysis. All enrolled participants underwent standardised dry eye assessments as part of the study protocol.

Demographic data, systemic and ocular history
Demographic data for participants, including age, gender, race, education level, duration of diabetes, and any underlying diseases, were collected. Detailed histories were also obtained through participant interviews, noting the use of any topical eye drops or medications such as antihistamines, antidepressants, or decongestants.

Ocular examination
Participants who provided informed consent underwent thorough examinations at the ophthalmology clinic. Comprehensive ocular assessments, including visual acuity measurement, anterior segment evaluation, and dilated fundus examination, were conducted using a slit lamp biomicroscope (Topcon Corp, Japan) and condensing lenses. Patients with ocular pathologies meeting the exclusion criteria were excluded.

Dry eye parameters were evaluated subjectively and objectively using the Ocular Surface Disease Index (OSDI), Tear Break-Up Time (TBUT), and Schirmer I test without anaesthesia. The OSDI identified patients with dry eye symptoms based on their responses to a 12-item questionnaire.

TBUT was assessed to evaluate tear film stability. A fluorescein-impregnated strip moistened with non-preserved saline was applied, and the fluorescein dye was distributed by blinking. Patients were instructed to look straight ahead without blinking, and the time between the last blink and the appearance of the first dry spot or break in the tear film was measured under cobalt blue light using a slit lamp. A TBUT of more or equal to 10 seconds was considered normal, while less than 10 seconds indicated dry eye. The test was repeated three times, and the mean value was recorded.

The Schirmer I test assessed both basic and reflex tearing. After drying the inferior fornix, a sterile paper strip was placed at the lateral third of the lower eyelid. The length of the moistened portion of the strip was measured after five minutes. A wetting length less than or equal to 10 millimetres (mm) indicates a dry eye. The Schirmer I test was performed once for each participant.

A 10-minute interval was maintained between the Schirmer I test and the TBUT assessment. All tests were performed at baseline (preoperatively), as well as at one week and three months postoperatively.

Statistical Analysis
The data were cleaned and analysed using SPSS version 29.0. The distribution of continuous variables was checked through histograms. Continuous variables found to be normally distributed are presented as mean (standard deviation, SD), while categorical variables are shown as frequency and percentage.

The mean OSDI, TBUT, and Schirmer's test values at baseline, one week, and three months post-cataract surgery were compared among the diabetic retinopathy, no diabetic retinopathy, and non-diabetic groups using repeated measures ANOVA (RM ANOVA). Mauchly's test of sphericity was conducted to assess the assumption of sphericity. When the assumption was violated, Greenhouse-Geisser correction was applied for Mauchly's $W < 0.75$, while Huynh-Feldt correction was used for Mauchly's $W > 0.75$. Estimated marginal means were calculated to evaluate the main effects of group and time, with Bonferroni correction applied for multiple comparisons. Additionally, repeated measures Analysis of Covariance (ANCOVA) was performed, adjusting for HbA1c levels and the duration of diabetes. All statistical tests were two-sided, with p-values < 0.05 considered significant.

RESULTS

Demographic characteristics
A total of 126 patients were included in the study, comprising 44 non-diabetic, 40 diabetics without retinopathy (no DR), and 42 diabetics with retinopathy (DR). The overall mean ± SD (standard deviation) age of the cohort was 64.06 ± 5.30 years. The DR group had the youngest mean age at 63.29 ± 4.90 years, while the no DR group had the oldest mean age at 65.68 ± 4.10 years. Males comprised 54.0% of the entire cohort, and females were the majority in the no DR group (52.5%). Nearly all female participants were postmenopausal (94.8%). The majority of patients were Malay (72.2%), with the highest proportion observed in the DR group (78.6%) (Table I).

Most had completed secondary education (66.7%), while only a small percentage had tertiary education, ranging from 7.10% in the DR group to 11.4% in the non-diabetic group. Hypertension was the most prevalent comorbidity (75.4%), highest in the DR group (90.5%), followed by hyperlipidaemia (35.7%). Only 7.1% reported no comorbid conditions. Right-eye involvement was more common overall (53.2%), particularly in the DR group (61.9%), while left-eye involvement predominated in the non-diabetic group (59.1%) (Table I).

Table I: Demographic characteristics of the patients

Characteristics	Overall (n= 126)	Non-diabetic (n= 44)	No DR (n= 40)	DR (n= 42)
Age in years, mean ± SD	64.06 ± 5.30	63.34 ± 6.33	65.68 ± 4.10	63.29 ± 4.90
Gender, n (%)				
Female	58 (46.0)	18 (40.9)	21 (52.5)	19 (45.2)
Male	68 (54.0)	26 (59.1)	19 (47.5)	23 (54.8)
If female, menopause, n (%)				
No	3 (5.2)	2 (11.1)	0 (0.0)	1 (5.3)
Yes	55 (94.8)	16 (88.9)	21 (100.0)	18 (94.7)
Race, n (%)				
Malay	91 (72.2)	29 (65.9)	29 (72.5)	33 (78.6)
Chinese	24 (19.0)	14 (31.8)	6 (15.0)	4 (9.5)
Indian	11 (8.7)	1 (2.3)	5 (12.5)	5 (11.9)
Educational level, n (%)				
Primary	31 (24.6)	11 (25.0)	11 (27.5)	9 (21.4)
Secondary	84 (66.7)	28 (63.6)	26 (65.0)	30 (71.4)
Tertiary	11 (8.7)	5 (11.4)	3 (7.5)	3 (7.1)
Comorbidities, n (%)				
No illness	9 (7.1)	9 (20.5)	0 (0.0)	0 (0.0)
Hypertension	95 (75.4)	26 (59.1)	31 (77.5)	38 (90.5)
Hyperlipidaemia	45 (35.7)	12 (27.3)	17 (42.5)	16 (38.1)
CKD/ESRF	10 (7.9)	1 (2.3)	1 (2.5)	8 (19.0)
IHD/AF	19 (15.1)	5 (11.4)	6 (15.0)	8 (19.0)
Stroke	3 (2.4)	2 (4.5)	0 (0.0)	1 (2.4)
Thyroid disorder	2 (1.6)	0 (0.0)	1 (2.5)	1 (2.4)
Bronchial asthma	7 (5.6)	5 (11.4)	2 (5.0)	0 (0.0)
Others	8 (6.3)	7 (15.9)	0 (0.0)	1 (2.4)
Laterality, n (%)				
Right eye	67 (53.2)	18 (40.9)	23 (57.5)	26 (61.9)
Left eye	59 (46.8)	26 (59.1)	17 (42.5)	16 (38.1)

Abbreviations: DR: diabetic retinopathy, SD: Standard deviation, CKD: chronic kidney disease, ESRF: end-stage renal failure, IHD: ischaemic heart disease, AF: atrial fibrillation

Table II: Characteristics of the diabetic patients

Characteristics	Overall (Mean ± SD)	No DR (Mean ± SD)	DR (Mean ± SD)	p-value
Duration of DM in years	10.76 ± 7.47	7.38 ± 5.39	13.98 ± 7.81	<0.001*
HbA1c	6.50 ± 0.48	6.20 ± 0.33	6.79 ± 0.41	<0.001*

*Independent sample t-test
Abbreviations: DR: diabetic retinopathy, SD: Standard deviation, DM: diabetes mellitus, HbA1c: Haemoglobin A1c

Characteristics of the diabetic patients

Table II outlines the characteristics of the diabetic patients, all of whom were classified with Type 2 diabetes mellitus (DM). The mean duration of diabetes was significantly longer in the DR group (13.98 ± 7.81 years) compared to the no DR group (7.38 ± 5.39 years, $p < 0.001$). Similarly, the mean HbA1c level was significantly higher in the DR group (6.79 ± 0.41) than in the no DR group (6.20 ± 0.33, $p < 0.001$).

Dry eye parameters

Table III summarizes the changes in dry eye parameters (OSDI, TBUT, and Schirmer's test) before and after cataract surgery among the diabetic retinopathy (DR), no diabetic retinopathy (no DR), and non-diabetic groups, analysed using repeated measures ANOVA (RM ANOVA).

Lower OSDI scores indicate an improvement in dry eye symptoms. Among all groups, the DR group consistently had the highest OSDI scores; however, these scores reflected only mild dry eye symptoms compared to the other groups. From baseline to one-week post-surgery, OSDI scores showed slight reductions across all groups, but these changes were not statistically significant. By three months post-surgery, however, all groups exhibited significant improvements in OSDI scores compared to baseline, as indicated by reduced scores: DR group (mean decrease of 6.41, $p = 0.015$), no DR group (mean decrease of 6.74, $p < 0.001$), and non-diabetic group (mean decrease of 6.68, $p = 0.001$). Between one week and three months post-surgery, OSDI scores continued to decline significantly across all groups ($p < 0.001$). These findings suggest a gradual but significant improvement in dry eye symptoms following cataract surgery.

From baseline to one-week post-surgery, TBUT significantly decreased in the DR group ($p = 0.028$) and no DR group ($p = 0.019$), while no significant change was observed in the non-diabetic group ($p > 0.950$). By three months post-surgery, TBUT values significantly improved compared to baseline in

Table III: Dry eye parameters before and after cataract surgery in the operated eyes of diabetic and non-diabetic patients (within factors)

Comparison	DR			No DR			Non-diabetic		
	Mean ± SD	MD (95% CI)	p value	Mean ± SD	MD (95% CI)	p value	Mean ± SD	MD (95% CI)	p-value
OSDI									
Baseline vs	13.15 ± 16.57	0.26 (-3.25, 3.76)	>0.950	12.16 ± 11.63	0.73 (-2.10, 3.57)	>0.950	9.89 ± 11.59	2.23 (-1.11, 5.56)	0.309
1 week	12.90 ± 12.13			11.43 ± 9.95			7.67 ± 8.19		
Baseline vs	13.15 ± 16.57	6.41 (1.02, 11.80)	0.015	12.16 ± 11.63	6.74 (2.73, 10.76)	<0.001	9.89 ± 11.59	6.68 (2.60, 10.75)	<0.001
3 months	6.75 ± 4.50			5.42 ± 3.91			3.22 ± 3.98		
1 week vs	12.90 ± 12.13	6.15 (2.80, 9.50)	<0.001	11.43 ± 9.95	6.01 (2.90, 9.13)	<0.001	7.67 ± 8.19	4.45 (1.73, 7.16)	<0.001
3 months	6.75 ± 4.50			5.42 ± 3.91			3.22 ± 3.98		
TBUT									
Baseline vs	7.31 ± 2.40	0.84 (0.07, 1.60)	0.028	7.72 ± 2.95	1.17 (0.16, 2.19)	0.019	7.05 ± 2.47	-0.02 (-0.87, 0.83)	>0.950
1 week	6.48 ± 2.22			6.55 ± 2.34			7.07 ± 2.11		
Baseline vs	7.31 ± 2.40	-2.03 (-3.03, -1.03)	<0.001	7.72 ± 2.95	-1.23 (-2.27, -0.19)	0.016	7.05 ± 2.47	-1.75 (-2.84, -0.66)	<0.001
3 months	9.35 ± 2.50			8.96 ± 2.22			8.80 ± 2.11		
1 week vs	6.48 ± 2.22	-2.87 (-3.69, -2.05)	<0.001	6.55 ± 2.34	-2.41 (-3.18, -1.63)	<0.001	7.07 ± 2.11	-1.73 (-2.34, -1.12)	<0.001
3 months	9.35 ± 2.50			8.96 ± 2.22			8.80 ± 2.11		
Schirmer's test									
Baseline vs	9.21 ± 5.31	0.14 (-1.58, 1.86)	>0.095	10.33 ± 5.76	0.93 (-0.99, 2.84)	0.705	8.18 ± 4.67	-0.71 (-2.57, 1.17)	>0.095
1 week	9.07 ± 3.99			9.40 ± 4.70			8.89 ± 4.49		
Baseline vs	9.21 ± 5.31	-1.41 (-3.13, 0.32)	0.146	10.33 ± 5.76	-0.48 (-2.33, 1.38)	>0.095	8.18 ± 4.67	-2.43 (-4.25, -0.62)	0.005
3 months	10.62 ± 3.46			10.80 ± 3.15			10.61 ± 3.53		
1 week vs	9.07 ± 3.99	-1.55 (-2.49, -0.60)	<0.001	9.40 ± 4.70	-1.40 (-2.46, -0.34)	0.006	8.89 ± 4.49	-1.73 (-2.63, -0.82)	<0.001
3 months	10.62 ± 3.46			10.80 ± 3.15			10.61 ± 3.53		

Abbreviations: OSDI: Ocular Surface Disease Index, DR: diabetic retinopathy, SD: Standard deviation, MD: Mean difference, CI: Confidence interval, vs: versus, TBUT: Tear break-up time

Table IV: Dry eye parameters in the operated eyes of diabetic and non-diabetic patients (between factors)

Comparison	Estimated Mean (95% CI)	MD (95% CI)	p-value
OSDI			
DR vs	10.93 (8.37, 13.50)	1.26 (-3.24, 5.76)	>0.950
No DR	9.67 (7.04, 12.30)		
DR vs	10.93 (8.37, 13.50)	4.01 (-0.39, 8.40)	0.086
Non-diabetic	6.93 (4.42, 9.43)		
No DR vs	9.67 (7.04, 12.30)	2.74 (-1.71, 7.19)	0.412
Non-diabetic	6.93 (4.42, 9.43)		
TBUT			
DR vs	7.71 (7.11, 8.31)	-0.03 (-1.08, 1.02)	>0.950
No DR	7.74 (7.13, 8.36)		
DR vs	7.71 (7.11, 8.31)	0.08 (-0.95, 1.10)	>0.950
Non-diabetic	7.64 (7.05, 8.22)		
No DR vs	7.74 (7.13, 8.36)	0.11 (-0.93, 1.15)	>0.950
Non-diabetic	7.64 (7.05, 8.22)		
Schirmer's test			
DR vs	9.64 (8.50, 10.77)	-0.54 (-2.53, 1.45)	>0.950
No DR	10.18 (9.01, 11.34)		
DR vs	9.64 (8.50, 10.77)	0.41 (-1.54, 2.35)	>0.950
Non-diabetic	9.23 (8.12, 10.34)		
No DR vs	10.18 (9.01, 11.34)	0.95 (-1.02, 2.92)	0.736
Non-diabetic	9.23 (8.12, 10.34)		

Abbreviations: CI: confidence interval, MD: Mean difference, OSDI: Ocular Surface Disease Index, DR: diabetic retinopathy, TBUT: Tear break-up time

all groups. The DR group showed the greatest increase (2.03 seconds, $p < 0.001$), followed by the non-diabetic group (1.75 seconds, $p = 0.001$) and the no DR group (1.23 seconds, $p = 0.016$). Between one week and three months post-surgery, TBUT values continued to recover significantly in all groups. The DR group demonstrated the largest improvement (2.87 seconds, $p < 0.001$), followed by the no DR group (2.41 seconds, $p < 0.001$) and the non-diabetic group (1.73 seconds, $p < 0.001$). These results indicate an initial post-surgical decline in TBUT among diabetic patients, followed by substantial recovery by three months.

Schirmer's test values showed minimal changes from baseline to one week post-surgery, with no statistically significant differences across all groups. By three months post-surgery, the DR and no DR groups demonstrated slight increases in Schirmer's test values, but these changes were not statistically significant ($p = 0.146$ and $p > 0.950$, respectively). In contrast, the non-diabetic group showed a significant improvement in Schirmer's test values during this period ($p = 0.005$).

Between one week to three months post-operation, Schirmer's test values increased significantly across all groups. The DR group showed an improvement of 1.55 ($p = 0.001$), the no DR group improved by 1.40 ($p = 0.006$), and the non-diabetic group exhibited the largest increase of 1.73 ($p < 0.001$). Overall, the non-diabetic group demonstrated the most notable increase in tear production by three months post-surgery.

Table IV presents the results of dry eye parameters analysed using repeated measures ANOVA (RM ANOVA) across the groups. No significant differences were found in any dry eye parameters between the groups ($p > 0.050$).

Table V presents the comparison of dry eye parameters (OSDI, TBUT, and Schirmer's test) among the DR, no DR, and non-diabetic groups at baseline, one week, and three months post-surgery.

At baseline, OSDI scores were comparable across all groups, with no statistically significant differences ($p > 0.050$). At one week post-surgery, although the DR group had slightly higher OSDI scores than the no DR and non-diabetic groups, the differences were not statistically significant ($p > 0.950$ for DR versus no DR, $p = 0.057$ for DR versus non-diabetic). By three months post-surgery, significant differences were observed between the DR and non-diabetic groups ($p < 0.001$) and between the no DR and non-diabetic groups ($p = 0.050$). However, no significant difference was found between the DR and no DR groups ($p = 0.446$). At three months, diabetic groups (DR and no DR) exhibited significantly higher OSDI scores compared to the non-diabetic group.

At baseline, TBUT values were comparable across all groups, with no significant differences detected ($p > 0.050$). One week after surgery, TBUT values decreased in both the DR and no DR groups; however, the differences between the groups remained statistically insignificant. By three months after surgery, TBUT values improved in all groups, but no significant intergroup differences were observed at any time point ($p > 0.050$). Overall, TBUT values were similar across the DR, no DR, and non-diabetic groups at all time points ($p > 0.050$, Table V).

At baseline, Schirmer's test values were similar across all groups, with no statistically significant differences observed ($p > 0.050$). One week post-surgery, minimal changes in Schirmer's test values were observed, with no statistically significant differences between the groups ($p > 0.950$). By three months post-surgery, Schirmer's test results were

Table V: Dry eye parameters before and after cataract surgery in the operated eyes of diabetic and non-diabetic patients (within-between factors)

Comparison	OSDI			TBUT			Schirmer's test		
	Mean ± SD	MD (95% CI)	p value	Mean ± SD	MD (95% CI)	p value	Mean ± SD	MD (95% CI)	p-value
Baseline									
DR vs	13.15 ± 16.57	0.99 (-6.23, 8.22)	>0.950	7.31 ± 2.40	-0.41 (-1.81, 0.99)	>0.950	9.21 ± 5.31	-1.11 (-3.93, 1.70)	>0.950
No DR	12.16 ± 11.63			7.72 ± 2.95			10.33 ± 5.76		
DR vs	13.15 ± 16.57	3.26 (-3.79, 10.31)	0.792	7.31 ± 2.40	0.27 (-1.10, 1.64)	>0.950	9.21 ± 5.31	1.03 (-1.72, 3.78)	>0.950
Non-diabetic	9.89 ± 11.59			7.05 ± 2.47			8.18 ± 4.67		
No DR vs	12.16 ± 11.63	2.27 (-4.88, 9.41)	>0.950	7.72 ± 2.95	0.68 (-0.71, 2.06)	0.714	10.33 ± 5.76	2.14 (-0.64, 4.93)	0.192
Non-diabetic	9.89 ± 11.59			7.05 ± 2.47			8.18 ± 4.67		
1 week									
DR vs	12.90 ± 12.13	1.47 (-4.00, 6.93)	>0.950	6.48 ± 2.22	-0.07 (-1.27, 1.12)	>0.950	9.07 ± 3.99	-0.33 (-2.69, 2.03)	>0.950
No DR	11.43 ± 9.95			6.55 ± 2.34			9.40 ± 4.70		
DR vs	12.90 ± 12.13	5.23 (-0.11, 10.57)	0.057	6.48 ± 2.22	-0.59 (-1.76, 0.57)	0.657	9.07 ± 3.99	0.19 (-2.12, 2.49)	>0.950
Non-diabetic	7.67 ± 8.19			7.07 ± 2.11			8.89 ± 4.49		
No DR vs	11.43 ± 9.95	3.76 (-1.64, 9.17)	0.281	6.55 ± 2.34	-0.52 (-1.70, 0.66)	0.863	9.40 ± 4.70	0.51 (-1.82, 2.85)	>0.950
Non-diabetic	7.67 ± 8.19			7.07 ± 2.11			8.89 ± 4.49		
3 months									
DR vs	6.75 ± 4.50	1.33 (-0.89, 3.55)	0.446	9.35 ± 2.50	0.39 (-0.83, 1.61)	>0.950	10.62 ± 3.46	-0.18 (-2.00, 1.64)	>0.950
No DR	5.42 ± 3.91			8.96 ± 2.22			10.80 ± 3.15		
DR vs	6.75 ± 4.50	3.53 (1.36, 5.69)	<0.001	9.35 ± 2.50	0.55 (-0.65, 1.74)	0.799	10.62 ± 3.46	0.01 (-1.77, 1.78)	>0.950
Non-diabetic	3.22 ± 3.98			8.80 ± 2.11			10.61 ± 3.53		
No DR vs	5.42 ± 3.91	2.20 (0.01, 4.39)	0.050	8.96 ± 2.22	0.16 (-1.05, 1.37)	>0.950	10.80 ± 3.15	0.19 (-1.61, 1.98)	>0.950
Non-diabetic	3.22 ± 3.98			8.80 ± 2.11			10.61 ± 3.53		

Abbreviations: OSDI: Ocular Surface Disease Index, SD: Standard deviation, MD: Mean difference, CI: Confidence interval, DR: diabetic retinopathy, vs: versus, TBUT: Tear break-up time

comparable across all groups, with no significant differences detected ($p > 0.950$). Overall, Schirmer's test values remained consistent across the DR, no DR, and non-diabetic groups at baseline, one week, and three months post-surgery.

DISCUSSION

This study evaluated changes in dry eye parameters using OSDI scores, TBUT, and Schirmer's test among non-diabetic, diabetic without retinopathy (no DR), and diabetic retinopathy (DR) patients following phacoemulsification. Findings revealed important insights into the variations in dry eye symptoms and signs across these groups, with a gradual improvement in ocular surface health observed in the long-term post-surgery.

The cohort predominantly consisted of Malay participants, with hypertension being the most common comorbidity, especially in the DR group, aligning with previous studies linking systemic hypertension to the progression of DR.[5] Gender distribution showed a male predominance except in the non-diabetic group, where females outnumbered males. Notably, many of the females in the cohort were postmenopausal, a factor that may contribute to an increased risk of dry eye symptoms.[6,7] Sriprasert et al. highlighted that postmenopausal women are at a higher risk for ocular surface issues due to hormonal changes, which could influence both subjective symptoms and objective dry eye measures.[6] The study also found a slight age difference across groups, with the DR group being slightly younger, suggesting earlier onset or more rapid progression of retinopathy in some individuals, potentially contributing to the observed ocular surface changes post-surgery. Patients who are older and female are at an increased risk of developing dry eye symptoms following cataract surgery.[8]

Despite early postoperative dry eye symptoms, no significant differences in OSDI, TBUT, or Schirmer's test were found among groups at any time point. However, a significant improvement from baseline to three months suggests transient postoperative ocular surface stress with subsequent stabilisation. Jiang et al. reported a dry eye incidence of 17.1% in diabetic and 8.1% in non-diabetic patients one week post-cataract surgery, which resolved in both groups by three months.[3] Kasetsuwan et al. similarly found a 9.8% incidence of postoperative dry eye, peaking at one week before improving.[9] Their findings align with ours, showing no significant differences in TBUT, Schirmer's test, or OSDI early postoperatively. These results suggest that dry eyes following cataract surgery are predominantly transient, driven by postoperative inflammation and ocular surface stress.

For the OSDI parameter, no significant changes were observed between baseline and one week post-surgery in any group. However, a significant decrease was evident from baseline to three months and from one week to three months across all groups, suggesting that while early postoperative dry eye symptoms persist, substantial relief occurs within three months. Li et al. attributed postoperative dry eye to poor adherence to prescribed eye drops, aligning with the understanding that surgical inflammation initially worsens symptoms before recovery stabilises them.[10,11] At three months, DR and no DR patients reported greater discomfort than non-diabetic patients. Jiang et al. showed that diabetic patients had significantly increased OSDI scores post-surgery, with non-diabetics returning to baseline within a month.[3] This may reflect diabetic pathology affecting nerve function and inflammatory responses, prolonging subjective dry eye symptoms even as TBUT and Schirmer's test improve.[8,12] Zamora et al. reported significant post-surgical improvement, consistent with our findings.[12] In our cohort, DR patients had milder preoperative dry eye symptoms than no DR and non-diabetic groups, as indicated by the OSDI scores.[13,14] Notably, patients with altered preoperative values indicative of dry eye were more likely to experience prolonged ocular surface disease postoperatively.[9]

TBUT analysis revealed significant differences among groups. Both the DR and no DR groups experienced a marked reduction in TBUT from baseline to one week postoperatively, likely due to postoperative inflammation and tear film instability,[10] whereas non-diabetic patients showed no significant change. However, all groups exhibited substantial TBUT improvement from baseline to three months, reflecting progressive tear film recovery. Kohli et al. found a transient decline in dry eye parameters immediately post-surgery, with recovery evident by six weeks, while Garg et al. reported a significant TBUT reduction at one week, returning to baseline by one month.[2,15] These studies, however, did not distinguish between diabetic and non-diabetic patients. Identified risk factors contributing to dry eye included age, exposure to operating microscope light, and the duration of surgery.

Liu et al. highlighted slower TBUT recovery in diabetic patients compared to non-diabetic patients, which aligns with our findings.[16] Similarly, Shaaban and Aziz observed a significant TBUT decline in the first week, followed by improvement at one month, though values had not fully returned to baseline by three months, particularly in diabetics.[8] Our findings reinforce the immediate impact of cataract surgery on TBUT and variations in recovery timelines. While early declines are common, gradual improvement over subsequent months underscores the need for close monitoring and tailored dry eye management, particularly for diabetic patients with slower recovery.

Schirmer's test showed no significant changes from baseline to one week post-surgery in any group, suggesting a minimal immediate impact on tear production. However, by three months, non-diabetic patients showed significant improvement, highlighting recovery of baseline tear production levels as healing advances. A similar trend from one week to three months was observed across all groups, suggesting prolonged recovery facilitates better lacrimal gland function and tear secretion.[17,18] Zhang et al. found that diabetic patients had significantly lower preoperative TBUT and Schirmer's test scores than non-diabetics, with minimal post-surgical changes. Glycaemic control did not significantly influence these outcomes.[19] Additionally, the lack of significant differences in TBUT and Schirmer's test between the three groups at all time points suggests that while diabetic status influences subjective symptoms (OSDI), objective signs may not show the same degree of variance between groups. This finding emphasizes the complex nature

of dry eye disease in diabetic populations, where symptomatic relief may not correlate linearly with clinical tests.[20] Cung et al. reported similar findings, showing that approximately one-third of patients developed mild keratoconjunctivitis sicca following cataract surgery. Schirmer's test scores significantly decreased during the first week but returned to baseline by three months.[17]

These findings emphasize that while early postoperative dry eye symptoms are common, significant improvement occurs within three months across all groups. The slower recovery in TBUT for DR patients suggests microvascular and neuropathic changes associated with diabetic eye conditions, which could delay or modify the response to ocular surgery.[21] Post-surgical dry eye is influenced by multiple factors. During phacoemulsification, exposure of the corneal surface and conjunctiva to the operating microscope's light and air can lead to tear film instability.[21,22] Surgical incisions can damage corneal nerves, causing denervation and resulting in changes to the ocular surface. Additionally, mydriatic agents and topical anaesthesia can decrease tear production and promote tear film instability.[17,22] Inflammation following surgery can elevate cytokine levels on the ocular surface, which undermines tear film stability and intensifies dry eye symptoms.[9] Oxidative stress during surgery may damage corneal epithelium and conjunctival cells, worsening dry eye, especially in diabetic patients. The prolonged use of corticosteroids and antibiotics post-surgery may impact gland function, contributing to tear film disruption. Corneal nerve damage may also lead to neurotrophic keratopathy, which reduces corneal sensitivity and interferes with the blink reflex, limiting effective tear distribution and prolonging dry eye symptoms.[19] Environmental and mechanical factors, including frequent blinking during recovery and exposure to air conditioning, add stress to the tear film. For diabetic patients, delayed wound healing and persistent inflammation can exacerbate these challenges to the ocular surface.[18]

Differences in the diabetes duration and HbA1c levels were noted between the DR and no DR groups. To account for these variables, repeated measures ANCOVA were performed for all dry eye parameters, but no significant differences were found. However, the DR group consistently exhibited slightly poorer scores, suggesting a trend toward greater ocular surface dysfunction. Similarly, Yu et al. found more pronounced tear film dysfunction in patients with proliferative diabetic retinopathy than in those with non-proliferative diabetic retinopathy highlighting the increased risk of dry eye in advanced diabetic complications.[18] Unlike our findings, their study reported significant differences between the DR and no DR groups.

Unlike previous studies that primarily focused on early postoperative dry eye symptoms, our findings demonstrate a significant improvement in dry eye parameters as early as three months post-phacoemulsification. Most studies assessed recovery within a shorter follow-up period of less than three months.[2,12,19] Elminshawy et al. reported that TBUT levels exceeded baseline values in both diabetic and non-diabetic groups, further supporting the potential for ocular surface recovery.[4] This suggests that postoperative dry eye symptoms may be transient and reversible with appropriate management and care.

This study's three-month follow-up may not fully capture long-term changes in dry eye parameters or the sustained effects of diabetes and cataract surgery on ocular surface health. The reliance on OSDI, TBUT, and Schirmer's test limits evaluation by excluding patient-reported symptom burdens or advanced diagnostics like meibography and tear osmolarity. Furthermore, TBUT and Schirmer's test may not adequately address the inflammatory aspects of dry eye, which are particularly relevant in diabetic patients. Future studies should prioritise targeted postoperative interventions to enhance tear film stability, particularly in diabetic and retinopathy patients. Longer follow-up is needed to assess sustained ocular surface changes. Incorporating patient-reported outcomes and stratifying diabetic patients based on glycaemic control or disease severity could provide deeper insights. Advanced diagnostics, such as tear osmolarity and meibography, may further clarify dry eye pathophysiology and guide more effective management strategies.

CONCLUSIONS

Three months after phacoemulsification, significant improvements in OSDI scores, TBUT, and Schirmer's test values were observed, indicating a recovery in dry eye status. Diabetic patients experienced more pronounced early postoperative changes but demonstrated comparable recovery trends to non-diabetics by three months. Unlike previous studies, which primarily reported transient postoperative dry eye symptoms, our findings uniquely highlight significant long-term improvements. These findings highlight the importance of monitoring dry eye parameters in diabetics, particularly during the early postoperative period, to optimise outcomes and patient satisfaction.

ACKNOWLEDGEMENT

We would like to thank Associate Professor Dr Siti Azrin Ab Hamid from the Unit of Biostatistics and Research Methodology, School of Medical Sciences, Universiti Sains Malaysia for her valuable assistance with data analysis.

REFERENCES

1. Ishrat S, Nema N, Chandravanshi SCL. Incidence and pattern of dry eye after cataract surgery. Saudi J Ophthalmol 2019; 33(1): 34-40.
2. Garg P, Gupta A, Tandon N, Raj P. Dry eye disease after cataract surgery: study of its determinants and risk factors. Turk J Ophthalmol 2020; 50(3): 133-42.
3. Jiang D, Xiao X, Fu T, Mashaghi A, Liu Q, Hong J. Transient tear film dysfunction after cataract surgery in diabetic patients. PLoS One 2016; 11(1): e0146752.
4. Elminshawy AI, Hegazy HS, Aly MM. Comparison of corneal changes after phacoemulsification in diabetic and nondiabetic eyes. Al-Azhar Assiut Med J 2022; 20: 134-141.
5. Wat N, Wong RLM, Wong IYH. Associations between diabetic retinopathy and systemic risk factors. Hong Kong Med J 2016; 22(6): 589-99.
6. Sriprasert I, Warren DW, Mircheff AK, Stanczyk FZ. Dry eye in postmenopausal women: a hormonal disorder. Menopause 2016; 23(3): 343-51.

7. Din NM, Sa'aid SHB, Shen LC, Shu YS, Rahman MBA, Azmi ABM, et al. Hormone replacement therapy and dry eye in postmenopausal women: study in a tertiary centre in Malaysia. Int J Med Students 2013; 1(1): 12-5.
8. Shaaban YM, Aziz BF. Tear film assessment before and after phacoemulsification in patients with age-related cataracts. BMC Ophthalmol 2024; 24(1): 280.
9. Kasetsuwan N, Satitpitakul V, Changul T, Jariyakosol S. Incidence and pattern of dry eye after cataract surgery. PLoS One 2013; 8(11): e78657.
10. Li XM, Hu L, Hu J, Wang W. Investigation of dry eye disease and analysis of the pathogenic factors in patients after cataract surgery. Cornea 2007; 26(9 Suppl 1): S16-20.
11. Cetinkaya S, Mestan E, Acir NO, Cetinkaya YF, Dadaci Z, Yener HI. The course of dry eye after phacoemulsification surgery. BMC Ophthalmol 2015; 15(1): 68.
12. Zamora MG, Caballero EF, Maldonado MJ. Short-term changes in ocular surface signs and symptoms after phacoemulsification. Eur J Ophthalmol 2020; 30(6): 1301-7.
13. Wolffsohn JS, Arita R, Chalmers R, Djalilian A, Dogru M, Dumbleton K, et al. TFOS DEWS II diagnostic methodology report. Ocular Surf 2017; 15(3): 539-74.
14. Lajmi H, Chelly Z, Choura R, Mansour KB, Hmaied W. Relationship between OSDI score and biomicroscopic ocular surface damages in glaucomatous patients treated with preserved antiglaucomatous eye drops. J Fr Ophtalmol 2021; 44(9): 1326-31.
15. Kohli P, Arya SK, Raj A, Handa U. Changes in ocular surface status after phacoemulsification in patients with senile cataract. Int Ophthalmol 2019; 39(6): 1345-53.
16. Liu X, Gu YS, Xu YS. Changes of tear film and tear secretion after phacoemulsification in diabetic patients. J Zhejiang Univ Sci B 2008; 9(4): 324-8.
17. Cung LX, Nga NTT, Nga DM, Hiep NX, Pham DT. Cataract surgery destabilises temporary the tear film of the ocular surface. Klin Monbl Augenheilkd 2021; 238(3): 282-7.
18. Yu L, Chen X, Qin G, Xie H, Lv P. Tear film function in type 2 diabetic patients with retinopathy. Ophthalmologica 2008; 222 (4): 284-91.
19. Zhang K, Zhang S, Yu J, Lu Y, Zhu X. Changes of the tear film lipid layer thickness after cataract surgery in patients with diabetes mellitus. Acta Ophthalmol 2021; 99(2): e202-8.
20. Mencucci R, Vignapiano R, Rubino P, Favuzza E, Cantera E, Aragona P, et al. Iatrogenic dry eye disease: dealing with the conundrum of post-cataract discomfort. Ophthalmol Ther 2021; 10(2): 211-23.
21. Zhang X, Zhao L, Deng S, Sun X, Wang N. Dry eye syndrome in patients with diabetes mellitus: prevalence, etiology, and clinical characteristics. J Ophthalmol 2016; 2016: 8201053.
22. Cho YK, Kim MS. Dry eye after cataract surgery and associated intraoperative risk factors. Korean J Ophthalmol 2009; 23(2): 65-73.

Treating newly diagnosed Diffuse Large B-cell Lymphoma in the elderly patient with R-mini-CHOP: A single centre analytical retrospective observational study

Lai Nai Lim, MMed (Malaya)[1], Bee Ping Chong, MMed (Malaya)[2]

[1]Department of Medicine, Hospital Kuala Lumpur, [2]Department of Medicine, Pusat Perubatan Universiti Malaya

ABSTRACT

Introduction: Diffuse large B-cell lymphoma (DLBCL) forms the bulk of non-Hodgkin lymphoma (NHL) cases encountered in clinical practice among the elderly. For the majority of cases of DLBCL, treatment comprising of Rituximab, Cyclophosphamide, Doxorubicin, Vincristine and Prednisolone (R-CHOP) is suggested as first line chemotherapy. However, chemotherapy in the elderly population may be hampered by multiple factors, including reduced bone marrow reserves, significant comorbidities, and greater side effects from chemotherapy. Treatment as such aims to offer disease control and prolong life whilst minimising treatment related complications in this group of patients. Treatment with R-mini-CHOP, a reduced dose form of R-CHOP offers survival benefits and is recommended for treatment of elderly DLBCL patients and those who are frail. Our study examined local Malaysian experience of treating the newly diagnosed elderly DLBCL patient with R-mini-CHOP.

Materials and Methods: We retrieved retrospective data of all DLBCL patients aged >65 years old from the electronic medical records in Pusat Perubatan Universiti Malaya who received R-mini-CHOP. Treatment response was assessed by the overall response rate (ORR), defined as the proportion of patients attaining complete and partial remission after six cycles of treatment. We excluded patients with transformed lymphomas and relapsed refractory disease. For secondary analysis, we examined patients' treatment response according to their baseline demographic characteristics, development of complications during therapy as well as their survival in months from diagnosis.

Results: Our study identified 33 patients in the period of January 2017 till June 2023. The mean age of the sample cohort was 78 years old (Range from 66 to 86 years old). Majority of the samples had advanced stage lymphoma at initial diagnosis with n=21/33 (63.6%) having stage III and IV disease. At the end of treatment, one patient did not have assessment scans and hence was excluded from analysis. n=16/32 patients (50.0%) had attained ORR when analysed by intention to treat, n=14/32 (43.7%) attained complete response and n=2/32 (6.25%) attained partial response. When analysed for treatment response, those who attained ORR were more likely to have Stage 1 or 2 disease (p value = 0.028) and had statistically significant increased progression free survival (28.5 vs 5.5 months, p value <0.01) and overall survival (28.5 vs 9.0 months, p value = 0.03) compared to those who did not attain ORR. In terms of treatment associated complications, n=9/32 (28.1%) of patients developed severe infection necessitating hospitalization, n=14/32 (43.7%) developed at least Grade 2 and above cytopenias, and n=13/32 (41.6%) developed some other adverse side effects, most of which were mild to moderate in terms of severity.

Conclusion: The ORR for our patients treated with R-mini-CHOP was lower than other cohorts. We hypothesise that R-mini-CHOP alone may not offer adequate lymphoma control in our sample, especially for treatment of advanced stage DLBCL. Age alone is not an objective assessment of suitability for treatment; therefore, we suggest the use of geriatric prognostication tools to better ascertain patient groups who would benefit from full dose R-CHOP chemotherapy to improve response and survival.

KEYWORDS:
Lymphoma, Large B-cell, Diffuse, Aged, R-mini-CHOP

INTRODUCTION

Non-Hodgkin lymphoma (NHL) comprise a large group of lymphoproliferative disorders affecting a wide spectrum of patients. In the elderly population, diffuse large B-cell lymphoma (DLBCL) forms the bulk of NHL cases encountered in clinical practice. DLBCL is a high-grade lymphoma that commonly presents with a rapidly enlarging mass commonly of nodal origin, but there are cases of DLBCL arising from extranodal and extramedullary tissues in any part of the body.

Management of this heterogenous group of patients in the geriatric population poses a challenge for treating physicians and often requires shared decision making between physician and patient. For the majority of cases of DLBCL patients, chemotherapy with R-CHOP has long been suggested as first line chemotherapy.[1] However, chemotherapy in the elderly population may be hampered by multiple factors, including reduced bone marrow reserves, significant comorbidities (e.g. Heart failure precluding use of anthralcycline-based chemotherapy, renal and hepatic impairment possibly requiring dose adjustment of chemotherapy drugs), and higher morbidity from side effects

This article was accepted: 25 March 2025
Corresponding Author: Lai Nai Lim
Email: jere_lai91@hotmail.com

of chemotherapy (e.g. Chemotherapy using vinca-alkaloid based treatments resulting in peripheral neuropathy).[2] Treating the frail and elderly DLBCL patient with a reduced dose of chemotherapy seeks to offer adequate disease control and prolong survival whilst minimising treatment related toxicities.

Treatment with R-mini-CHOP involves administering chemotherapy at a pre-specified lower dose (Approximately 50% reduction in the dose of Cyclophosphamide, Doxorubicin, and Vincristine). In its pivotal single arm prospective trial in 2010, R-mini-CHOP showed survival benefit in patients above 80 years old, and offered a good compromise between treatment efficacy and safety. After a median follow up of 20 months (Range 0-45) in N=149 patients, the trial reported a median overall survival of 29 months and progression free survival of 21 months. Fifty eight deaths occurred in the cohort (n=58/149, 38.9%) for which n=33/58 (56.8%) were attributable to disease progression and n=12/58 (20.7%) due to treatment related complications. Overall response rate (ORR) was achieved in n=109/149 (73%) of patients. In terms of treatment toxicity, n=59/149 (39.5%) developed severe neutropenia and n=11/149 (7.3%) developed febrile neutropenia, n=56/149 (37.5%) developed thrombocytopenia and n=133/149 (89.2%) had anaemia, most of which were Grade 1-2 in terms of severity.[3] Since then, R-mini-CHOP has been adopted as a treatment modality in guidelines for patients above 80 years old or in those younger than 80 years old but with other significant comorbidities or impaired performance status.[4-6]

Our study is a retrospective observational study looking at the characteristics and outcomes of treating elderly DLBCL patients with R-mini-CHOP in a local Malaysian population.

MATERIALS AND METHODS

We identified all DLBCL patients who had received R-mini-CHOP through a retrospective review of chemotherapy charts and daycare visits data. Patients younger than 65 years of age, those who had relapsed-refractory DLBCL or those with transformed lymphoma were excluded. The choice of patients aged 65 years old and above receiving R-mini-CHOP as opposed to 80 years old as per international guidelines was at the discretion of the treating haematologists, seeking to minimise treatment related toxicities. This arbitrary institutional practice takes into consideration that the Malaysian authorities use a cutoff point of 60 years chronological age to define the 'older persons'.[7]

Patient information was then retrieved from the electronic medical records at Pusat Perubatan Universiti Malaya. The diagnosis of DLBCL was confirmed through histological and appropriate immunophenotyping testing. Data was collected for each patient by manual review of patient records – Age, sex, comorbidities at diagnosis, baseline Eastern Cooperative Oncology Group (ECOG) performance score, baseline lactate dehydrogenase (LDH) level, lymphoma staging at diagnosis, extranodal disease involvement at diagnosis, date of treatment, treatment received, number of treatment cycles received, treatment response at interim and end of treatment, development of cytopenias and sepsis, adverse outcomes from treatment, months of survival from start of treatment, and death as of 31/12/2023. Data on survival was censored after 31/12/2023.

Response was assessed by the ORR according to the Lugano classification, defined as the proportion of patients attaining complete and partial remission after six cycles of treatment.[8] Assessment of response was by demonstration of reduction in size and uptake of fluorodeoxyglucose F18 (FDG) by diseased tissue on imaging with either fluorodeoxyglucose-18 positron emission tomography (FDG-PETCT) or computed tomography (CT) scan. During the course of treatment, should a patient develop complications from therapy, the adverse event was recorded and severity assessed using the Common Terminology Criteria for Adverse Events (CTCAE) version 5 criteria.[9] The CTCAE grading is a five-level scale from 1 to 5: Grade 1 mild, Grade 2 moderate, Grade 3 severe, Grade 4 life threatening, and Grade 5 death related to adverse events.

Data analysis was performed with SPSS software version 27. Patient demographics was expressed with descriptive statistics. At end of treatment, patients who attained ORR (Responders) were compared with those who did not attain ORR (Non-responders) for differences in terms of baseline demographics and treatment related complications. Differences between responders and non-responders were assessed by Chi-square test/ Fishers exact test for categorical data and respective parametric/non parametric test for scale data. Where appropriate, an odds ratio was calculated at significance level $\alpha = 0.05$. Survival data was expressed as median number of months of survival. Progression free survival is defined as the duration of survival till disease relapse/ progression or death of patient. Overall survival is the duration of survival till death from any cause.

RESULTS

A total of 36 patients were identified from January 2017 till June 2023. Two patients were excluded as they had transformed lymphoma. Of the 34 patients identified who received treatment with R-mini-CHOP, one had refractory disease and hence was excluded from analysis (Figure 1).

The mean age of our sample cohort was 78 years old (Range 66 to 88 years old). n=13/33 (39.3%) of the patients were female. Majority of patients had some comorbid medical condition at diagnosis of DLBCL with hypertension in n=18/33 patients (54.5%), diabetes mellitus in n=12/33 patients (36.3%), heart disease in n=9/33 patients (27.2%) and chronic kidney disease in n=3/33 patients (9.1%). Notably, n=19/33 (57.6%) of patients had two and more underlying comorbidities at initial diagnosis of DLBCL. At initial diagnosis, n=21/33 of patients (63.6%) had stage 3 and 4 disease and extranodal disease was present in n=25/33 patients (75.7%).

n=22/33 of patients (66.7%) receiving R-mini-CHOP completed at least six cycles of treatment (Figure 2(i)). One patient was lost to follow up before assessment imaging and was excluded from our final analysis. Of the remaining patients who did not complete treatment, n=5/10 (50.0%)

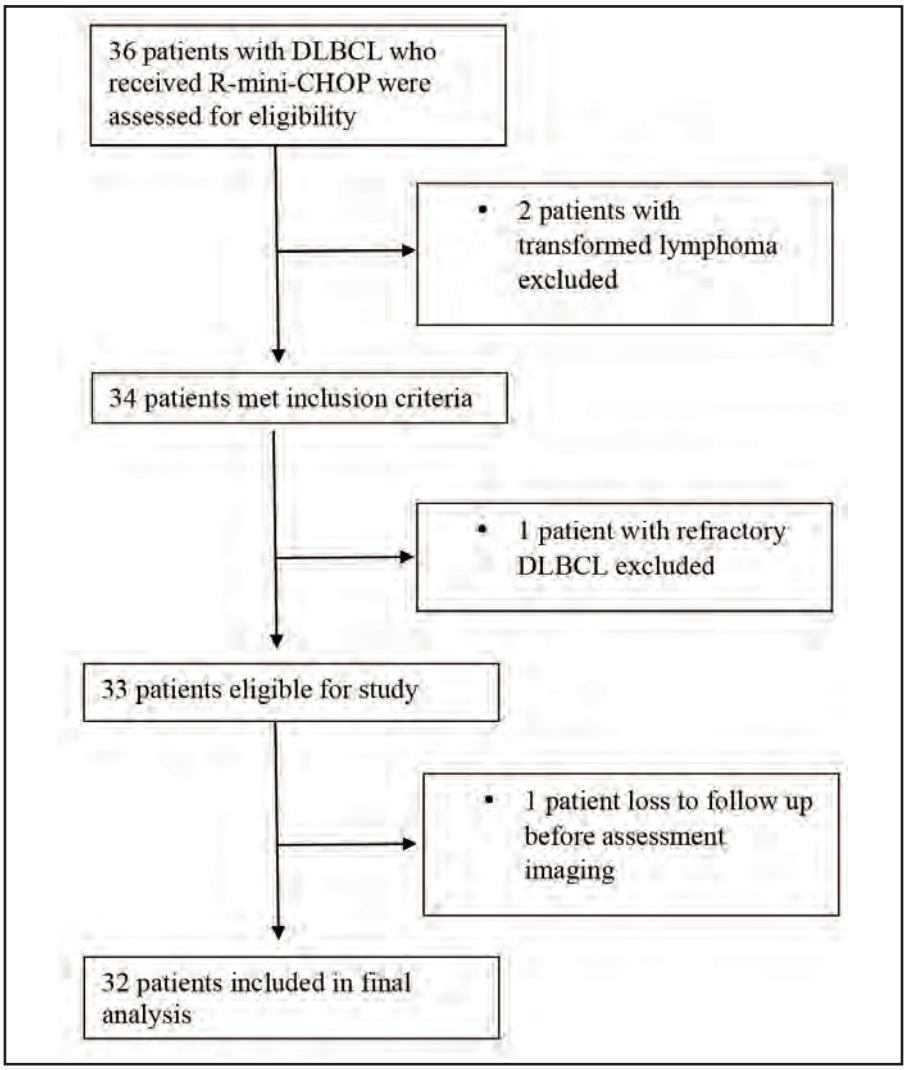


Fig. 1: Screening and classification of patients

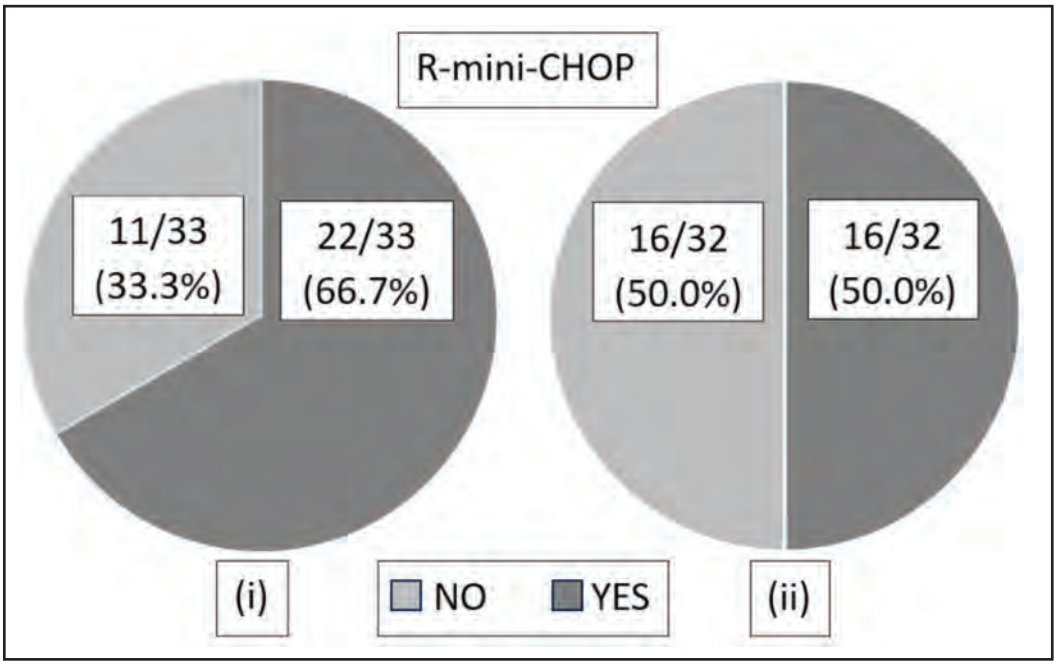


Fig. 2: (i) Proportion of patients receiving R-mini-CHOP completing at least 6 cycles of treatment. (ii) Proportion of patients receiving R-mini-CHOP attaining ORR

Table I: Comparison between treatment non-responder and responders for R-mini-CHOP in terms of baseline demographic characteristics

	Non Responder (n= 16)	Responder (n=16)	Odds Ratio [95% confidence interval]	p-value[a]
Mean age[b] – year (Standard deviation)	76.7 (4.95)	78.5 (4.02)	-	0.264
Gender – female n (%)	7 (43.8)	6 (37.5)	0.77 [0.19-3.17]	0.719
Comorbidities at diagnosis				
Diabetes mellitus – n (%)	5 (31.3)	6 (37.5)	1.32 [0.31-5.70]	0.710
Hypertension – n (%)	9 (56.3)	8 (50.0)	0.78 [0.19-3.13]	0.723
Cardiovascular disease[c]- n (%)	3 (18.8)	5 (31.3)	1.97 [0.38-10.17]	0.685
Chronic kidney disease[d]- n (%)	2 (12.5)	1 (6.3)	0.47 [0.04-5.73]	>0.995
ECOG performance status[e]				
≤ 2 – n (%)	15 (93.8)	16 (100)	-	>0.995
3-4 – n (%)	1 (6.3)	0 (0)		
Median Lactate dehydrogenase (LDH) at diagnosis – IU/L (Interquartile range, IQR)	293 (350)	209 (142)	-	0.105[f]
Extranodal involvement at diagnosis – n (%)	14 (87.5)	10 (62.5)	0.24 [0.04- 1.43]	0.22
Stage of disease at diagnosis				
1-2 – n (%)	3 (18.8)	9 (56.3)	5.57 [1.13-27.52]	0.028
3-4 – n (%)	13 (81.3)	7 (43.8)	0.18 [0.04-0.89]	

[a] Unless specified, difference between non – responders and responders were tested using Chi square test or Fisher exact test at statistical significance p< 0.05
[b] Independent samples T test was used to compare means in sample
[c] Cardiovascular disease was defined as having previously diagnosed with ischemic heart disease, heart failure, valvular heart disease, arrhythmia
[d] Chronic kidney disease was defined as having previous glomerular filtration rate <60 ml/min/1.73m2 for more than 3 months
[e] Eastern Cooperative Oncology Group performance status at diagnosis
[f] Independent samples Mann-Whitney U test used to compare medians in sample

Table II: Comparison between treatment non-responder and responders for R-mini-CHOP in terms of treatment related complications and survival in months

	Non Responder (n= 16)	Responder (n=16)	Odds Ratio [95% confidence interval]	p-value
Developed cytopenia during treatment[a] - n (%)	8 (50.0)	6 (37.5)	0.60 [0.15-2.46]	0.476
Developed sepsis or hospitalised for infection during treatment – n (%)	6 (37.5)	3 (18.8)	0.39 [0.08-1.93]	0.433
Developed other severe adverse effects during treatment[b] – n(%)	3 (18.8)	1 (6.3)	0.29 [0.03-3.13]	0.600
Death[c] – n (%)	10 (62.5)	1 (6.3)	0.04 [0.01-0.39]	<0.01
Progression free survival – median months (IQR)	5.5 (4)	28.5 (23)	-	<0.01[d]
Overall survival – median months (IQR)	9.0 (28)	28.5 (22)	-	0.03[d]

[a] Cytopenias defined as haemoglobin ≤ 10g/dL, absolute neutrophil count < 1 x 10^9/L, platelet < 150 x 10^9/L or more than one of the above defined cytopenias during the course of treatment
[b] Severe adverse effects was defined as ≥ Grade 3 according to CTCAE Version 5
[c] Death as of 31/12/2023
[d] Independent samples Mann-Whitney U test used to compare medians in sample

were due to disease progression, n=2/10 (20.0%) due to drug intolerance, and the remaining n=3/10 (30.0%) had completed treatment as determined by the treating physician. At end of treatment, n=16/32 patients (50.0%) had attained ORR by intention to treat analysis; n=14/32 (43.7%) attained complete response and n=2/32 (6.25%) attained partial response (Figure 2(ii)). When analysis was limited only to the 22 patients who completed six cycles of treatment; limiting effects from suboptimal drug administration either due to intolerance or change in chemotherapy regime, n=12/22 (54.5%) had attained complete remission, n=2/22 (9.1%) had partial remission, giving an adjusted ORR of 63.6%.

Table I demonstrates comparison between non-responders and responders in terms of baseline demographic characteristics. Among the various variables, stage of disease at diagnosis was found to correlate with treatment response whereby responders had a higher proportion of patients with Stage 1 and 2 disease (p-value = 0.028). There was no significant difference between non-responders versus responders in terms of mean age, gender, comorbidities at diagnosis, baseline LDH, extranodal involvement, or baseline ECOG performance status.

Table II illustrates safety outcomes comparing treatment related complications between non-responders and responders as well as associated survival outcomes. n=14/32 (43.7%) developed at least Grade 2 and above cytopenias, n=9/32 (28.1%) of patients developed severe infection necessitating hospitalization, and n=13/32 (41.6%) developed some other adverse side effects while undergoing treatment, from which n=4 were Grade 3 and above in terms of CTCAE classification. Most of these adverse effects were

mild to moderate (Grade 1-2) diarrhoea, vomiting, paraesthesia, and alopecia; however there was 1 episode of Grade 4 upper gastrointestinal bleeding, 1 episode of Grade 3 vomiting and 2 Grade 3 thromboembolic events. There was no statistically significant difference between responders and non-responders in terms of development of cytopenias [n=6 (37.5%) vs n=8 (50.0%), p-value = 0.476], development of sepsis [n=3 (18.8%) vs n=6 (37.5%), p-value = 0.433] and development of other Grade 3 above complications [n=1 (6.3%) vs n=3 (18.8%), p-value = 0.600].

In terms of survival, the median progression free survival in months for the entire cohort was 13 (IQR 27) and median overall survival was 26 (IQR 26). Responders had statistically significant increased progression free survival (28.5 vs 5.5 months, p-value <0.01) and overall survival (28.5 vs 9.0 months, p-value = 0.03) compared to non-responders (Table II).

DISCUSSION

In our sample cohort, majority of patients completed six cycles of treatment, n= 22/33 (66.7%) but only n= 16/32 (50.0%) attained ORR. In a retrospective review of the Dutch cancer registry, Al-Sarayfi et al (10) examined response of patients receiving R-mini-CHOP (Median number of cycles six, range one to eight) against a propensity matched cohort receiving full dose R-CHOP. Our sample ORR for R-mini-CHOP is lower than the ORR of 72% in the aforementioned study.[10] We are unable to ascertain causality for the lower ORR recorded in our cohort. However, as the reason for patients not completing treatment in our cohort was mainly related to disease progression, n=5/10 (50.0%), we hypothesise that R-mini-CHOP may not offer adequate lymphoma control in our sample, especially for treatment of advanced stage DLBCL.

The study by Al-Sarayfi et al (10) attempted to determine if the better tolerability of R-mini-CHOP would be offset by possible reduced disease response. While there was reported poorer ORR, progression free survival and overall survival in the R-mini-CHOP cohort as compared to the R-CHOP cohort, the study did not report the occurrence of treatment related complications or toxicities in the sample, which may hypothetically be higher in the R-CHOP cohort.[10]

When compared with the earlier Phase 2 trial by Peyrade et al (3), our cohort had lower progression free survival (13 vs 21 months) but comparable overall survival (26 vs 29 months).[3] Death occurred in n=11/32 (34.3%) of our patient cohort similar to the aforementioned study (38.6%).

In our cohort, patients who responded to R-mini-CHOP had improved progression free survival (28.5 vs 5.5 months) and overall survival (28.5 vs 9.0 months) compared to non-responders. In terms of treatment toxicity, n=14/32 (43.7%) of our patients developed at least Grade 2 and above cytopenias and n=9/32 (28.1%) developed sepsis requiring hospitalisation. This finding supports earlier and more aggressive supportive interventions with blood products transfusion and the use of prophylactic granulocyte colony stimulating factors during treatment. For those at high risk of developing infective complications, a consideration should be given for antimicrobial prophylaxis therapy.[11]

We have identified several limitations in our study. Firstly, our small sample size of N=33 patients may be underpowered to detect true effect sizes. Furthermore, assessment of patient's performance status or frailty was done subjectively by the treating haematologist using only the patient's ECOG score. Patients who are more frail at diagnosis may have been selected for treatment with R-mini-CHOP, hence negatively impacting the treatment outcomes.

Due to the retrospective nature of our study design, there is a chance of biases introduced in our study. Misclassification bias can happen during data collection where variables can be inappropriately coded. Selection bias can be introduced as patients were sampled only from a single medical centre in an urban setting, and results may not be generalisable to other patient populations. Study outcomes may be affected by other unsampled variables leading to confounding, and a causal link between treatment and outcomes cannot be determined.[12]

There is a lack of information of the DLBCL cell of origin or genetic profiling, which increasingly is shown to affect disease outcomes.[13] It is unclear if the lower than expected ORR for our patient cohort might possibly be due to higher prevalence of adverse genetic and molecular profiles in our population.

The use of R-mini-CHOP in a younger patient cohort than that reported in literature and guidelines (65 years old vs 80 years old) may offer explanation for the lower ORR seen in our study, as the attenuated treatment may not offer adequate disease control compared to full dose chemotherapy. Lastly, due to the widespread disruption of health services during the COVID-19 pandemic, it is unclear from our study how the pandemic had impacted treatment decisions, follow up or outcomes.

Moving forward, further dedicated geriatric oncology studies are warranted to improve outcomes for this diverse group of patients, whom unfortunately are underrepresented in many clinical trials.[14-15] To improve assessment of elderly patients for selection of treatments, use of screening tools like G8, Elderly Prognostic Index or Clinical Frailty Score should be encouraged, with frail patients identified then referred to a geriatrician for comprehensive assessment.[14-15] Proper patient selection for treatment is important as fit elderly patients benefit similarly from curative treatment regimes such as full dose R-CHOP chemotherapy as compared to younger patients.[16] A follow up prospective cohort study of all elderly DLBCL patients treated with the various chemotherapy regimes including R-CHOP and R-mini-CHOP could better identify patient profiles best suited for each treatment arm.

CONCLUSION

Whilst preliminary, our findings suggest that R-mini-CHOP may not adequately control DLBCL, especially in patients who present with advanced stage disease. As such, due consideration should be given in selecting suitable patients

who may otherwise benefit from full dose R-CHOP chemotherapy in the newly diagnosed elderly patient with DLBCL.

Our study highlights part of the complexities of managing malignancy in a diverse and possibly frail patient population, emphasising that there is no 'one size fits all' treatment. We concur with reported literature highlighting the role of geriatric assessments in a multi-disciplinary team to better individualise treatment. We also report Malaysian data, which could serve as a benchmark against other reported patient outcomes elsewhere.

ACKNOWLEDGEMENTS

The authors have no conflict of interests to declare. We would like to thank respective faculty members from the Department of Medicine and Haematology Unit of Pusat Perubatan Universiti Malaya for their assistance in proofreading this manuscript.

REFERENCES

1. Morrison VA, Hamlin P, Soubeyran P, Stauder R, Wadhwa P, Aapro M, et al. Approach to therapy of diffuse large B-cell lymphoma in the elderly: the International Society of Geriatric Oncology (SIOG) expert position commentary. Ann Oncol 2015; 26(6): 1058-68.
2. Morrison VA, Hamlin P, Soubeyran P, Stauder R, Wadhwa P, Aapro M, et al. Diffuse large B-cell lymphoma in the elderly: impact of prognosis, comorbidities, geriatric assessment, and supportive care on clinical practice. An International Society of Geriatric Oncology (SIOG) expert position paper. J Geriatr Oncol 2015; 6(2): 141-52.
3. Peyrade F, Jardin F, Thieblemont C, Thyss A, Emile JF, Castaigne S, et al. Attenuated immunochemotherapy regimen (R-miniCHOP) in elderly patients older than 80 years with diffuse large B-cell lymphoma: a multicentre, single-arm, phase 2 trial. Lancet Oncol 2011; 12(5): 460-68.
4. Fox CP, Chaganti S, McIlroy G, Barrington SF, Burton C, Cwynarski K, et al. The management of newly diagnosed large B-cell lymphoma: A British Society for Haematology guideline. Br. J. Haematol 2024; 204(4): 1178-92.
5. Tan D, Tan SY, Lim ST, Kim SJ, Kim WS, Advani R, et al. Management of B-cell non-Hodgkin lymphoma in Asia: resource-stratified guidelines. Lancet Oncol 2013; 14(12): e548-61.
6. Guerra VA, Ocampo M, Cusnir M. Outcomes of R-Mini-CHOP therapy for elderly patients with diffused large B cell lymphoma (DLBCL) and high-grade follicular lymphoma (FL). Blood 2021; 138(Supplement 1): 4581.
7. National Digital Department of Malaysia. General information on senior citizen healthcare [cited February 2025]. Available from: https://www.malaysia.gov.my/portal/content/27752
8. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: The Lugano classification. J Clin Oncol 2014; 32(27): 3059-67.
9. U.S. Department of Health and Human Services. Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, November 27, 2017 [cited December 2023]. Accessed from: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_60
10. Al-Sarayfi D, Brink M, Chamuleau MED, Brouwer R, van Rijn RS, Issa D, et al. R-miniCHOP versus R-CHOP in elderly patients with diffuse large B-cell lymphoma: A propensity matched population-based study. Am J Hematol 2024; 99(2): 216-22.
11. Kossow KW, Bennett JG, Hoffmann MS. Antimicrobial prophylaxis in lymphoma by chemotherapy regimen. J. Cell. Immunol 2023; 5(4).
12. Talari K, Goyal M. Retrospective studies - utility and caveats. J R Coll Physicians Edinb 2020; 50(4): 398-402.
13. Cutmore NH, Krupka JA, Hodson DJ. Genetic profiling in diffuse large B-cell lymphoma: The promise and the challenge. Mod Pathol 2023; 36(1): 100007.
14. Sedrak MS, Freedman RA, Cohen HJ, Muss HB, Jatoi A, Klepin HD, et al. Older adult participation in cancer clinical trials: A systematic review of barriers and interventions. CA Cancer J Clin 2021; 71(1): 78-92.
15. Abel GA, Klepin HD. Frailty and the management of hematologic malignancies. Blood 2018; 131(5): 515-24.
16. Merli F, Pozzi S, Catellani H, Barbieri E, Luminari S. The role of geriatric assessment in the management of diffuse large B-cell lymphoma. Cancers (Basel) 2023; 15(24): 5845.

The impact of facial acne scars on quality of life, anxiety, depression and its associated risk factors

Mabel Meibo Heah, MRCP[1], Ru Shing Ng, MPharm[2,3], Dy-Win Low, Adv M Derm[4], Norazlima Mohd Ali, Adv M Derm[1], Adawiyah Jamil, Adv M Derm[5]

[1]Department of Dermatology, Hospital Pulau Pinang, Georgetown, Penang, Malaysia, [2]Clinical Research Centre, Hospital Pulau Pinang, Georgetown, Penang, Malaysia, [3]Institute for Clinical Research, National Institute of Health, Shah Alam, Selangor, Malaysia, [4]Sunway Medical Centre Penang, Seberang Jaya, Penang, Malaysia, [5]Department of Medicine, Hospital Tuanku Muhriz Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

ABSTRACT

Introduction: Acne scars negatively impact psychosocial and emotional wellbeing. However, data on the impact of acne scarring on anxiety and depression as well as quality of life are limited. This study assessed the effects of facial acne scars on quality of life, anxiety, and depression, and identifies risk factors associated with scar severity.

Materials and Methods: We conducted an observational cross-sectional study between February 2023 and January 2024 at dermatology clinics in two public hospitals. A total of 175 patients with facial acne scars were recruited. Data collection included patient demographics, acne severity, scar severity (SCAR-S) and questionnaires such as Dermatology Life Quality Index (DLQI) and Hospital Anxiety and Depression Scale (HADS). The association between scar severity, quality of life, anxiety, and depression was analyzed using Chi-square tests and ordinal logistic regression.

Results: The median age of participants was 26.9 years and 56% were female. Most patients had mild (37.7%) to moderate (28.6%) acne scars. Among patients with severe/very severe scars, 69.2% reported a significant impact on quality of life (DLQI>10). A significant association was observed between scar severity and anxiety (p=0.009) as well as depression (p<0.001). There was a positive correlation between HADS and DLQI scores (r=0.602, p<0.001). Delayed or absent treatment after acne onset was a significant risk factor for acne scar severity.

Conclusion: Facial acne scars are associated with impairment in quality of life and increased anxiety and depression. Timely and effective acne treatment is essential to reduce the severity of scarring and its psychosocial burden.

KEYWORDS:
Acne, scars, quality of life, anxiety, depression

INTRODUCTION

Acne vulgaris is a chronic inflammatory disease of the pilosebaceous unit that frequently affects adolescents and young adults. Estimated prevalence of acne ranged from 35% up to 100%, varying between countries and age groups.[1] Acne scar may be permanent depending on its severity and persists to adulthood leading to a negative impact on psychosocial and emotional wellbeing. However, little is known about the impact of acne scarring on psychosocial health due to limited literature reviews and data.

Early intervention of acne scars may prevent consequential psychosocial debilitation for our patients. Scarring can be a major concern for patients as it persists and is often undermined and overlooked. A few studies have investigated the impact of facial acne scarring on quality of life. These studies showed that quality of life deteriorates as the severity of acne scarring increases.[2-4] Most patients reported embarrassment, self-consciousness and poor self-esteem. To the best of our knowledge, the direct relationship between acne scarring with anxiety and depression has not been investigated.

Prevention remains a primary strategy for acne scarring management. Recognizing potential risk factors early is imperative in preventing post acne scar formation. While risk factors for the development of acne vulgaris have been widely studied, little is known about the risk factors for acne scar development. From the very limited literature available, risk factors such as acne severity, time between acne onset and first effective treatment, recurrent acne, gender, family predisposition and dietary and social habits play a role in acne scar severity.[5-8]

In this study, we aimed to determine the impact of facial acne scars on quality of life and its association with anxiety and depression. We aimed also to assess risk factors that may be associated with acne scar severity.

MATERIALS AND METHODS

Study Design

This observational cross-sectional study was conducted between February 2023 to January 2024 at the dermatology clinics in two tertiary public hospitals. All patients with acne scars were screened for eligibility. Patients with facial acne scars, aged between 18-55 years old with acne in remission either on or off treatment were included. Exclusion criteria were patients with active acne, post inflammatory erythema or hyperpigmentation and patients who were unable to complete the study questionnaires. Informed consent was

This article was accepted: 25 March 2025
Corresponding Author: Mabel Heah Meibo
Email: mabelheah_88@hotmail.com

obtained. Approval from the Medical Research and Ethics Committee (MREC) of the Ministry of Health (MOH) Malaysia was obtained prior to initiation of the study. The study was conducted in compliance with ethical principles outlines in the Declaration of Helsinki and Malaysian Good Clinical Practice guidelines.

Data on socio-demographics, history of acne and its treatment were recorded. Patients were required to grade their previous acne severity based on worst acne ever experienced. Standard photographs for the classification of acne severity which were developed by Hayashi et al. were shown to the patients.[9] There were 2 photographs each for mild, moderate, severe and very severe acne. Patients scored their acne severity during most active disease by referencing to these photographs.

Patients completed two self-administered questionnaires, the Dermatology Life Quality Index (DLQI) and Hospital Anxiety and Depression Scale (HADS) which were aimed to assess their response in relation to their current acne scars.[10,11] The DLQI comprised of 10 questions, each with four possible answers with scores from 0-3. The total overall score is 0-30, with higher scores indicating greater impairment in quality of life. The clinical interpretation of the DLQI scores are as follows: 0-1, no effect at all on the patient's life; 2-5, a small effect on the patient's life; 6-10, a moderate effect on the patient's life; 11-20, a very large effect on the patient's life and 21-30, an extremely large effect on the patient's life. The HADS questionnaire had a total of 14 questions, 7 for each anxiety and depression domain. The scores were summed to produce two subscales corresponding to anxiety and depression; 0-7 for normal; 8- 10 for borderline abnormal; 11-21 for abnormal. Participants were required to endorse a response based on their current acne scars.

All patients were examined by a single investigator to determine their acne scar severity based on the Scale for Acne Scar Severity (SCAR-S). SCAR-S is a validated grading tool, which include both atrophic and hypertrophic scars.[12] A score of 0 is for clear skin, 1 for almost clear, 2 for mild, 3 for moderate, 4 for severe and 5 for extremely severe. The types of scars; namely ice pick, boxcar, rolling, hypertrophic/keloid were documented. Patients may have more than one type of acne scar.

Statistical Analyses

Despite an extensive literature search, we found no similar studies which can be used as reference for our sample size calculation. The methodology and outcome measures of the existing literature differ from that of our proposed study. Thus, our proposed pilot study uses a sample size of 30 subjects per category of acne scar severity (ranging from 1 to 5). The minimum sample size needed was 150 subjects. A total of 175 patients were recruited.

Descriptive analysis summarized sociodemographic and clinical characteristics. Categorical variables were presented as frequency and percentage while continuous variables were expressed as median and interquartile range (IQR). The relationship between grades of acne scarring and HADS-Anxiety groups, HADS-Depression groups and DLQI groups were analysed using Chi-square test. Bonferroni correction was used for pairwise comparison. DLQI scores were categorized into 3 groups; score of 0 – 5, score 6 – 10 and score 11 – 30. HADS scores were categorized into 2 groups; score 0 – 7 and score 8 – 20 for both HADS-A and HADS-D. Pearson coefficient assessed the correlation between HADS and DLQI scores.

Ordinal logistic regression was used to model the association between acne scar severity and risk factors. Acne scar severity was categorized into three groups: almost clear and mild, moderate, and severe and very severe. Potential risk factors analysed included sociodemographic and clinical characteristics. This method was also employed to identify the risk factors associated with the development of different acne scar severity which were ordinally scaled. The model was appropriate. Results are presented in odds ratio and corresponding 95% confidence interval (CI). Univariate ordinal logistic regression was performed to evaluate factors associated with scar severity. A multivariate ordinal logistic regression was then conducted with all the 13 variables to control confounding factors.

RESULTS

Socio-Demographic and Clinical Characteristics

A total of 175 patients with facial acne scars were included in this study (Table I). The median age was 26.9 (12.1) years. Most of the participants were females (56%). The majority ethnic group was Malay, followed by Chinese and Indian/others. Most did not smoke (88.6%) nor consume alcohol (88.6%). Nuts, cereals and butter were least consumed; mostly never or occasionally, while most patients consumed fast food (60.6%) and milk (45.7%) at least once or twice a week.

Patient self-reported assessment showed that grade of acne severity during active disease were mostly moderate (42.3%), followed by severe (31.4%), mild (16%) and lastly; very severe (10.3%). Most patients either took more than a year to seek treatment (29.1%); or never sought treatment (29.1%) prior to presentation to our clinic. The majority of patients had mild (37.7%) to moderate (28.6%) acne scars, followed by severe (21.7%) and very severe (8%) scarring based on SCAR-S (Table II). The most common types of scars were ice pick (72.8%), rolling (57.1%), boxcar (30.9%) and hypertrophic/keloid (2.3%).

Impact of Facial Acne Scar Severity on Quality of Life (QOL)

Significant association was observed between the severity of acne scar and quality of life (Table III). Mean DLQI score was higher with worsening of scar severity. Overall, 69.2% of patients with severe and very severe scars reported a DLQI score > 10 (very / extremely large impact). A proportion of patients with mild acne scarring; 24.7% and 28% of patients with moderate scarring also reported a DLQI score > 10.

Relationship between Facial Acne Scar Severity on Anxiety and Depression

We observed a significant association between scar severity with anxiety and depression (Table III). Post-hoc test with Bonferroni correction resulted in a significance level set at p

Table I: Socio-demographic characteristics of the study population

Characteristics	N= 175 n (%) or median (IQR)
Age (year)	26.9 (12.1) [a]
Gender	
Female	98 (56.0)
Male	77 (44.0)
Race	
Malay	98 (56.0)
Chinese	62 (35.4)
Indian/ others	15 (8.5)
BMI (kg/m2)	22.8 (5.7) [a]
Smoking status	
No	155 (88.6)
Yes	15 (8.6)
Ex-smoker	5 (2.9)
Alcohol	
No	155 (88.6)
Yes	20 (11.4)
Diet – Nuts	
Never/ occasionally	89 (50.9)
Once or twice a week	75 (42.9)
Most or all days	11 (6.3)
Diet – Fast food	
Never/ occasionally	56 (32.0)
Once or twice a week	106 (60.6)
Most or all days	13 (7.4)
Diet – Cereals	
Never/ occasionally	84 (48.0)
Once or twice a week	66 (37.7)
Most or all days	25 (14.3)
Diet – Butter	
Never/ occasionally	82 (46.9)
Once or twice a week	78 (44.6)
Most or all days	15 (8.6)
Diet – Milk	
Never/ occasionally	46 (26.3)
Once or twice a week	80 (45.7)
Most or all days	49 (28.0)

BMI Body Mass Index, IQR Interquartile Range
[a] The distribution is skewed to the right

< 0.017. There was significant association with anxiety (χ^2 = 6.775; p = 0.009) and depression (χ^2 = 14.773; p < 0.001) when patients with almost clear/mild acne scars were compared with patients with severe and very severe acne scars. Marginally significant association in development of anxiety (χ^2 = 5.759; p = 0.016) and depression (χ^2 =6.069; p = 0.014) were observed in comparing patients with severe and very severe scars versus patients with moderate scars. There was no difference in anxiety (χ^2 = 0.001; p = 0.977) and depression (χ^2 = 1.314; p = 0.252) between patients with almost clear/mild scars and patients with moderate scars. In the anxiety domain, patients scored highest with questions of 'feeling tensed' and 'excessive worrying'; while within the depression domain, 'loss of interest' and 'depressed mood' scored the highest.

Correlation between HADS and DLQI scores

There was significant positive correlation between HADS total scores and DLQI total scores (r =0.602, p < 0.001). This suggests that the greater the HADS scores, the greater the DLQI scores.

Risk Factors for the Development of Facial Acne Scars

Thirteen potential risk factors were analysed using univariate ordinal logistic regression (Table IV). Alcohol consumption, grade of acne severity during active disease and delayed treatment initiation (>1 year) were significantly associated with acne scar severity. Factors such as gender, race, smoking status, acne sites, family history of acne/ acne scar and dietary habits did not show any statistical significant association with acne scar severity.

Multivariate logistic regression revealed that the odds of developing more severe acne scar were 37.61 (95% CI: 6.16 - 229.72), 89.39 (95% CI: 14.16 - 564.28) and 880.98 (95% CI: 88.73 - 8747.36) times higher among patients with moderate, severe and very severe acne compared with patients with mild acne. Patients who received treatment after a year of acne onset or who were never treated had 2.78 (95% CI:1.14 - 6.80) and 3.89 (95% CI: 1.44 – 10.49) times higher odds to develop more severe acne scars respectively; compared to those who received treatment within 6 months of acne onset.

Table II: Clinical characteristics of the study population

Characteristics	N= 175, n (%)
Grade of acne severity during active disease	
Mild	28 (16.0)
Moderate	74 (42.3)
Severe	55 (31.4)
Very severe	18 (10.3)
Sites affected by acne	
Face	109 (62.3)
Face & trunk	66 (37.7)
Duration of acne prior to seeking treatment (doctors'/ over-the-counter treatments)	
0 – 3 months	23 (13.1)
3 – 6 months	24 (13.7)
6 – 9 months	12 (6.9)
9 – 12 months	14 (8.0)
> 1 year	51 (29.1)
Never treated	51 (29.1)
Family history of acne/ acne scars	
No	73 (41.7)
Parent(s)	23 (13.1)
Sibling(s)	61 (34.9)
Parent(s) & sibling(s)	18 (10.3)
Grade of acne scar severity	
Almost clear	7 (4.0)
Mild	66 (37.7)
Moderate	50 (28.6)
Severe	38 (21.7)
Very severe	14 (8.0)
Quality of life, DLQI	
0 – 1 no effect	27 (15.4)
2 – 5 small effect	42 (24.0)
6 – 10 moderate effect	38 (21.7)
11 – 20 very large effect	60 (34.3)
21 – 30 extremely large effect	8 (4.6)
Anxiety, HADS-A score	
0 – 7 normal	103 (58.9)
8 – 10 borderline abnormal	42 (24.0)
11 – 21 abnormal	30 (17.1)
Depression, HADS-D score	
0 – 7 normal	120 (68.6)
8 – 10 borderline abnormal	25 (14.3)
11 – 21 abnormal	30 (17.1)

DLQI Dermatology Life Quality Index, HADS-A Hospital Anxiety and Depression Scale-Anxiety, HADS-D Hospital Anxiety and Depression Scale-Depression

Table III: Relationship between acne scar severity with quality of life / anxiety and depression

Parameters	Acne scar severity N = 175			p-value[b]
	Almost clear & mild n (%)	Moderate n (%)	Severe & very severe n (%)	
Quality of life, DLQI				
No / small effect	39 (53.4)	21 (42.0)	9 (17.3)	< 0.001
Moderate effect	16 (21.9)	15 (30.0)	7 (13.5)	
Very / extremely large effect	18 (24.7)	14 (28.0)	36 (69.2)	
Anxiety, HADS-A				
0 – 7 normal	48 (65.8)	33 (66.0)	22 (42.3)	0.015
8 – 21 abnormal	25 (34.2)	17 (34.0)	30 (57.7)	
Depression, HADS-D				
0 – 7 normal	59 (80.8)	36 (72.0)	25 (48.1)	< 0.001
8 – 21 abnormal	14 (19.2)	14 (28.0)	27 (51.9)	

DLQI Dermatology Life Quality Index, HADS-A Hospital Anxiety and Depression Scale-Anxiety, HADS-D Hospital Anxiety and Depression Scale-Depression
b Chi-square test for independence

Table IV: Factors associated with acne scar severity

Variables	β	Crude OR	(95% CI OR)	p-value
Gender				
Male	Ref.	1.00		
Female	-0.06	0.94	(0.54; 1.64)	0.835
Race				
Indian/ others	Ref.	1.00		
Malay	0.27	1.31	(0.46; 3.75)	0.612
Chinese	0.60	1.82	(0.61; 5.37)	0.281
Smoking status				
Ex-smoker	Ref.	1.00		
No	0.62	1.86	(0.32; 10.98)	0.485
Yes	1.81	6.08	(0.79; 46.55)	0.075
Alcohol				
Yes	Ref.	1.00		
No	-0.94	0.39	(0.16; 0.95)	0.035
Grade of acne severity during active disease				
Mild	Ref.	1.00		
Moderate	2.76	15.80	(3.50; 71.23)	<0.001
Severe	3.67	39.28	(8.49; 181.85)	<0.001
Very severe	5.26	192.95	(30.30; 1228.74)	<0.001
Sites affected by acne				
Face & trunk	Ref.	1.00		
Face	-0.21	0.81	(0.46; 1.42)	0.456
Time to seek treatment for acne				
0 – 6 months	Ref.	1.00		
6 – 12 months	0.14	1.15	(0.47; 2.82)	0.761
> 1 year	0.78	2.19	(1.05; 4.59)	0.039
Never treated	0.18	1.19	(0.57; 2.51)	0.641
Family history with acne/ acne scars				
Parent(s) & sibling(s)	Ref.	1.00		
Parent(s)	-0.46	0.63	(0.21; 1.88)	0.430
Sibling(s)	-0.15	0.86	(0.34; 2.15)	0.758
No	-0.11	0.89	(0.36; 2.20)	0.816
Diet – Nuts				
Most or all days	Ref.	1.00		
Once or twice a week	-0.14	0.87	(0.25; 3.04)	0.814
Never/ occasionally	-0.40	0.67	(0.19; 2.32)	0.493
Diet – Fast food				
Most or all days	Ref.	1.00		
Once or twice a week	0.55	1.73	(0.54; 5.52)	0.341
Never/ occasionally	1.12	3.06	(0.91; 10.27)	0.062
Diet – Cereals				
Most or all days	Ref.	1.00		
Once or twice a week	0.76	2.13	(0.87; 5.26)	0.088
Never/ occasionally	0.29	1.34	(0.56; 3.19)	0.501
Diet – Butter				
Most or all days	Ref.	1.00		
Once or twice a week	0.46	1.59	(0.55; 4.61)	0.383
Never/ occasionally	0.33	1.40	(0.48; 4.04)	0.528
Diet – Milk				
Most or all days	Ref.	1.00		
Once or twice a week	-0.09	0.91	(0.47; 1.79)	0.785
Never/ occasionally	-0.45	0.64	(0.30; 1.35)	0.238

B coefficients, CI Confidence Interval, OR Odds Ratio, Ref Reference

DISCUSSION

Impact of Acne Scar on Quality of Life

Acne scars have been shown to impact quality of life, with DLQI scores ranging from 5.6 to 6.2.[2-4] Our study revealed an overall mean DLQI score of 8.8, suggesting a moderate effect which is comparable to other skin diseases such as Behcet's syndrome (DLQI 5.7), Darier's disease (DLQI 5.89), Hailey-Hailey disease (DLQI 6.06), rosacea (DLQI 6.1) and cutaneous lupus erythematosus (DLQI 6.5).[13]

Our findings showed the impact on the domains of quality of life increases with acne scar severity. A reasonable explanation behind this is that treatment options for acne scars are limited, mostly time consuming and costly especially with physical modalities such as peels and lasers, microneedling, dermabrasion and platelet-rich plasma injections. They require multiple visits to the dermatologist and pose a great financial burden and in patients who are severely scarred, results may be less than optimal. In terms of

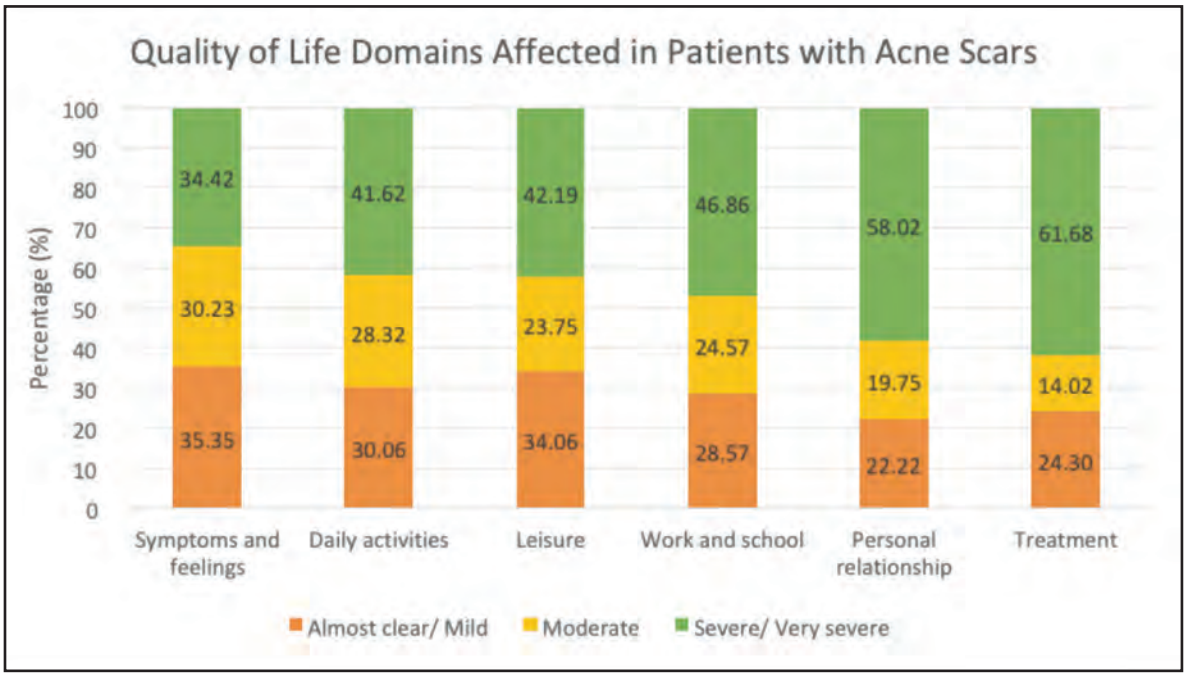


Fig. 1: Domains of Quality of Life Affected in Patients with Acne Scars

personal relationships, due to the reserved and introverted nature of our Asian culture, patients may find it a challenge to open up their feelings to friends, family and partner. A Singaporean study which evaluated the impact of post acne scars in young adults reported similar findings in this domain. Their lowest mean DLQI score was 0.26 for question nine on sexual difficulties.[2] This local Asian population data further supports our postulation. In the work / school domain, patients with significant acne scars may be perceived differently by society and this can negatively impact perceived attributes, skills and prospects. As a result, patients may find it challenging to appear in public and would try to avoid places where they could be seen, compounded by their own feelings of self-consciousness and embarrassment. Tan et al. evaluated the psychological well-being and social impact of atrophic acne scarring in a multi-national study which reported similar findings. In that study, a large group of patients reported significant impact with engaging and exposing daily activities such as swimming, yoga and sunbathing. Some patients also felt that they have been treated unfairly at work.[14]

Overall, quality of life is largely affected as acne scar severity worsens. However, a notable percentage of patients with mild scar also reported a high impact on quality of life. This finding was observed in the study by Tan et al. which found that even patients with mild acne scars reported high levels of impairment in quality of life.[3] Most of these patients were young females. Females often place a high value on their appearance. Acne scars can lead to decreased self-esteem and body image issues especially during their teenage years when emotional development is a pivotal growth. This in return, causes further emotional distress and affects social interactions be it at home, school or work.

The impact of acne scars on anxiety and depression is largely unknown. Most data revolved around the impact of acne vulgaris on psychosocial health such as anxiety, depression and quality of life. A multi-centre study reported significant psychological burden of common skin diseases with higher prevalence of depression and anxiety as well as suicidal ideation in the acne vulgaris group compared to many other common skin conditions such as psoriasis, atopic dermatitis, leg ulcers.[15] These psychosocial effects can be attributed to both active acne lesions as well as post acne scarring. In a previous study evaluating psychosocial judgements and perceptions of acne vulgaris among adolescents, the authors observed that teenagers, especially females with acne are often misperceived as shy, stressed, lonely, introverted, boring and nerdy. Teenagers with acne also reported lower self-confidence, difficulty finding jobs and engaging in personal relationships.[16]

Greater impairment in quality of life suggests an increased risk for anxiety and depression as demonstrated by our results. The impact of acne on quality of life, anxiety and depression is well documented.[17-19] Most patients from these studies had moderate to severe acne and demonstrated a common and significant association between DLQI and anxiety as well as depression. Those who developed anxiety ranged from 26-32% while 9-29% of the patients developed depression. A significant risk factor for the development of anxiety and depression is poor quality of life. They did not find any correlation between acne severity and the development of anxiety and depression, suggesting that the patient's own perception towards acne plays an important role in the negative effects of acne on psychosocial health. Negative societal perception towards those with acne and scars have also largely influenced the way patients perceive of themselves. Coupled with the distinctive impact of social

media and peer pressure these days which seem to glorify models and celebrities with crystal clear skin, the negative influence on psychosocial health and quality of life in patients with acne and scars is amplified. Findings from this study reminds us to not overlook the psychosocial health impacts of not just acne but post acne scars as well in our patients, with prompt referral to the psychiatrist should the need arise.

Observations from our study showed that patients with moderate and severe acne that were left untreated or treated late were at risk for higher acne scar severity. These findings were consistent with the findings of Dreno et al.[5] Family history of acne, male gender, acne relapses and alcohol have been identified as risk factors for formation of acne scars in some studies.[6-8] This is contrary to our findings which may be due to other factors that were not documented in our study such as accessibility to dermatological services and treatment or skincare habits which may also influence scar formation. These may have masked the effects of the risk factors which were analysed in our study like family history.

Preventing acne requires a combination of skincare, lifestyle changes and public health efforts. Early intervention and education with public initiatives, including school programs, affordable and accessible dermatological care, timely institution of treatment and evidence-based skincare awareness can reduce acne prevalence and improve skin health.

LIMITATIONS

Recall bias is a limitation of this study as patients graded their past acne severity based on memory. However, as scar severity was determined by clinical examination and psychosocial assessments were performed when patients no longer have active acne, our results would reflect the psychosocial impact of scars.

The objective of the study was to determine the impact of scars on quality of life and its association with anxiety and depression. Thus, we included patients with acne scars only, without active acne. This is to try to eliminate the impact of active acne on psychosocial issues.

The DLQI and HADS assessed in this study are based upon patients' response towards their current acne scars, rather than their previous acne. As the patients did not have active acne during the study, therefore acne severity had to be determined retrospectively.

CONCLUSION

This study highlights the impact of facial acne scars on psychosocial health. Impairment in quality of life and the development of anxiety and depression are inter-related. Greater acne scar severity is strongly associated with increased impairment in quality of life and a higher risk of developing anxiety and depression. Severe acne and delay in seeking treatment are independent risk factors for acne scar severity. Acne prevention and early intervention are key strategies against scars. Physicians should approach patients holistically, providing not only prompt and effective treatment in acne and its scars but also in terms of mental health.

ACKNOWLEDGEMENT

We would like to thank the Director General of Health Malaysia for his permission to publish this article.

FUNDING

No funding was received for conducting this study.

CONFLICT OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

1. Stathakis V, Kilkenny M, Marks R. Descriptive epidemiology of acne vulgaris in the community. Australas J Dermatol 1997; 38(3): 115-23.
2. Chuah SY, Goh CL. The Impact of Post-acne scars on the quality of life among young adults in Singapore. J Cutan Aesthet Surg 2015; 8(3): 153-8.
3. Tan J, Beissert S, Cook-Bolden F, Chavda R, Harper J, Hebert A, et al. Impact of facial atrophic acne scars on quality of life: a multi-country population-based survey. Am J Clin Dermatol 2022; 23(1): 115-23.
4. Hayashi N, Miyachi Y, Kawashima M. Prevalence of scars and "mini-scars", and their impact on quality of life in Japanese patients with acne. J Dermatol 2015; 42(7): 690-6.
5. Dreno B, Layton A, Bettolli V, Lozada VT, Kang S. Evaluation of the prevalence, risk factors, clinical characteristics and burden of acne scars among active acne patients in Brazil, France and the USA. J Am Acad Dermatol 2017; 76: AB132.
6. Tan J, Kang S, Leyden J. Prevalence and risk factors of acne scarring among patients consulting dermatologists in the USA. J Drugs Dermatol 2017; 16(2): 97-102.
7. Liu L, Xue Y, Chen Y, Chen T, Zhong J, Shao X, et al. Prevalence and risk factors of acne scars in patients with acne vulgaris. Skin Res Technol 2023; 29(6): e13386.
8. Say YH, Heng AHS, Reginald K, Wong YR, Teh KF, Rawanan Shah SM, et al. Modifiable and non-modifiable epidemiological risk factors for acne, acne severity and acne scarring among Malaysian Chinese: a cross-sectional study. BMC Public Health 2021; 21(1): 601.
9. Hayashi N, Akamatsu H, Kawashima M; Acne Study Group. Establishment of grading criteria for acne severity. J Dermatol 2008; 35(5): 255-60.
10. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol 1994; 19(3): 210-6.
11. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67(6): 361-70.
12. Tan JK, Tang J, Fung K, Gupta AK, Richard Thomas D, Sapra S, et al. Development and validation of a Scale for Acne Scar Severity (SCAR-S) of the face and trunk. J Cutan Med Surg 2010; 14(4): 156-60.
13. Basra MK, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. Br J Dermatol 2008; 159(5): 997-1035.
14. Tan J, Beissert S, Cook-Bolden F, Chavda R, Harper J, Hebert A, et al. Evaluation of psychological well-being and social impact of atrophic acne scarring: A multinational, mixed-methods study. JAAD Int 2021; 6: 43-50.

15. Dalgard FJ, Gieler U, Tomas-Aragones L, Lien L, Poot F, Jemec GBE, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol 2015; 135(4): 984-91.
16. Ritvo E, Del Rosso JQ, Stillman MA, La Riche C. Psychosocial judgements and perceptions of adolescents with acne vulgaris: A blinded, controlled comparison of adult and peer evaluations. Biopsychosoc Med 2011; 5(1): 11.
17. Martínez-García E, Arias-Santiago S, Herrera-Acosta E, Affleck A, Herrera-Ceballos E, Buendía-Eisman A. Quality of life of cohabitants of people living with acne. Acta Derm Venereol 2020; 100(17): adv00290.
18. Duman H, Topal IO, Kocaturk E, Duman MA. Evaluation of anxiety, depression, and quality of life in patients with acne vulgaris, and quality of life in their families. Dermatol Sin 2016; 34(1): 6-9.
19. Yazici K, Baz K, Yazici AE, Köktürk A, Tot S, Demirseren D, et al. Disease-specific quality of life is associated with anxiety and depression in patients with acne. J Eur Acad Dermatol Venereol 2004; 18(4): 435-9.

Mild cognitive impairment and its associated factors amongst the older people attending government health clinics in Kuantan

Yashdev Atri Roop[1], Mohd Shaiful Ehsan Shalihin[2], Edre Mohammad Aidid[3]

[1]IIUM Family Health Clinic, International Islamic University of Malaysia, Jalan Sultan Ahmad Shah, Kuantan, Pahang, [2]Department of Family Medicine, Kulliyyah of Medicine, International Islamic University of Malaysia, Jalan Sultan Ahmad Shah, Kuantan, Pahang, [3]Department of Community Medicine, Kulliyyah of Medicine, International Islamic University of Malaysia, Jalan Sultan Ahmad Shah, Kuantan, Pahang

ABSTRACT

Introduction: Malaysia is undergoing a demographic transition towards an aging population, resulting in an anticipated rise in the number of older individuals at risk of developing dementia. Mild cognitive impairment (MCI) represents an intermediate clinical stage between normal cognition and dementia, characterized by cognitive decline that does not significantly impair daily functional activities. Early detection of MCI is critical, as early-stage interventions and modifications of risk factors can yield promising outcomes.

Materials and Methods: This cross-sectional study evaluated the prevalence of MCI and its associated factors among 327 older adults attending healthcare clinics. Data were collected using the Elderly Cognitive Assessment Questionnaire, Geriatric Depression Scale, and Barthel Index. Binary logistic regression analysis was performed to identify significant factors of MCI.

Results: The prevalence of MCI in the study population was 18.7%, with the majority demonstrating poor control of comorbid conditions. Significant factors of MCI included being a widower (OR 0.4; 95%CI: 0.18, 0.94), increasing age (OR 0.92; 95% CI: 0.88, 0.97), and having diabetes (OR 3.49; 95% CI: 1.81, 6.73).

Conclusion: The findings highlight that a significant proportion of older adults are at risk of progressing to dementia but remain underdiagnosed during the early stages. Optimizing blood sugar control emerges as a crucial strategy to mitigate the progression of MCI to irreversible dementia. Implementation of active cognitive screening programs is essential for early identification and timely intervention.

KEYWORDS:
Mild, cognitive dysfunction, aged, clinical, risk

INTRODUCTION

Cognition encompasses thinking, knowledge acquisition, and memory. Cognitive decline is commonly observed in older individuals due to aging or underlying physical and mental conditions. It is typically marked by memory disturbances and is prevalent among the elderly.[1] Dementia is a primary cause of cognitive deterioration, with Alzheimer's disease being its most common form.[2] Mild cognitive impairment (MCI) represents an intermediate phase between normal aging-related cognitive changes and dementia. Early detection of dementia, whether primary or secondary, is crucial, as timely intervention can lead to better outcomes.[1]

The global prevalence of MCI ranges from 14.71% to 38.6%, as reported in studies from Saudi Arabia, China, and India.[3-5] In Malaysia, prevalence rates in rural and urban settings vary between 22.4% and 64.7%.[6,7] Advancing age is a key determinant of MCI, with studies from Saudi Arabia and Malaysia reporting increased odds of MCI with age.[3,6] Additionally, research from New York, Malaysia, and Taiwan suggests that females may have a higher prevalence and faster progression of MCI compared to males (OR 2.95; 95% Confidence Interval 95%CI:1.82, 4.78).[2,8,9] Lower education levels and fewer years of schooling are associated with an increased risk of MCI.[7,9] Strong family support may reduce this risk, while low social support has been linked to higher susceptibility.[10] Smoking has also been associated with cognitive decline, including MCI.[11] Regular physical activity is consistently linked to a lower risk of MCI.[4,9] Additionally, non-communicable conditions such as being underweight or overweight, hypertension, diabetes, and dyslipidemia contribute to cognitive impairment.[3,12,13] Depression may further exacerbate cognitive decline and significantly increase the risk of dementia.[14,15]

As Malaysia's population ages, the proportion of individuals aged 65 and above is expected to rise from 5.0% in 2010 to 14.5% by 2040.[16] In Malaysia, research has primarily focused on dementia and its risk factors. However, significant gaps remain, particularly in the comprehensive diagnosis of MCI in the community, including the incorporation of daily function scores and the identification of associated risk factors. Therefore, this study aims to assess the prevalence of MCI accordingly and its related determinants among elderly patients receiving care at government health facilities in Kuantan.

This article was accepted:
Corresponding Author: Mohd Shaiful Ehsan Bin Shalihin
Email: shaifulehsan@iium.edu.my

MATERIALS AND METHODS

Subjects

This cross-sectional study was conducted at government health clinics in Kuantan, Pahang, over six months beginning in January 2023. Participants were Malaysian citizens aged 60 years and older who attended these clinics. Individuals with known diagnoses of stroke, depression, dementia, or Alzheimer's disease were excluded from the study.

Sample Size

The sample size was determined using the single proportion formula, calculated with OpenEpi Version 3, an open-source calculator. Based on a prevalence rate of 14.6% reported by Hussin et al. in 2019, the minimum required sample size was 288.[17] To account for a 10% dropout rate, the final calculated minimum sample size was adjusted to 320.

Sampling Method

A multistage random sampling technique was employed. Initially, five clinics were randomly selected from a total of 12 clinics using the Excel randomization function. These clinics were chosen based on their high geriatric attendance rates and represented the target sample size of 320. Systematic random sampling was then applied, whereby every third elderly patient listed for the day's geriatric appointments was approached until the required number of participants per clinic (approximately 50 to 70) was achieved. The distribution of participants was proportional to the annual geriatric attendance rates for each clinic.

Study Instruments

Study's questionnaire comprised three sections:

1. Sociodemographic Data: This section captured information on participants' weight, height, age, ethnicity, education level, occupation, and marital status.
2. Medical History and Lifestyle Factors: This section assessed participants past medical history and lifestyle. It included the Activities of Daily Living questionnaire and the Geriatric Depression Scale.
3. Cognitive Assessment: Cognitive impairment was screened using the validated Malay version of the Elderly Cognitive Assessment Questionnaire (ECAQ). This interview-based tool, designed for high sensitivity and specificity, consists of three subsections: memory, orientation and information, and memory recall. The ECAQ has a scoring range of 0–10, with a cut-off score of 7 or lower indicating MCI. The Malay version's validation showed 85.3% sensitivity, 91.5% specificity, and an 82.8% positive predictive value, outperforming other tools such as the Montreal Cognitive Assessment (MoCA) and Mini-Mental State Examination (MMSE).

Elderly patients at the five selected clinics were approached by investigators who provided a verbal explanation of the study and distributed a detailed patient information sheet. Eligible individuals who met the inclusion criteria were asked to sign informed consent forms before participation.

Statistical Analysis

Data analysis was conducted using IBM SPSS version 27.0. A $p < 0.05$ was considered statistically significant. Descriptive statistics, including frequencies and percentages, were used to summarize the respondents' baseline characteristics. MCI was defined as a score ≤ 7 on the Malay version of the ECAQ, alongside an intact Barthel Index score and no clinically identified alternate causes as assessed by the physician. Bivariate analyses were performed using the Chi-Square test and independent t-test to identify potential factors associated with MCI. Variables with significant bivariate results were then included in a multiple logistic regression analysis to determine adjusted odds ratios and 95% confidence intervals.

Ethical Considerations

Ethical approval was granted by the Department of Family Medicine and the Kulliyyah of Medicine at the International Islamic University Malaysia (IIUM) (IREC 2022-180), as well as by the Medical Research and Ethics Committee (MREC) (NMRR ID-21-02107-4AY), the Pahang State Health Department, and the Kuantan District Health Office.

RESULTS

A total of 327 individuals participated in this study, with a mean age of 67 years (SD 6.1), ranging from 60 to 92 years. The majority of participants were of Malay ethnicity (76.8%) and married (84.4%), while a significant proportion had attained a secondary school level of education (73.4%). Only a small percentage of respondents lived alone (9.8%).

In terms of smoking status, 61.5% of participants were non-smokers at the time of the study, excluding 21.4% who had ceased smoking. The prevalence of obesity, based on body mass index (BMI), was notably high at 72.4%. Hypertension was present in 83.5% of respondents, with 33.0% experiencing uncontrolled systolic blood pressure and 12.5% having uncontrolled diastolic blood pressure. Diabetes was reported in 70.3% of the population, with poor glycemic control, as reflected by HbA1c levels, observed in 42.2% of these cases. Hyperlipidemia was the most prevalent comorbidity, affecting 93.0% of respondents.

All participants underwent assessments for depression and limitations in activities of daily living; no cases were identified. A summary of the sociodemographic and clinical characteristics of the study population is presented in Table I.

The prevalence of MCI observed in this study (18.7%). A significant association between age, marital status, education level, hypertensive status and diabetic status with MCI was observed ($p < 0.001$) as showed in Table II. Table III highlights the adjusted associations between sociodemographic and clinical factors and MCI among elderly individuals in Kuantan. The main factors for MCI are marital and diabetic status.

DISCUSSION

Prevalence of Mild Cognitive Impairment amongst the elderly

The MCI prevalence in this study (18.7%) aligns with previous Malaysian studies.[18,19] Rashid reported 11%, while Lee et al. and Sherina et al. found rates of 21.1% and 22.4%, respectively.6,18,19 Razali et al. reported a higher prevalence of 63.7%, likely due to differences in study settings.[7]

Table I Sociodemographic and clinical characteristics of the respondents (N=327)

Variables	n	%
Age(years) [a]	67.3	6.48
Gender		
Male	156	47.7
Female	171	52.3
Ethnicity		
Malay	251	76.8
Chinese	41	12.5
Indian	28	8.6
Others[b]	7	2.1
Marital Status		
Single	7	2.1
Married	276	84.4
Divorced / widowed	44	13.5
Education Level		
Primary school	61	18.7
Secondary school	240	73.4
Tertiary education	26	8.0
Living Alone		
Yes	32	9.8
No	295	90.2
Smoking History		
Smoker	56	17.1
Ex-smoker	70	21.4
Non-smoker	201	61.5
Weight Class (kg)		
30-49	18	5.5
50-69	140	42.8
70-89	139	42.5
90-109	25	7.6
110-129	5	1.5
BMI Classification		
Normal	47	14.4
Overweight	43	13.1
Obese type 1	128	39.1
Obese type 2	109	33.3
Hypertension		
Yes	273	83.5
No	54	16.5
Systolic Blood Pressure (mmHg)		
Normal	219	67.0
Uncontrolled	108	33.0
Diastolic Blood Pressure (mmHg)		
Normal	286	87.5
Uncontrolled	51	12.5
Diabetes		
Yes	230	70.3
No	97	29.7
HbA1C		
Not available	98	30.0
Normal	91	27.8
Uncontrolled	138	42.2
Hyperlipidaemia		
Yes	304	93.0
No	23	7.0
GDS[c] Class		
Normal	327	100.0
ADL[d] Class		
Normal	327	100.0

Note: Data were expressed as n (%) unless otherwise indicated
[a] Mean (SD)
[b] Others include Orang asli and other minority groups
[c] GDS – Geriatric depression scale
[d] ADL – Activity of daily living

Table II: Association between sociodemographic, clinical factors and MCI (N=327)

Variables	MCI		χ^2 Statistics (df = 1)	p-value
	Yes n (%)	No n (%)		
Age(years) [a]	71.85 (6.96)	66.26 (5.90)	4.930b	0.027[c]
Gender				
Male	29 (47.5)	127 (47.7)	0.001	0.977
Female	32 (52.5)	139 (52.3)		
Ethnicity				
Malay	48 (78.7)	203 (76.3)	2.821	0.420
Chinese	6 (9.8)	35 (13.2)		
Indian	7 (11.5)	21 (7.9)		
Others	0 (0.0)	7 (2.6)		
Marital Status				
Single	2 (3.3)	5 (1.9)	25.002	<0.001[c]
Married	39 (63.9)	237 (89.1)		
Widowed	20 (32.8)	24 (9.0)		
Living alone				
Yes	9 (14.8)	23 (8.6)	2.097	0.148
No	52 (85.2)	243 (91.4)		
Education level				
Primary school	23 (37.7)	38 (14.3)	18.760	<0.001[c]
Secondary school	36 (59.0)	204 (76.7)		
Tertiary education	2 (3.3)	24 (9.0)		
Smoking History				
Smoker	11 (18.0)	45 (16.9)	4.743	0.093
Ex-smoker	19 (31.1)	51 (19.2)		
Non-smoker	31 (50.8)	170 (63.9)		
Hypertension				
Yes	60 (98.4)	213 (80.1)	12.034	<0.001c
No	1 (1.6)	53 (19.9)		
Systolic Blood Pressure(mmHg)				
Normal	32 (52.5)	187 (70.3)	7.141	0.008[c]
Uncontrolled	29 (47.5)	79 (29.7)		
Diastolic Blood Pressure (mmHg)				
Normal	47 (77.0)	239 (89.8)	7.414	0.006[c]
Uncontrolled	14 (23.0)	27 (10.2)		
Hyperlipidaemia				
Yes	54 (88.5)	250 (94.0)	2.263	0.133
No	7 (11.5)	16 (6.0)		
Weight(kg)				
30-49	5 (8.2)	13 (4.9)	9.150	0.057
50-69	23 (37.7)	117 (44.0)		
70-89	28 (45.9)	111 (41.7)		
90-109	2 (3.3)	23 (8.6)		
110-129	3 (4.9)	2 (0.8)		
130-149	0 (0.0)	0 (0.0)		
Body Mass Index				
Underweight	0 (0.0)	0 (0.0)	2.362	0.501
Normal	6 (9.8)	41 (15.4)		
Overweight	6 (9.8)	37 (13.9)		
Obese1	26 (42.6)	102 (38.3)		
Obese 2	23 (37.7)	86 (32.3)		
Diabetes Mellitus				
Yes	28 (45.9)	202 (75.9)	21.459	<0.001[c]
No	33 (54.1)	64 (24.1)		

Note: Data were expressed based on Chi Square test otherwise indicated
[a] Mean (SD)
[b] F independent sample T test
[c] P < 0.05

Community-based studies, including ours, tend to report lower prevalence than hospital-based studies, where participants often have chronic illnesses requiring specialized care.[6,7,18]

Regionally, Xue et al. reported a pooled MCI prevalence of 14.7% in China, with slight variations based on methodology.[5] In Singapore, Lim et al. found a prevalence of 7.7% using the Elderly Cognitive Assessment Questionnaire (ECAQ), though its predominantly Chinese sample limits generalizability.[18] Studies using the Montreal Cognitive Assessment (MoCA) report MCI prevalence ranging from 6.7% to 25.2%, consistent with our findings.[5,18]

Table III Logistic regression analysis of factors associated with MCI

Variables	Odd ratio	Wald	Adjusted Odds Ratio	95%CI		p-value
				Lower	Upper	
Age(years)	-0.094	14.972	0.910	0.868	0.955	< 0.001[a]
Marital Status						
Married	Reference					
Single	-0.235	0.053	0.790	0.107	5.812	0.817
Widowed	-0.963	5.411	0.382	0.170	0.859	0.020[a]
Education level						
Primary school	-0.966	1.275	0.381	0.071	2.035	0.259
Secondary school	-0.385	0.255	0.681	0.139	3.331	0.635
Tertiary education	Reference					
Hypertension						
No	Reference		1			
Yes	1.856	6.398	0.158	0.805	50.829	0.079
Systolic Blood Pressure(mmHg)						
Normal	Reference		1			
Uncontrolled	0.597	2.612	1.897	0.881	3.749	0.106
Diastolic Blood Pressure (mmHg)						
Normal	Reference		1			
Uncontrolled	0.736	2.625	2.088	0.857	5.090	0.105
Diabetes Mellitus						
No	Reference		1			
Yes	1.250	13.918	3.490	1.810	6.729	<.001[a]
Constant	5.594	10.721	385.298			0.001[a]

Notes: Estimates of odds ratio from a binary logistic regression adjusted; aStatistically significant at P < 0.05; OR = odds ratio; CI = confidence interval

Associated Factors For Mild Cognitive Impairment (Univariate analysis)

A significant association between age and MCI was observed (p < 0.001), with the prevalence of MCI increasing in older age groups (Table II). For instance, the proportion of individuals with MCI rose from 16.4% in those aged 60–64 to 27.9% in the 70–74 age group and remained high in those aged 75–84. This aligns with existing evidence that advancing age is one of the strongest risk factors for cognitive decline, attributed to neurodegenerative changes, reduced brain plasticity, and vascular factors.19 Interestingly, the prevalence of MCI was slightly lower in the oldest age group (85+), possibly reflecting survivor bias, where the healthiest individuals tend to live longer, or under-detection due to reduced health-seeking behavior in this subgroup.

Marital status was significantly associated with MCI (p < 0.001), with divorced or widowed individuals showing a much higher prevalence of MCI (32.8%) compared to married individuals (9.0%). This highlights the potential protective role of marriage, which may provide emotional support, companionship, and assistance in managing health-related issues. In contrast, divorced or widowed individuals may experience loneliness and reduced social interaction, which are known risk factors for cognitive decline.[20]

Education level also showed a strong association with MCI (p < 0.001). Those with primary education exhibited the highest prevalence of MCI (37.7%), compared to 59.0% for secondary and 3.3% for tertiary education. This finding supports the cognitive reserve hypothesis, which posits that higher education provides a buffer against neurodegeneration by enhancing the brain's ability to cope with pathological changes.[21] Furthermore, lower education may be linked to poorer brain development, reduced synaptic density, and diminished neuroplasticity, contributing to the brain's decreased ability to compensate for age-related changes and neurodegeneration, thus increasing susceptibility to MCI.[22] The low prevalence of MCI among tertiary-educated individuals underscores the importance of lifelong learning and mental stimulation in promoting cognitive health.

Hypertension was significantly associated with MCI (p < 0.001), with 98.4% of individuals with MCI reporting a history of hypertension compared to 80.1% of those without MCI. Additionally, uncontrolled systolic (p = 0.008) and diastolic (p = 0.006) blood pressure were significantly linked to the presence of MCI. These findings align with evidence highlighting the role of hypertension in vascular damage and reduced cerebral blood flow, which contribute to the development of vascular cognitive impairment.[23] Previous studies have emphasized that effective management of blood pressure may mitigate the risk of cognitive impairment in aging populations.[23,24] These findings underscore the importance of routine screening for MCI in hypertensive individuals to facilitate early diagnosis and implement targeted interventions aimed at preventing cognitive decline.

Diabetes mellitus emerged as a significant factor of MCI (p < 0.001), with a much higher prevalence among individuals with MCI (45.9%) compared to those without MCI (24.1%). This association can be attributed to several mechanisms, including chronic hyperglycemia, insulin resistance, and vascular complications, which negatively affect brain health and increase the risk of cognitive impairment.[25] These findings emphasize the need for stringent glycemic control and regular cognitive screening among diabetic patients.

Significant Factors for Mild Cognitive Impairment
Marital Status
A significant association was observed between marital status and MCI, with divorced or widowed individuals showing higher odds of developing MCI compared to their married counterparts (AOR :0.418; 95% CI: 0.181, 0.962). These findings align with existing literature indicating that marital status serves as a proxy for social support, which may safeguard against cognitive decline. The absence of a spouse is often associated with increased loneliness, reduced social interaction, and diminished emotional support—factors linked to accelerated cognitive deterioration.[20]

Previous studies have reported that divorced or widowed individuals, particularly older adults, are more likely to experience MCI than married individuals.[26] Contributing factors may include the psychological stress of divorce, reduced social support, fewer social interactions, and limited mental stimulation, all of which are protective against MCI. Moreover, the financial and emotional hardships often associated with divorce or widowhood can exacerbate cognitive decline. These findings underscore the importance of community-based social interventions aimed at supporting divorced or widowed elderly individuals to mitigate their risk of cognitive impairment.

Diabetes Mellitus
Diabetes mellitus emerged as the strongest predictor of MCI in this study, with diabetic patients showing over three times the odds of having MCI compared to non-diabetic individuals (AOR:3.434; 95% CI: 1.775, 6.643). This finding is consistent with global evidence linking diabetes to cognitive impairment through mechanisms such as chronic hyperglycemia, insulin resistance, microvascular damage, and increased oxidative stress.[27,28] These processes contribute to both vascular dementia and Alzheimer's disease, explaining the high prevalence of cognitive decline among diabetic populations.

Chronic hyperglycemia and insulin resistance in diabetes lead to neuroinflammation, oxidative stress, and vascular damage, which impair synaptic plasticity and neuronal health, ultimately affecting memory and cognitive function. Additionally, diabetes accelerates amyloid-beta deposition and tau phosphorylation, both of which are hallmark pathological processes in Alzheimer's disease.[29] Given these mechanisms, integrating cognitive screening into diabetes management protocols and prioritizing interventions for optimal glycemic control is essential.

LIMITATIONS
This study was conducted exclusively in government primary healthcare clinics in Kuantan, thereby excluding patients who receive follow-up care in private practices or hospitals. While these clinics serve a demographically diverse population, the findings may not reflect the prevalence of MCI across the entire Kuantan district. Furthermore, the study's scope was confined to Kuantan, the principal district of Pahang State. Expanding future research to include other districts in Pahang and a wider range of medical facilities could provide more comprehensive insights into the prevalence and determinants of MCI.

CONCLUSION
Our findings revealed that 18.7% of the elderly population in Kuantan is affected by MCI. Key factors of MCI identified in this study were marital status and diabetes mellitus. Notably, Malaysia's national screening guidelines and algorithms do not currently incorporate routine screening for MCI in community healthcare settings. With the nation's aging population, the combination of modifiable and non-modifiable risk factors is expected to result in a marked increase in dementia cases. These findings underscore the need to revise targeted screening protocols for the elderly, with an emphasis on cardiovascular and dementia risk factors. Specifically, the results justify implementing routine MCI screening for elderly individuals, particularly those with diabetes, within community healthcare settings. To optimize healthcare delivery, we propose incorporating MCI screening into the existing diabetic foot screening schedule. Such an approach would streamline preventive efforts and facilitate early detection, enabling timely interventions to mitigate cognitive decline in at-risk populations.

ACKNOWLEDGEMENTS
We would like to acknowledge our Kulliyyah and Research management center for full support of this research. This study is funded by a research grant from the Ministry of Higher Education (MOHE) – Fundamental Research Grant Scheme (FRGS); FRGS/1/2021/SKK05/UIAM/03/1. The authors also thank the research management center IIUM for facilitating the grant management (FRGS21-211-0820).

REFERENCES
1. Chehrehnegar N, Nejati V, Shati M, Rashide V, Lofti M, et al. Early detection of cognitive disturbances in mild cognitive impairment: A systematic review of observational studies. Psychogeriatrics 2020; 20: 212-28.
2. Deng Y, Zhao S, Cheng G, Yang J, Li B, et al. The Prevalence of Mild Cognitive Impairment among Chinese People: A Meta-Analysis. Neuroepidemiology 2021; 55(2): 79-91.
3. Alkhunizan M, Alkhenizan A, Basudan L. Prevalence of mild cognitive impairment and dementia in Saudi Arabia: A community-based study. Dement Geriatr Cogn Disord Extra 2018; 8(1): 98-103.
4. Mohan D, Iype T, Varghese S, Usha A, Mohan M, et al. Cross-sectional study to assess prevalence and factors associated with mild cognitive impairment among older adults in an urban area of Kerala, South India. BMJ Open 2019; 9(3).
5. Xue J, Li J, Liang J, Chen S. The prevalence of mild cognitive impairment in China: A systematic review. Aging Dis 2018; 9(4): 706-15.
6. Sherina MS, Rampal L, Mustaqim A. Cognitive impairment among the elderly in a rural community in Malaysia. Med J Malaysia 2004; 59(2): 252-7.
7. Razali R, Baharudin A, Nik Jaafar NR, Sidi H, Rosli AH, Khoo BH, et al. Factors associated with mild cognitive impairment among elderly patients attending medical clinics in Universiti Kebangsaan Malaysia Medical Centre. Sains Malaysiana 2012; 41(5): 641-7.
8. Lin KA, Choudhury KR, Rathakrishnan BG, Marks DM, Petrella JR, et al. Marked gender differences in progression of mild cognitive impairment over 8 years. Alzheimers Dement (Amst) 2015; 1(2): 103-10.
9. Lim SC, Gan WY, Chan YM. The effects of socio-demographic characteristics, nutritional status, physical activity, and physical function on cognitive function of community-dwelling older adults in the Klang Valley, Malaysia. Malays J Med Health Sci 2020; 16(Suppl. 6).

10. Malek Rivan NF, Shahar S, Rajab NF, Ajit Singh DK, Che Din N, et al. Cognitive frailty among Malaysian older adults: Baseline findings from the LRGS TUA cohort study. Clin Interv Aging 2019; 14: 1343-52.
11. Muhammad T, Govindu M, Srivastava S. Relationship between chewing tobacco, smoking, consuming alcohol, and cognitive impairment among older adults in India: A cross-sectional study. BMC Geriatr 2021; 21(1).
12. Albai O, Frandes M, Timar R, Roman D, Timar B, et al. Risk factors for developing dementia in type 2 diabetes mellitus patients with mild cognitive impairment. Neuropsychiatr Dis Treat 2019; 15: 167-75.
13. Kjeldsen SE, Narkiewicz K, Burnier M, Oparil S. Systolic blood pressure control prevents cognitive decline and slows development of white matter lesions in the brain: The SPRINT MIND study outcomes. Blood Press 2019; 28(6): 356-7.
14. Chow YY, Verdonschot M, McEvoy CT, Peeters G. Associations between depression and cognition, mild cognitive impairment, and dementia in persons with diabetes mellitus: A systematic review and meta-analysis. Diabetes Res Clin Pract 2022.
15. Ruthirakuhan M, Herrmann N, Vieira D, Gallagher D, Lanctôt KL, et al. The roles of apathy and depression in predicting Alzheimer disease: A longitudinal analysis in older adults with mild cognitive impairment. Am J Geriatr Psychiatry 2019; 27(8): 873-82.
16. Department of Statistics Malaysia. Malaysian Population. 2022. Accessed February 15, 2024. Available from: https://www.dosm.gov.my/portal-main/landingv2.
17. Hussin NM, Shahar S, Yahya HM, Che Din N, Ajit Singh DK, et al. Incidence and predictors of mild cognitive impairment (MCI) within a multi-ethnic Asian populace: A community-based longitudinal study. BMC Public Health 2019; 19(1).
18. Rashid AK. Impairment among the elderly Malays living in rural Malaysia. Med J Malaysia 2012; 67(2).
19. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA 2014; 312(23): 2551-61.
20. Shankar A, Hamer M, McMunn A, Steptoe A. Social isolation and loneliness: relationships with cognitive function during 4 years of follow-up in the English Longitudinal Study of Ageing. Psychosom Med 2013; 75(2): 161-70.
21. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol 2012; 11(11): 1006-12.
22. Chen TB, Yang H, Cheng Y, Cooper C, et al. The impact of education on mild cognitive impairment and its progression to dementia: A systematic review and meta-analysis. J Alzheimers 2022; 85(2): 763-75.
23. Iadecola C, Gottesman RF, Seshadri S. Hypertension and brain dysfunction in the pathogenesis of cognitive impairment and Alzheimer's disease: A review. Hypertension 2021; 78(4): 955-64.
24. Wei J, Yin X, Liu Qi, Tan L, Jia C, et al. Association between hypertension and cognitive function: A cross-sectional study in people over 45 years old in China. J Clin Hypertens 2018; 20(11): 1575-83.
25. Cheng G, Huang C, Deng H, Wang H. Diabetes as a risk factor for dementia and mild cognitive impairment: a meta-analysis of longitudinal studies. Intern Med J 2012; 42(5): 484-91.
26. Brown SL, Lin IF, Vielee A, Mellencamp KA. Midlife Marital Dissolution and the Onset of Cognitive Impairment. Gerontologist 2021; 61(7): 1085-1094.
27. Jin CY, Yu SW, Yin JT, Yuan XY, Wang XG. Corresponding risk factors between cognitive impairment and type 1 diabetes mellitus: A narrative review. Heliyon 2022; 8(8): e10073.
28. Abdul Munif M, Verma L, Faizan Ahmad M, Khan AA, Singh A. Association between risk factors and cognitive impairment among type 2 diabetes mellitus patients. Int J Health Serv Res Policy 2022; 7(2): 173-80.
29. Rojas M, Chávez-Castillo M, Bautista J, Ortega Á, Nava M, et al. Alzheimer's disease and type 2 diabetes mellitus: Pathophysiologic and pharmacotherapeutics links. World J Diabetes. 2021; 12(6): 745-766.

ORIGINAL ARTICLE

Effect of oral colostrum application every 2 hours and 4 hours in order to achieve trophic feeding in preterm infants: A randomized controlled trial

Astri Tantri Indriani, MD[1,2,3], Setya Wandita, MD[2,3], Ekawaty Lutfia Haksari, MD[2,3], Tunjung Wibowo, MD[2,3], Alifah Anggraini, MD[2,3], Helen Natalia, MD[3]

[1]Faculty of Medicine, Universitas Sebelas Maret, [2]RSUP Dr. Sardjito, Yogyakarta, Indonesia, [3]Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada

ABSTRACT

Introduction: Enteral feeding in preterm neonates starts with trophic feeding, which is the practice of feeding minute volumes of enteral feeds (starting at 10-25mL/kg/day) through an orogastric tube. Colostrum has protective effects, such as anti-inflammatory, immunomodulatory, and antimicrobial effects. The oral colostrum application is a safe, effective and economical therapy. However, the most optimal frequency of the oral colostrum application is not yet conclusive. This study aims to evaluate the effects of applying colostrum orally every 4 and 2 hours in order to achieve trophic feeding in preterm infants.

Materials and Methods: In this randomized controlled trial with an open-label design, very-low-birth-weight neonates admitted to RSUP Dr. Sardjito from March to August 2023 were allocated to receive oral colostrum applications either every two hours or every four hours. Subjects were randomized into study groups using a random block size of four through computer-generated in a 1:1 ratio. The primary outcome was the time to achieve trophic feeding. The extraneous variables were necrotizing enterocolitis, sepsis, hemodynamically significant Patent Ductus Arteriosus (hsPDA) and gender. Data analysis was conducted using SPSS.

Results: A total of 40 neonates were analyzed for primary outcome. Of these, 20 neonates received oral colostrum applications every 2 hours, and the other 20 subjects were fed every 4 hours. Bivariate analysis showed that colostrum application given every 4 hours achieved the trophic feeding 0.47 day faster than the colostrum application every 2 hours. However, the difference between the two feeding methods was not statistically significant (p=0.703).

Conclusion: There is no significant difference in achieving trophic feeding in preterm neonates (less than 34 weeks) whether the colostrum was given every 2 or 4 hours.

KEYWORDS:
Colostrum, very low birth weight, trophic feeding, clinical trial, preterm

INTRODUCTION

A healthy oral cavity serves as the first line of defense against infections. When enteral feeding via oral cannot be administered to the sick or preterm infants, it may compromise the integrity of the buccal mucosa and serve as a focal point for infection. The oropharyngeal colostrum application helps maintain oral cavity moisture and reduces bacterial colonization.[1]

Numerous studies have evaluated the effects of the oral care using sterile water, colostrum, or breast milk. The usage of colostrum or breast milk is preferred due to its immunological benefits. The oropharyngeal colostrum application with breast milk involves the direct administration or spreading of breast milk onto the oral mucosa to keep it clean and intact. This procedure is safe and well-tolerated, even in clinically unstable infants.[1,2]

Applying colostrum to the oral cavity stimulates the immune response of the oropharyngeal lymphoid tissue and prevents microbial adhesion by providing a layer of immunoglobulin A (IgA) on the oropharyngeal mucosa. Oropharyngeal colostrum immunotherapy is a good strategy to deliver antimicrobial and anti-inflammatory protective factors, stimulate immune cells in oropharyngeal lymphoid tissue, support immune system maturation, and aid digestive system maturation.[3,4] Additionally, it is beneficial for the development of taste and smell sensations, initiates early enteral feeding, accelerating full enteral feeding, and speeds up the return to birth weight.[1,2]

The application of colostrum to the oropharynx is a simple, economical, and positive treatment, but the frequency of colostrum application is still inconclusive. Further studies are needed for producing reliable scientific evidence. Some studies indicate that more frequent colostrum application in the oropharynx leads to better outcomes.[5] A meta-analysis in June 2022 by Huo et al., including 11 randomized controlled trials, showed that oropharyngeal colostrum administration reduces the risk of necrotizing enterocolitis, late-onset sepsis, ventilator-associated pneumonia, accelerates full enteral feeding, and reduces the length of hospital stay.[6] A more recent meta-analysis in July 2022 by Cai et al. stated similar findings.[7] In both, conducted meta-analyses, variations in the

This article was accepted: 30 March 2025
Corresponding Author: Astri Tantri Indriani
Email: tantriindrianiastri@gmail.com;astri_tan3@stafff.uns.ac.id

interval of oropharyngeal colostrum administration were observed among studies.[6,7]

Most oropharyngeal colostrum application is done at intervals of every 2 hours. However, there is another study using a 4-hour interval, proving that colostrum administration every 4 hours increases lactoferrin in saliva and has a positive impact on expediting enteral feeding.[4] Hence, oral colostrum application has positive effects, but currently there is no research on the optimal frequency of colostrum application that can be applied in Indonesia.

This study aims to evaluate the effects of oral colostrum application frequency, every 4 hours and every 2 hours, in order to achieve trophic feeding in preterm infants <34 weeks gestational age.

MATERIALS AND METHODS

This experimental study using an open-label Randomized Controlled Trial (RCT) design was conducted from March to August 2023. The study was done in a level III neonatal referral center in Indonesia at Neonatal Intensive Care Unit (NICU) of RSUP Dr. Sardjito. Inclusion criteria were infants with gestational age <34 weeks, birth weight between 1000-1500 grams, colostrum used was from the mother's own breast milk (not donors' milk), and parents/guardians willing to participate in the study by signing informed consent. This study excluded infants with major congenital abnormalities (Table I), infants born to Human Immunodeficiency Virus-positive mothers, infants whose mothers consumed breastfeeding contraindicated drugs (Table I), severely ill mothers unable to provide breast milk, and mothers unable to supply breast milk within 48 hours after delivery. The research was conducted after obtaining approval from the Ethics Committee of the Faculty of Medicine, Public Health, and Nursing, UGM, with approval number KE-FK-0811-EC-202. The trial was registered retrospectively at the U.S. National Institutes of Health (ClinicalTrials.gov) with number NCT06379178.

In this study, data was collected based on consecutive sampling until the required sample size was reached. The sample size was determined using the formula for the independent t-test with alpha= 0.05 and beta=0.2, we calculated a minimum of 14 samples in each group, totaling 28.[8] To account for potential loss to follow-up, the researcher added 10% more samples, resulting in a minimum total of 32 samples. Subjects were randomized into study groups using a random block size of four through computer-generated in a 1:1 ratio. This study used four randomization blocks, with code A for colostrum application every 2 hours and code B for colostrum application every 4 hours. There were six permutation sequences: AABB, ABBA, BAAB, BABA, ABAB, BBAA. Samples were subsequently allocated to either the 2-hour or 4-hour colostrum application group based on the code in the randomized sequence. This study was also an open-label controlled trial with no blinding.

The study necessitated colostrum from the mother's own breast milk, a 1 ml syringe to measure colostrum volume, a tool to facilitate even colostrum application, and a suction catheter to remove excess saliva. Breast milk expression was initiated within 6 hours post-delivery, adjusted based on the mother's condition. Nurses or doctors assisted with the initial breast milk expression using standard hospital-grade breast pumps or manual methods. Expression was performed by the mother or with the assistance of her spouse every 2-3 hours (eight times per day), with a minimum duration of 10 minutes on each breast.

The expressed breast milk was collected in a container, drawn into a 1 ml, 3 ml, or 5 ml syringe depending on the obtained volume, and then labeled for identification. For colostrum application, the nurse took 0.2 ml of colostrum using a 1 ml syringe. Colostrum application prioritized the freshest breast milk, allowing direct administration to the research subject. The remaining breast milk can be used within 2-4 hours at room temperature (16-29 °C) or stored in the refrigerator or freezer according to breast milk storage standards. If frozen, the breast milk should be thawed in the refrigerator before being extracted in the amount of 0.2 ml using a 1 ml syringe for application to the research subject. If stored in the refrigerator, rather than the freezer, 0.2 ml of breast milk should be drawn with a 1 ml syringe, allowed to sit for a designated period (15-30 minutes), and subsequently administered to the research subject.[9]

The nurses followed a pre-procedure checklist for colostrum application in the NICU, prepared and administered the colostrum for each research subject scheduled for colostrum application. If the research subject used mechanical ventilation, the colostrum application procedure was carried out by two nurses. Before colostrum application, the oral cavity of the research subject was examined. If excess fluid or saliva was found, suction was performed first, or a sterile gauze was used to remove excess fluid while eliminating dry skin or debris on the lips. Then, with the aid of a 1 ml syringe after removing the needle, colostrum was evenly applied to both right and left cheeks (buccal mucosa), the roof of the mouth, gingival surface, and lips for approximately 1-2 minutes every 2 hours (12 times per day) or every 4 hours (6 times per day) depending on the code in the randomization sequence. Colostrum in the mouth will be distributed as it mixes with saliva. Colostrum application to the oral cavity began within 24-48 hours after birth and continued for a total of 5 days.

In case of an unstable clinical condition such as apnea, tachypnea, or desaturation, stabilization was carried out first following the Neonatal Intensive Care Unit (NICU) RSUP Dr. Sardjito Standard Operational Procedure (SOP). After stabilization, colostrum application can be resumed. In cases of feeding intolerance, management was carried out according to the SOP for handling feeding intolerance. If breast milk is available and there are no contraindications to enteral feeding, it can be administered promptly.

Observations of subjects were recorded using the NICU daily monitoring sheet, with detailed notes on various evaluations such as age, days of care, vital signs, feeding intolerance, diagnosis, fluid balance, diuresis, and so forth. Primary research data was from the Case Report Form (CRF) based on the Electronic Medical Record (EMR) and NICU daily

Table I: Exclusion criterias

Major congenital abnormalities			
Atresia esophagus	Atresia anorectal	Association Atresia intestinal	Anencephaly
Conjoined twins	Cyanotic congenital heart disease	Diaphragmatic hernia	Gastroschisis
Hypoplasia Lung	Klinefelter	Lung agenesis	Omphalocele
Orofacial cleft	Pyloric stenosis	Syndrome Hirschrpung disease	Trisomy 13
Trisomy 18	Turner syndrome		
Contraindicated drugs			
Abacavir	Atenolol	Simvastatin	Valganciclovir
Acitretin	Azathioprine	Dihydroergotamine	Macitentan
Alprazolam	Ado-trastuzumab emantasine	Diphenoxylate	Methamphetamine
Amitriptyline	Amantadine	Dolutegravir	Miltefosine
Amoxapine	Amphetamine	Efavirenz	Nelfinavir
Aripiprazel	Anakinra	Emtricitabine	Nevirapine
Aspirin	Atazanavir	Enfuvirtide	Pentamidine
THC (marijuana)	Atorvastatin	Entecavir	Rifabutin
Tenofovir	Chlorpromazine	Ergotamine	Ritonavir
Tramadol	Clonazepam	Fosamprenavir	Saquinavir
Lopinavir	Cobicistat	Heroin	Smallpox vaccine
Lamivudine	Cocaine	Indinavir	
Zidovudine	Delavirdine	Stavudine	

Table II: Baseline Characteristics of Research Subjects

Characteristic	Colostrum application every 2 hours	Colostrum application every 4 hours
Gender, n (%)		
Female	6 (15)	11 (27,5)
Male	14 (35)	9 (22,5)
Birth weight in g, mean ±SD	1.314,80 ± 117,921	1.215,30 ± 140,067
Gestational age in weeks, n (%)		
28 weeks - <31+6 weeks	16 (40)	10 (25)
32 weeks - <33+6 weeks	4 (10)	10 (25)
Methods of delivery, n (%)		
Spontaneous delivery	11 (27,5)	6 (15)
Abdominal delivery	9 (22,5)	14 (35)
Maternal infection risk factors, n (%)		
With risk factors	14 (35)	12 (30)
Without risk factors	6 (15)	8 (20)
Maternal age in years, mean ±SD	31,45± 6,022	32,15± 7,013
Sepsis, n (%)		
Yes	20 (50)	19 (47,5)
No	0 (0)	1 (2,5)
hemodynamically significant Patent Ductus Arteriosus, n (%)		
Yes	5 (12,5)	6 (15)
No	15 (37,5)	14 (35)
Necrotizing enterocolitis, n (%)		
Yes	3 (7,5)	1 (2,5)
No	17 (42,5)	19 (47,5)

g=gram; SD= standard deviation

Table III: Bivariate analysis (Independent t-test) of colostrum application and days to achieve trophic feeding

Variable	Mean (days) to achieve trophic feeding		95% CI	p-value
	2 hours	4 hours		
Colostrum application	7	6,53	-2,011 -2,944	0,703

monitoring sheet. The research subjects were prospectively observed, with data being regularly monitored every day until the primary outcomes were achieved, with a maximum observation period of 21 days.

The primary outcome is the time to achieve trophic feeding defined as reaching a target volume of 25mL/kg/day with enteral feeding tolerance. Tolerance of enteral feeding was assessed if gastric residual was ≤5mL/kg or residual was <50% of the volume given, no blood was found in the residual, no vomiting or abdominal distension occurred, and there was no decrease, delay, or cessation of enteral feeding. While external variables include necrotizing enterocolitis, sepsis, hsPDA, and gender, represented as categorical variables.

Table IV: Bivariate analysis (Independent t-test) of external variable and dependent variable (days to achieve trophic feeding)

Variable	Mean (days) to achieve trophic feeding		95% CI	p-value
	Yes	No		
Subjects with colostrum application every 2 hours				
Necrotizing enterocolitis	10,50	6,53	-2,292 – 10,225	0,197
Sepsis	7,00	-	-	-
hsPDA	10,33	6,29	-1,093 – 9,188	0,114
Male gender	6,42	8,40	-2,575 – 6,542	0, 368
Subjects with colostrum application every 4 hours				
Necrotizing enterocolitis	9,00	6,36	-3,215 – 8,501	0,348
Sepsis	6,71	4,00	-3,132 – 8,561	0,334
hsPDA	6,00	6,67	-4,430 – 3,096	0,708
Male gender	6,71	6,38	-3,367 – 2,688	0,812

The collected data was entered into a database and processed using Statistical Package for the Social Sciences (SPSS) software version 27. Data characteristics were analyzed using univariate tests to determine frequency distribution. Hypothesis testing involves bivariate analysis to determine the mean time to achieve trophic feeding (Independent t-test). A significance level of $p<0.05$ is considered statistically significant. The study followed the principle of intention to treat, meaning that the entire sample will be analyzed.

RESULTS

There were 47 preterm infants <34 weeks gestational age being treated in the Neonatal Intensive Care Unit (NICU) at RSUP Dr. Sardjito over a six month period between March and August 2023, including those born at RSUP Sardjito and in external healthcare facilities referred to our unit.

Out of the 47 very low birth weight (VLBW) infants, 7 were excluded, including 1 infant with esophageal atresia, 1 infant born to a critically ill mother who passed away, resulting in no available breast milk, 1 with mother consuming medication contraindicated for breastfeeding, and 4 infants receiving donor breast milk. The 40 eligible VLBW infants were randomly divided into two treatment groups: 20 received colostrum application every 2 hours, while the rest received it every 4 hours based on the randomization sequence from the permutation code. During the follow-up process, 8 subjects died before reaching trophic feeding, with a breakdown of 7.5% mortality in the colostrum application every 2 hours group and 12.5% in the every 4 hours group. The study applied the intention-to-treat principle, so the analyzed sample size included all 40 research subjects. Figure 1 presents the flowchart of the study. Table II shows the baseline characteristics of the neonates.

Table III demonstrates the bivariate analysis result using an Independent t-test of the oral colostrum application and the time to achieve trophic feeding. The analysis results indicate that subjects with colostrum application every 4 hours reached trophic feeding, on average, 0.47 days earlier than subjects with oral colostrum application every 2 hours. However, the difference in means is not statistically significant ($p=0.703$).

Table IV shows bivariate analysis results (Independent t-test) of the external variable with time to achieve trophic feeding. It indicates that subjects with necrotizing enterocolitis, sepsis, male gender, and hsPDA did not show statistically significant differences on the mean to achieve trophic feeding, as evidenced by the $p\geq0.05$ values in both groups of subjects receiving colostrum every 2 hours and every 4 hours (Table IV).

DISCUSSION

Our study was conducted on a population of preterm infants <34 weeks gestation and birth weight less than 1500 grams (VLBW) with the frequency of oral colostrum application every 2 hours and every 4 hours. In previous research, three studies provided colostrum every 2 hours, with two of them were conducted on VLBW populations and one study on extremely low birth weight infants (ELBW). There were six studies administering colostrum every 3 hours, with two of them were conducted on VLBW populations, and four other studies did not mention the weight of the subjects. Two studies administered colostrum every 4 hours, with one study was conducted on ELBW populations and one study did not specify the weight of the subjects.[10] The baseline characteristic for birth weight between the two groups in our study was normally distributed.

The gestational age in our study was 28-31+6 weeks in 65.6% of the cases and 32 weeks - <33+6 weeks in 34.4%, whereas in previous studies, both the researches with colostrum application every 2 hours and 4 hours, had populations with an average gestational age of <34 weeks, with only one study by Rodriguez using a population <28 weeks.[11] Thus, all participants included were within the range of preterm infants studied in previous research.

The colostrum volume used in previous studies was 0.2 ml, with only one study by Romero using a colostrum volume of 0.3 ml.[10,12] In this study, the volume given was 0.2 ml for colostrum administration every 2 hours and 4 hours. The duration of colostrum administration in previous studies varied between 1-3 days, 4-7 days, 8-10 days, and more than 10 days.[10] In this study, the duration of colostrum administration was 5 days for both every 2 hours and 4 hours colostrum applications. To date, the researchers have not found a similar study.

The objective of our current study is to investigate the effects of colostrum application every 2 hours and every 4 hours on the attainment of trophic feeding. We have not found a similar previous study, but there are several studies that have been conducted related to feeding status, including full enteral feeding and feeding intolerance.[4,12,13]

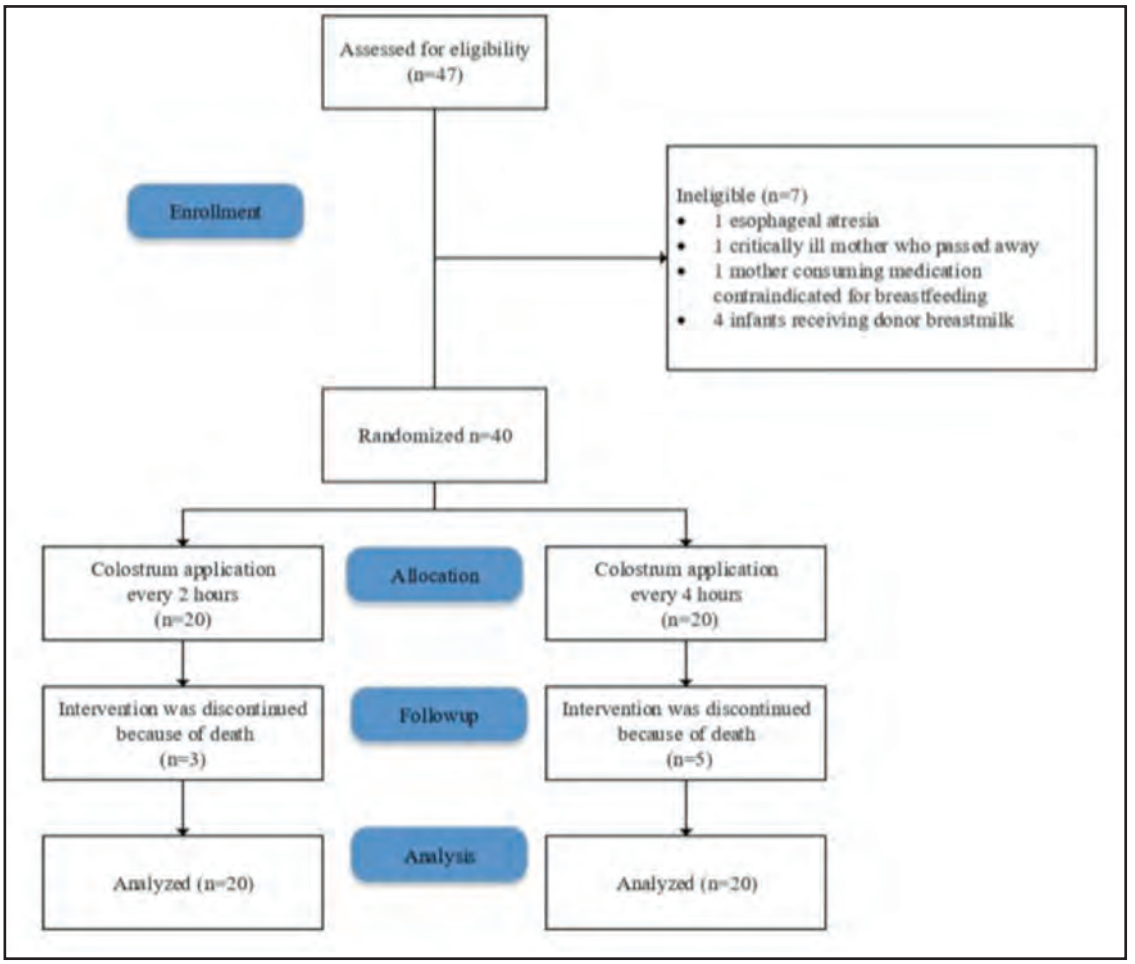


Fig. 1: Research flow until data analysis

The result of a meta-analysis conducted by Fu et al indicates that there is no significant difference in feeding status indicators between colostrum administration every 2 hours and 3 hours.[10] However, when colostrum is administered every 4 hours, a significant difference is observed, with a faster attainment of trophic feeding. These meta-analysis results differ from our study's findings, where the application of colostrum every 2 hours and every 4 hours shows no statistically significant difference (Table III). This discrepancy may be attributed to the differences in the study population, the volume of colostrum used, and the duration of colostrum administration.

The application of colostrum every 2 hours and every 4 hours in the duration of 5 days in patients with necrotizing enterocolitis and sepsis showed no difference in achieving trophic feeding (Table IV). A meta-analysis by Fu et al. suggests that the colostrum application every 4 hours with a shorter duration will affect feeding status, while a duration of 8-10 days can prevent necrotizing enterocolitis and late-onset sepsis.[10] It is possible that our study may show differences if the duration of colostrum administration is longer. Further research is needed to examine the feeding status, necrotizing enterocolitis, or sepsis with colostrum application every 2 hours and 4 hours with a longer duration.

The research results indicate that colostrum application every 2 hours and 4 hours showed no significant difference in achieving trophic feeding. Since our findings show no significant difference in clinical outcome, we suggest that colostrum be administered every 4 hours based on practical considerations, including the limited colostrum availability in the early postpartum. A study conducted in Indonesia examined the average amount of breast milk produced in spontaneous deliveries, with an average of (± 0.155) ml at 2 hours postpartum, (± 1.272) ml at 16 hours postpartum, and (± 1.369) ml at 24 hours postpartum.[14] Another reason for recommending colostrum application every 4 hours is that premature or very low birth weight infants are often born to ill mothers. While NICU workload is not a clinical consideration, the implementation of colostrum application every 4 hours can offer operational benefits, reducing the nursing workload, as providing nutrition is a time-consuming task.[15]

A study indicates that the recommended clinical management for infants is to adopt a minimal handling approach. Minimal handling is a common practice in the NICU to provide restful sleep for infants without disturbance from medical, care, and other examination activities. Minimal handling can prevent the occurrence of intraventricular hemorrhage (IVH) in the acute phase, within

72 hours after birth.[16] The clinical implications of this research could serve as a basis or guideline for clinical practice in colostrum application for very low birth weight infants (VLBW) and suggest that colostrum application should be performed every 4 hours. This aligns with the recommendations from the meta-analysis conducted by Fu et al.[10]

Challenges encountered during our research included difficulties in educating and convincing mothers and families to perform breast milk expression. Often, mothers lacked confidence, so we addressed this issue by providing education to mothers and families immediately after the baby's birth. We also provided intensive family support several times until the 5th day, regularly reminding mothers and families about breast milk expression and collection. On average, after the 4th day, mothers and families became more enthusiastic and disciplined in collecting breast milk.

To the best of our knowledge, this study is one of the first to explore the frequency of oral colostrum application every 2 hours and 4 hours in infants born before 34 weeks of gestation. As of now, we have not found any publications or ongoing studies on the frequency of colostrum applications. However, we acknowledge that our study has some limitations. It was conducted at a single center and had an abnormal distribution of data. The research method used was the randomized control trial open label, and blinding was not possible due to the limitations in the research location and team. While the absence of blinding increases the potential for a placebo effect, we minimized the bias by using objective outcome measures and implementing a standardized procedure checklist. Karyotyping was not performed to determine major congenital anomalies in subjects meeting the exclusion criteria due to the limitations in diagnostic facilities. Subjects meeting the exclusion criteria with a list of major congenital anomalies were determined based on clinical manifestations of related syndromes.

CONCLUSION

To conclude, the frequency of oral colostrum application every 4 hours does not differ from every 2 hours in order to achieve trophic feeding in infants born before 34 weeks of gestation. Colostrum application in the oral cavity should preferably be done every 4 hours because the amount of breast milk produced in the early postpartum period is not abundant. Premature or very low birth weight infants are often born to ill mothers. The implementation of colostrum application every 4 hours is a minimal handling approach compared to every 2 hours and can reduce the workload of NICU nurses. Further research is needed to evaluate the impact of longer oral colostrum administration durations on feeding status, necrotizing enterocolitis, or late-onset sepsis in colostrum applications every 2 hours and 4 hours.

ACKNOWLEDGEMENTS

The authors express sincere appreciation to Neonatal Intensive Care Unit RSUP Dr. Sardjito for the invaluable support and families who allowed their infants to participate in the study.

SCOPE STATEMENT

This original research presents a critical aspect of neonatal care — evaluating the effects of oral colostrum application frequency, administered every 4 hours versus every 2 hours, on achieving trophic feeding in preterm infants <34 weeks gestational age by using an open-label randomized clinical trial. The study explored numerous factors, including gestational age, birthweight, gender, maternal age and risk factors, methods of delivery, necrotizing enterocolitis, sepsis, hemodynamically significant patent ductus arteriosus among the subjects. Providing evidence into the most effective frequency of colostrum administration for these vulnerable infants, the study highlights the potential benefits for implications in clinical practice in optimizing feeding status of preterm infants. The study findings hold a promising suggestion in protocols and practices within neonatal intensive care settings.

REFERENCES

1. Digal KC, Upadhyay J, Singh P, Shubham S, Grover R, Basu S. Oral Care with Mother's Own Milk in Sick and Preterm Neonates: A Quality Improvement Initiative. Indian J Pediatr 2021; 88(1): 50-7.
2. Nasuf AWA, Ojha S, Dorling J. Oropharyngeal colostrum in preventing mortality and morbidity in preterm infants. Cochrane Database of Syst Rev 2018; 9(9): CD011921.
3. Garofalo NA and Caplan MS. Oropharyngeal Mother's Milk: State of the Science and Influence on Necrotizing Enterocolitis. Clin. Perinatol 2019; 46(1): 77-88.
4. Zhang Y, Ji F, Hu X, Cao Y, Latour JM. Oropharyngeal Colostrum Administration in Very Low Birth Weight Infants: A Randomized Controlled Trial. Pediatr Crit Care Med 2017; 18(9): 869-75.
5. Xavier Ramos MS, Martins CDC, Souza ES, Vieira GO, Gomes-Filho IS, Figueiredo ACMG, et al. Oropharyngeal colostrum immunotherapy and nutrition in preterm newborns: meta-analysis. Rev Saude Publica 2021; 55: 59.
6. Huo M, Liu C, Mei H, Zhang Y, Liu C, Song D, et al. Intervention Effect of Oropharyngeal Administration of Colostrum in Preterm Infants: A Meta-Analysis. Front Pediatr 2022; 10: 895375.
7. Cai M, Lin L, Peng Y, Chen L, Lin Y. Effect of Breast Milk Oral Care on Mechanically Ventilated Preterm Infants: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Front Pediatr 2022; 10: 899193.
8. Madiyono B, Sastroasmoro S. 2014. Dasar-dasar Metodologi Penelitian klinik. 5th Edition. Sagung Seto.
9. Eglash A, Simon L, Academy of Breastfeeding Medicine. ABM clinical protocol #8: Human milk storage information for home use for full-term Infants, Revised 2017. Breastfeed Med 2017; 12(7): 390-5.
10. Fu ZY, Huang C, Lei L, Chen LC, Wei LJ, Tao M, et al. The effect of oropharyngeal colostrum administration on the clinical outcomes of premature infants: A meta-analysis. Int. J. Adv. Nurs. Stud 2023; 144: 104527.
11. Rodriguez NA, Groer MW, Zeller JM, Engstrom JL, Fogg L, Du H, et al. A Randomized Controlled Trial of the Oropharyngeal Administration of Mother's Colostrum to Extremely Low Birth Weight Infants in the First Days of Life. Neonatal Intensive Care: j. perinatol.-neonatol 2011; 24(4): 31-5.
12. Romero-Maldonado S, Soriano-Becerril DM, García-May PK, Reyes-Muñoz E, Muñoz-Ortíz EG, Carrera-Muiños S, et al. Effect of Oropharyngeal Administration of Colostrum in Premature Newborns ≤32 Weeks of Gestation on the Immune Response and Neonatal Morbidity: A Double-Blind Randomized Clinical Trial. Front Pediatr 2022; 10: 891491.

13. Sharma D, Kaur A, Farahbakhsh N, Agarwal S. Role of oropharyngeal administration of colostrum in very low birth weight infants for reducing necrotizing enterocolitis: a randomized controlled trial. Am J Perinatol 2020; 37(7): 716-21.
14. Kustriyani M and Wulandari P. The First 24 Hours Post Partum Mother's Breast Milk Production at Hospital. South East Asia Nursing Research 2020; 2(4): 151-5.
15. Yadav A, Siddiqui N, Debata KP. Two-hourly vs Three-hourly Feeding in Very Low Birthweight Neonates: A Randomized Controlled Trial. Indian Pediatr 2021; 58(4): 320-4.
16. Isayama T. The clinical management and outcomes of extremely preterm infants in Japan: past, present, and future. Transl Pediatr 2019; 8(3): 199-211.

Development and validation of a questionnaire assessing disease knowledge and self-care (ARKSc) level among allergic rhinitis patients

Ahmad Hazim Hazlami Ahmad Nizar[1,2], Azliana Aziz[1], Baharudin Abdullah[1], Azidah Abdul Kadir[3], Wan Mohd Zahiruddin Wan Mohammad[4], Nor Shahida Abd Mutalib[2]

[1]Department of Otorhinolaryngology-Head & Neck Surgery, School of Medical Sciences, Universiti Sains Malaysia Health Campus, Kubang Kerian, Kelantan, Malaysia, [2]Department of Otorhinolaryngology-Head Neck Surgery, Hospital Sultan Abdul Halim, Sungai Petani, Kedah, Malaysia, [3]Department of Family Medicine, School of Medical Sciences, Universiti Sains Malaysia Health Campus, Kubang Kerian, Kelantan, Malaysia, [4]Department of Community Medicine, School of Medical Sciences, Universiti Sains Malaysia Health Campus, Kubang Kerian, Kelantan, Malaysia

ABSTRACT

Introduction: Understanding a patient's knowledge and self-care level regarding allergic rhinitis (AR) is essential these factors significantly influence treatment outcomes. Currently, there is a lack of validated measurement tools specifically designed to assess disease knowledge and self-care among patients with AR. Therefore, this study aimed to develop and validate a new questionnaire - the Allergic Rhinitis Knowledge and Self-care (ARKSc) questionnaire- for this purpose.

Material and Methods: A questionnaire development and validation study was conducted in two phase. The first phase involved developing a self-administered questionnaire through literature review and consultations with an expert panel. Content validation was evaluated by a group of content experts using the content validity index (CVI), while face validity was assessed by AR patients using the Face Validity Index (FVI). In the second phase, construct validity of the final ARKSc questionnaire was examined at Hospital Sultan Abdul Halim, Kedah, Malaysia and Universiti Sains Malaysia Specialist Hospital (HPUSM), Kelantan, Malaysia involving 136 AR patients. Exploratory factor analysis (EFA) and reliability analysis were performed to assess the factorial structure and internal consistency of the questionnaire.

Results: The preliminary questionnaire included 16 questions (22 items) assessing AR knowledge and 11 questions on self-care. During content validation, three items with low item-CVI (I-CVI) score were removed. The average Scale-CVI (S-CVI/Ave) for both knowledge and self-care domains was 0.83. The scale-level face validity index value (S-FVI/Ave) 0.95, indicating excellent clarity and comprehensibility. Following construct validation, the final version of questionnaire consisted of 11 items in the knowledge section and 4 items in self-care section. The Cronbach's alpha was 0.659 for the knowledge section, and 0.663 for the self-care section, reflecting acceptable internal consistency.

Conclusion: The newly developed and validated ARKSc questionnaire is a valid and reliable instrument for assessing disease knowledge and self-care among patients with allergic rhinitis. This study provides a foundation for future development of more refined tools and underscores the importance of evaluating these domains to enhance disease management and patient outcomes.

KEYWORDS:
Allergic rhinitis, Knowledge, Self-care, Questionnaire, Development, Validation

INTRODUCTION

Allergic rhinitis (AR) is a hypersensitivity disorder affecting the upper respiratory tract and is one of the most common chronic conditions globally, with a reported prevalence ranging from 5% to 40% of the population.[1] In Malaysia, the average prevalence of AR is approximately 7.1% and has shown a rising trend over recent decades.[2] AR contributes significantly to the socioeconomic burden, particularly among the working population, leading to increased healthcare expenditure and productivity loss. Despite its prevalence and associated impact, AR remains underdiagnosed and undertreated worldwide.[3] The clinical manifestations of AR include nasal congestion, rhinorrhea, sneezing, postnasal drip and itchy or watery eye typically triggered by an allergen-induced, immunoglobulin E (IgE)-mediated inflammatory response.[4] These symptoms can adversely affects patients' daily activities, work productivity, self-esteem, social interactions and in children, their learning process.[5] Management of AR involves both pharmacological and non-pharmacological strategies, both of which are vital in achieving optimal symptoms control.[6] Non-pharmacological approaches, such as allergen identification and avoidance, complement pharmacological treatments that include antihistamines, intranasal corticosteroids, and decongestants.

Adequate patient knowledge about AR plays a crucial role in ensuring effective treatment and disease management.

This article was accepted: 17 April 2025
Corresponding Author: Azliana Aziz
Email: az_aziz@usm.my

Inadequate understanding may lead to inappropriate self-treatment and poor disease control, despite the availability of effective therapeutic options. Basic knowledge of AR encompasses its causes, preventive strategies, available treatments and potential complications. A lack of such knowledge often results in poor self-care, ultimately diminishing quality of life. Therefore, evaluating and understanding patients' knowledge and self-care practices is essential for delivering optimal care.

From literature reviews, most existing questionnaires adopt a broad focus, assessing general knowledge, attitudes and behaviours related to allergic rhinitis. Many of these instruments are non-standardised, not validated or designed for one-time use. Furthermore, they are often not adapted to specific clinical settings or reviewed by multidisciplinary experts. In response to these limitations, this study aims to develop and validate a disease-specific questionnaire to assess knowledge and self-care among AR patients, tailored for use in the Malaysian healthcare setting.

MATERIALS AND METHODS

The development and validation of the questionnaire were conducted in two main stages. The first stage focused on the development of the questionnaire, while the second stage involved its validation, which included content and face validity assessment, followed by psychometric validation using Exploratory Factor Analysis (EFA).

Phase 1 : Development of questionnaire

The initial phase involved the development of a self-administered questionnaire, based on comprehensive literature review and expert panels consultations. The literature was reviewed to identify relevant concepts to guide item selection and the formation of questionnaire sections. The expert panel comprised three Otorhinolaryngologists, one Family Medicine physician and one Community Medicine physician.

The newly developed questionnaire aimed to assess the patient's knowledge and self-care related to AR. It was structured into three sections. The first section is the demographic data. This section included eight items capturing demographic characteristics such as age, gender, ethnicity, education level, occupation, family history of atopy (bronchial asthma, atopic eczema or AR), smoking status and time since diagnosis. Section two is regarding the knowledge of AR. This section consisted of 16 questions divided into four domains: aetiology, symptoms, complications and treatment. A total of 22 items were included and each was assessed using a three-option scoring system ("Yes", "No" and "Not sure"). Section three is about Self-care of AR. This section contained 11 items designed to assess self-care practices. A four-point Likert scale was used for responses: "Always", "Often", "Sometimes" and "Never". In total, the initial draft of the questionnaire comprised of 27 questions encompassing 33 items - 22 in the knowledge section and 11 in the self-care section.

Regarding conceptual definitions, knowledge is defined as the information, understanding, and skills aquired through education or experience.[7] The knowledge items aim to assess the patient's understanding of the causes, symptoms, complications and treatment of AR. Self-care, as defined by the World Health Organization (WHO), refers to the ability of individuals, families and communities to promote health, prevent disease, maintain health, and cope with illness and disability, with or without the support from healthcare providers.[8] The self-care items in the questionnaire were developed to assess patients' behaviours, routine and management strategies in relation to AR. The questionnaire was initially developed in the Malay language, considering its intended use in the Malaysian healthcare context.

Phase 2

Content and face validity

Content validity is defined as the extent to which the items of an assessment instrument are relevant to and representative of the intended construct for a specific assessment purpose.[9] This process is crucial as it provides preliminary evidence of construct validity, offers insights into the clarity and appropriateness of the items, and facilitates the refinement of the instrument through expert feedback.[10] The content validity index (CVI) is the most widely used method to evaluate content validity.[11] In this study, five experts were recruited to assess content validity – three Otorhinolaryngologists, one Family Medicine physician and one Community Medicine physician. The Otorhinolaryngologists and family medicine physician served as subject matter experts, while the community medicine physician contributed expertise in questionnaire design.

The draft questionnaire was pretested by the panel to identify potential issues in wording, relevance and interpretation. Both qualitative and quantitative approaches were employed. Each item was rated independently by the experts using a 4-point Likert scale: 1 = not relevant, 2 = somewhat relevant, 3 = relevant, and 4 = very relevant. In addition to that, the panel of experts also gave their recommendations and feedback on whether any modifications need to be made to the items or new items to be added to the questionnaire. Responses are dichotomized for CVI calculation, where the rating of 3 or 4 was scored as 1 (relevant) and ratings of 1 or 2 as 0 (not relevant). The Item-CVI (I-CVI) was calculated by dividing the number of experts who rated the item as relevant (3 or 4) by the total number of experts (12). For example, an I-CVI of 3/5= 0.6 indicates insufficient agreement, with a recommended threshold of 1.0 when using five experts. The Scale-CVI Average (S-CVI/Ave) was then calculated by summing all I-CVI and dividing by the total number of items. The acceptable cutoff for S-CVI/Ave is ≥ 0.83.[12]

Face validity index (FVI) assesses the clarity and comprehensibility of the questionnaire items from the respondent's perspective. This study followed the method proposed by Yusof (2019).[13] Ten AR patients from the Otorhinolaryngology (ORL) clinic at Hospital Sultan Abdul Halim (Kedah, Malaysia) participated in the face validity assessment. These were not part of the actual validation cohort. Each was provided with a printed version of the questionnaire and asked to evaluate each items based on

Table I: Content validation analysis by the panel of experts

	R1	R2	R3	R4	R5	I-CVI
Knowledge						
Q1	1	1	1	1	1	1
Q2	1	1	1	1	0	0.8
Q3	1	1	0	1	1	0.8
Q4	0	1	1	1	1	0.8
Q5	1	1	0	0	1	0.6
Q6	0	0	0	0	0	0
Q7.1	1	1	1	1	1	1
Q7.2	1	1	1	1	1	1
Q7.3	1	1	1	1	1	1
Q7.4	1	1	1	1	1	1
Q8.1	1	1	0	1	1	0.8
Q8.2	1	1	0	1	1	0.8
Q8.3	1	0	0	1	1	0.6
Q8.4	1	0	1	1	1	0.8
Q9	1	0	1	1	0	0.6
Q10	1	1	0	0	1	0.6
Q11	1	1	1	1	1	1
Q12	1	1	0	1	1	0.8
Q13	1	1	0	1	1	0.8
Q14	1	1	1	1	1	1
Q15	1	1	1	1	1	1
Q16	1	1	1	1	1	1
Q17	1	1	0	1	1	0.8
Q18	0	1	0	1	1	0.6
Q19	1	1	0	1	1	0.8
Self-care						
Q1	1	1	1	1	1	1
Q2	1	1	1	1	1	1
Q3	0	1	1	1	1	0.8
Q4	0	1	1	1	1	0.8
Q5	1	1	1	1	1	1
Q6	1	1	1	1	1	1
Q7	1	1	1	1	1	1
Q8	1	1	0	1	1	0.8
Q9	0	1	0	1	1	0.6
Q10	1	1	1	1	1	1
Q11	1	1	1	1	1	1
					S-CVI/Ave	0.83

R - rater
I-CVI - Item Content Validity Index
S-CVI/Ave - Average CVI

clarity and understanding, using a 5-point Likert scale. Items rated as 4 or 5 were considered "very clear and comprehensible" and scored as 1, while those rated from 1 to 3 scored as 0. The Item-face validity index (I-FVI) was calculated by dividing the number of respondents who rated the item as 4 or 5 by the total number of participants. The Scale-FVI Average (S-FVI/Ave) was derived by averaging all I-FVIs. A minimum S-FVI/Ave of 0.83 is considered acceptable.[13] This indicates that the questionnaire items are generally well understood and appropriately worded for the target population. All queries, comments and suggestions from participants during the face validity phase were reviewed and used to refine the questionnaire. The final draft at this stage included three sections with a total of 27 questions (33 items): demographic information (8 items), disease knowledge (22 items) and self-care practices (11 items). This revised version was then used in the subsequent psychometric validation phase.

Psychometric validation study

The psychometric validation was conducted through a cross-sectional study involving Ar patients who attended the ORL Department at Hospital Sultan Abdul Halim (HSAH) and Universiti Sains Malaysia Specialist Hospital (HPUSM), between 16 April 2021 to 31 December 2021. To determine the required sample size for Exploratory Factor Analysis (EFA), a commonly used ratio of 5 participants per item, was applied.[14] With a total of 27 questions across two domains, the target sample size was 135. Convenience sampling was used to recruit participants who met the following inclusion criteria: aged between 18 to 60 years old, clinically diagnosed with AR, positive skin prick test and able to comprehend either Malay or English. Patients with non allergic rhinitis, nasal polyposis or rhinosinusitis were excluded.

Each participant provided informed consent and received an explaination of the study's purpose, procedures, risks, benefits and confidentiality. The questionnaire was self-administered and the completion time was approximately 10 minutes. The finalised questionnaire was then subjected to construct validity analysis through EFA and reliability testing to assess internal consistency.

Table II: Face validation analysis by respondents who were not involved in the actual study

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10		I-FVI
Knowledge												
Q1	1	0	1	1	1	1	1	1	1	1		0.9
Q2	1	0	1	0	1	1	1	1	1	1		0.8
Q3	1	0	1	1	1	1	1	1	1	1		0.9
Q4	1	1	1	1	1	1	1	1	1	1		1
Q5	1	1	1	1	1	1	1	1	1	1		1
Q6	1	0	1	1	1	1	1	1	1	1		0.9
Q7.1	1	1	1	1	1	1	1	1	1	1		1
Q7.2	1	1	1	1	1	1	1	1	1	1		1
Q7.3	1	1	1	1	1	1	1	1	1	1		1
Q7.4	1	1	1	1	1	1	1	1	1	1		1
Q8.1	1	1	1	1	1	1	1	1	1	1		1
Q8.2	1	1	1	1	1	1	1	1	1	1		1
Q8.3	1	1	1	1	1	1	1	1	1	1		1
Q8.4	1	1	1	1	1	1	1	1	1	1		1
Q9	1	0	1	1	1	1	1	1	1	1		0.9
Q10	1	1	1	1	1	1	1	1	1	1		1
Q11	1	1	1	1	1	1	1	1	1	1		1
Q12	1	1	1	1	1	1	1	1	1	1		1
Q13	1	1	1	1	1	1	1	1	1	1		1
Q14	1	0	1	1	1	1	1	1	1	1		
0.9												
Q15	1	1	1	1	1	1	1	1	1	1		1
Q16	1	1	1	1	1	1	1	1	1	1		1
Q17	1	1	1	1	1	1	1	1	1	1		1
Q18	1	0	1	1	1	1	1	1	1	1		
0.9												
Q19	0	0	1	1	1	1	1	1	1	1		
0.8												
Self-care												
Q1	1	1	1	1	1	1	1	1	1	1		1
Q2	1	1	1	1	1	1	1	1	1	1		1
Q3	1	0	1	1	1	1	1	1	1	1		0.9
Q4	1	1	1	1	1	1	1	1	1	1		1
Q5	1	0	1	1	1	1	1	1	1	1		0.9
Q6	1	1	1	1	1	1	1	1	1	1		1
Q7	0	1	1	1	1	1	1	1	1	1		0.9
Q8	0	0	1	1	1	1	1	1	1	1		
0.8												
Q9	0	0	1	1	1	1	1	1	1	1		0.8
Q10	1	0	1	1	1	1	1	1	1	1		0.9
Q11	1	0	1	1	1	1	1	1	1	1		0.9
											S-FVI/Ave	0.95

R – rater
I-FVI - Item-Face Validity Index
S-FVI/Ave - Scale-Face Validity Index

Statistical analysis

Data entry and statistical analysis were performed using Microsoft Excel and IBM SPSS version 26.0 for Windows. Descriptive statistics were used to summarise the demographic characteristics of participants. Numerical data were represented as mean (standard deviation, SD) and frequency (n, %) for categorical variables.

Exploratory Factor Analysis (EFA) was used to assess the construct validity of the questionnaire in both the knowledge and self-care domains. Prior to EFA, Bartlett's Test of Sphericity was conducted to determine whether sufficient correlations among the items.[15] Additionally, the Kaiser-Meiyer-Olkin (KMO) Measure of Sampling Adequacy was used to evaluate the proportion of common variance among variables. A Bartlett's Test p-value of <0.05 and a KMO value ≥0.5 were considered indicative of suitability for factor analysis.[15] Factor loadings were used to extract underlying factors and items with commonalities and factor loadings ≥0.3 were retained for further analysis. The internal consistency of each domain was evaluated using Cronbach's alpha coefficient, with values ≥0.6 considered acceptable for newly developed instruments.[16]

Ethical considerations

This study received ethical approval from the Medical Research and Ethics Committee (MREC) under the National Medical Research Register (NMRR-21-243-58303 (IIR)) and the Human Research Ethics Committee of Universiti Sains Malaysia (JEPeM-USM) under protocol code USM/JEPeM/20090454.

Table III: Demographic data of the participants (N=136)

Variable	Mean (SD)
Age of the respondent	
18 - 29	54 (39.7)
30 - 40	39 (28.7)
41 - 50	20 (14.7)
51 - 60	23 (16.9)
Gender	
Male	40 (29.4)
Female	96 (70.6)
Ethnicity	
Malay	107 (78.7)
Chinese	12 (8.8)
Indian	16 (11.8)
Others	1 (0.7)
Education level	
Primary	5 (3.7)
Secondary	67 (49.3)
Tertiary	61 (44.9)
Others	3 (2.2)
Family history of atopy (Bronchial asthma)	
Yes	50 (36.8)
No	86 (63.2)
Family history of atopy (Atopic eczema)	
Yes	17 (12.5)
No	119 (87.5)
Family history of atopy (Allergic rhinitis)	
Yes	66 (48.5)
No	70 (51.5)
Smoking	
Yes	13 (9.6)
No	117 (86.0)
Ex-smoker	6 (4.4)
Time since diagnosis (years)	7.76 (7.24) *

* Mean (SD)

All participants provided written informed consent prior to participation. They were fully briefed on the purpose of the study, procedures, potential risks and benefits, and their right to withdraw at any time. All collected data were treated with strict confidentiality and used solely for research purposes.

RESULTS

Content validity

The Item-Level Content Validity Index (I-CVI) for the questionnaire items ranged from 0 to 1 (Table I). The Scale-Level Content Validity Index Average (S-CVI/Ave) was 0.83 for both the knowledge and self-care sections, meeting the recommended threshold.

In the knowledge section (25 items initially), 9 items achieved an I-CVI of 1.0, indicating perfect agreement among the expert panel. Ten items with an I-CVI of 0.8 and three items with an I-CVI value of 0.6 were retained following expert panel discussions, as they were considered important to the construct being measured and were subsequently revised. Three items with an I-CVI ≤ 0.6 were removed.

In the self-care section (11 items initially), 7 items had an I-CVI of 1.0, three items had an I-CVI of 0.8 and one item had an I-CVI of 0.6. All items below perfect agreement were retained with modifications after expert consultation. At this stage, the revised questionnaire consisted of 22 items in the knowledge section and 11 items in the self-care section.

Face validity

The Item-Level Face Validity Index (I-FVI) for the knowledge and self-care sections ranged from 0.8 to 1.0 (Table II). The overall Scale-Level Face Validity Index Average (S-FVI/Ave) was 0.95 for both sections, indicating excellent clarity and comprehensibility. Based on feedback form the participants, several items were modified to improve understanding and ease of interpretation.

Psychometric analysis

A total of 136 participants completed the questionnaire. The demographic characteristics of the respondents are presented in Table III. The majority were aged 18 – 29 years old (39.7%), female (70.6%), Malay (78.7%) and had attained secondary education level (49.3%). Most participants reported no family history of bronchial asthma (63.2%), atopic eczema (87.5%) or allergic rhinitis (51.5%) and 86% were non-smokers. The mean time since diagnosis was 7.76 years.

Validation and reliability

EFA was conducted separately for the knowledge and self-care sections. A total of 33 items initially underwent EFA. In the knowledge section, the KMO measure of sampling adequacy was 0.663 and Bartlett's Test of Sphericity was significance ($p<0.001$), confirming the data's suitability for factor analysis. Parallel analysis suggested a three-domain structure: domain 1-typical symptoms, domain 2- related symptoms, domain 3- general information. A total of 10

Table IV: EFA and Reliability of Knowledge and Self-care section

	Factor loading	Cronbach alpha
Knowledge Section		
Domain 1		0.626
Typical symptoms of allergic rhinitis are :		
Gejala-gejala tipikal alahan rinitis adalah seperti :		
Runny nose	0.546	
Hidung berair		
Sneezing	0.670	
Bersin		
Blocked nose	0.428	
Hidung tersumbat		
Nasal itchiness	0.583	
Gatal hidung		
Domain 2		0.787
These symptoms are also related to allergic rhinitis :		
Gejala-gejala berikut juga berkait rapat dengan alahan rinitis:		
Yellow coloured nasal discharge	0.618	
Hingus berwarna kuning		
Nose bleeding	0.897	
Pendarahan hidung		
Facial pain	0.607	
Sakit di bahagian muka		
Loss of smell	0.667	
Hilang deria bau		
Domain 3		
Intranasal corticosteroid spray causes long term side effects	0.723	0.609
Ubat semburan hidung steroid menyebabkan kesan sampingan dalam jangka masa panjang		
Medications to treat allergic rhinitis can cause drug dependence	0.567	
Penggunaan ubat-ubatan bagi merawat alahan rinitis boleh menyebabkan kebergantungan kepada ubat		
Allergic rhinitis is a curable disease	0.397	
Penyakit alahan rinitis boleh disembuhkan		
Overall Knowledge section Cronbach alpha		0.659
Self-care Section		
I adhere to my scheduled clinic appointments	0.576	
Saya hadir sesi temujanji klinik yang ditetapkan		
Avoiding exposure to dust is part of my daily routine	0.403	
Mengelakkan diri daripada terdedah kepada habuk adalah amalan saya		
I use the intranasal steroid spray as instructed by my doctor	0.788	
Saya menggunakan ubat semburan hidung steroid mengikut arahan doktor		
My allergic rhinitis self-care is good.	0.427	
Saya mengamalkan pengurusan kendiri terhadap alahan rinitis yang baik.		
Overall Self-care section Cronbach alpha		0.663

items were removed due to low factor loadings (<0.3). The communality values for all retained items were above 0.3, supporting the validity of the three- domain model. A factor loading value of more than 0.3 is considered acceptable.[17] One additional item was removed from Domain 3 to improve internal consistency. The final Cronbach's alpha for domain 1 was 0.626, domain 2 was 0.787, and domain 3 was 0.609. The overall Cronbach's alpha for all three domains in the knowledge section was 0.659.

In the self-care section, the KMO value was 0.611 and Bartlett's Test of Sphericity was also significant ($p<0.001$). The EFA revelaed that the self-care section was unidimensional. A total of 7 items with low factor loadings (<0.3) were excluded. The final Cronbach's alpha for the self -care section was 0.663. The factor loadings and reliability statistics are presented in Table IV. The final validated ARKSc questionnaire (Appendix 1) consists of 11 items in the knowledge section and 4 items in the self-care section.

DISCUSSION

Allergic rhinitis (AR) remains a significant public health concern due to its increasing prevalence over the past decade. Globally, the prevalence of AR ranges from 10% to 40%.[18] Several contributing factors have been proposed, including increased exposure to pollutants due urbanization, global climate changes, and dietary shifts.[4] AR imposes a considerable socioeconomic burden on healthcare systems and individuals, Its common association with comorbidities such as asthma, atopic dermatitis, allergic conjunctivitis, otitis media and rhinosinusitis further complicates disease

management. And increases treatment costs. Additional, impairment in school and work performance due to AR contributes to productivity loss and overall economic impact. Rhinitis is broadly classified into two major types: AR and nonallergic rhinitis (NAR). AR is characterised as an IgE-mediated disorder of the nasal mucosa triggered by allergen exposure.[19] In contrast, NAR is associated with chronic nasal symptoms, such as nasal obstruction and rhinorrhea, typically caused by nonallergic, non-infectious triggers such as weather changes or exposure to strong odors. The skin prick test remains a widely accepted diagnostic tool for AR due to its ease of use and less-invasive nature.[20] Unlike AR patients, those with NAR usually test negative for allergens via skin prick testing or allergen-specific antibody tests.[21] Therefore, only patients with positive skin prick test were included in this study to accurately distinguish AR from NAR. Patients with AR frequently seek treatment from general practitioners (GPs), health clinics, Otorhinolaryngology clinics or pharmacists, especially during symptoms exacerbations. Physicians play a pivotal role in providing accurate diagnoses and effective management strategies. However, a study conducted in Southeast Asian revealed a significant gap between guidelines-recommended AR management and actual clinical practice.[22] Primary care physicians, particularly GPs, were found to inadequately implement clinical guidelines, underscoring the need for improved disease education and training. Moreover, several studies, including those conducted in the United States and other countries, consistently reported that patients had poor knowledge and inappropriate practices regarding AR.[23-25] These findings highlight the critical importance of patient education in achieving optimum treatment outcomes.

Self-care is another crucial components in AR management. According to WHO, self-care refers to the ability of individuals to manage their health through awareness, self-control, and self-reliance.[26] Effective self-care can significantly reduce symptom burden and disease morbidity, contributing to improved long-term disease control.

This study involved the development and validation of a novel questionnaire to assess AR patients' knowledge and self-care practices. The initial version comprised of 22 items in knowledge domain and 11 items in self-care domain. During the content validity phase, items deemed irrelevant or difficult to understand were removed or revised based on expert panel feedback. As content validity is foundational to the overall validity of a measurement instrument, it was given high priority during tool development. The face validity phase, conducted among non-study patients, was used to assess clarity and relevance from the end-users' perspective. Both quantitative and qualitative feedback were incorporated, resulting in important adjustments to item wording, structure and content.Items with I-CVI or I-FVI scores below acceptable thresholds were revised or removed based on rater recommendations. The validation phase of the questionnaire is important as it express the degree to which a measurement measures what it aims to measure.[27] Meanwhile, the reliability measured the degree to which the result obtained by a measurement and procedure can be replicated.[27]

Following these steps, the questionnaire underwent EFA. In the knowledge section, three domains emerged- typical symptoms, related symptoms and general information, with acceptable Cronbach's alphas values (0.626, 0.787, and 0.609 respectively). The overall reliability of the knowledge section was 0.659, indicating acceptable internal consistency for a newly developed tool. The self-care section revealed a unidimensional structure, with a Cronbach's alpha of 0.663, also reflecting acceptable reliability.

The final validated questionnaire includes 11 items in the knowledge section and 4 items for the self-care section. The retained knowledge items focus primarily on AR symptoms, aligning with the key elements that affect patient recognition and response to their condition.. Despite the item reduction, the instrument remain robust and effective in assessing patient understanding and management practices.

A unique strength of the ARKSc questionnaire lies in its development and validation in Malay, with bilingual (Malay and English) version available to ensure accessibility. Designed specifically for AR, the items reflect real-world symptoms and treatment behaviours. The questionnaire was intentionally constructed to be clear, easy to score and comprehensible to the target populations, thereby improving response accuracy. In clinical settings, this tool can help assess whether patients have adequate disease education and whether further intervention is needed. It also directly evaluates patient-driven self-management behaviours.

Furthermore, this validated tool provides a platform for future research to assess knowledge and self-care levels among AR patients, identify associated factors, and evaluate the impact of educational or behavioural interventions.

CONCLUSION

The newly developed and validated ARKSc questionnaire is a valid and reliable instrument for measuring disease knowledge and self-care levels among patients with allergic rhinitis. This study offers an essential foundation for future refinement and supports the importance of assessing these sections as part of comprehensive AR management. Enhancing patients' knowledge and self-care capacity is fundamental to improving treatment outcomes and reducing the burden of allergic rhinitis.

ACKNOWLEDGEMENT

Thank you to Universiti Sains Malaysia for the grant (Geran Penyelidikan Sarjana Perubatan PPSP (GPSP) 1001/PPSP/8070011) which makes the study's objectives achievable.

The questionnaire also has been copyrighted and registered under MyIPO (LY2024P02979).

DECLARATION OF CONFLICT OF INTEREST

The authors declare that all primary investigators and co-researchers have no conflict of interest.

REFERENCES

1. Annesi-Maesano I, Didier A, Klossek M, Chanal I, Moreau D, Bousquet J. The score for allergic rhinitis (SFAR): a simple and valid assessment method in population studies. Allergy. 2002; 57(2): 107-14.
2. Katelaris CH, Lai CKW, Rhee C-S, Lee SH, De Yun W, Lim-Varona L, et al. Nasal Allergies in the Asian–Pacific Population: Results from the Allergies in Asia–Pacific Survey. Am J Rhinol Allergy 2011; 25(5_suppl): S3-15.
3. Maurer M, Zuberbier T. Undertreatment of rhinitis symptoms in Europe: findings from a cross-sectional questionnaire survey. Allergy 2007; 62(9): 1057-63.
4. Keith PK, Desrosiers M, Laister T, Schellenberg RR, Waserman S. The burden of allergic rhinitis (AR) in Canada: perspectives of physicians and patients. Allergy, Asthma Clin Immunol 2012; 8(1): 7.
5. Caballer B de la H, Rodríguez M, Fraj J, Cerecedo I, Antolín-Amérigo D, Colas C. Allergic Rhinitis and Its Impact on Work Productivity in Primary Care Practice and a Comparison with Other Common Diseases: The Cross-sectional Study to Evaluate Work Productivity in Allergic Rhinitis Compared with Other Common Diseases (CAPRI) Study. Am J Rhinol Allergy 2012; 26(5): 390-4.
6. Schutzmeier P, Kutzora S, Mittermeier I, Becker J, Bergmann K, Böse-O'Reilly S, et al. Non-pharmacological interventions for pollen-induced allergic symptoms: Systematic literature review. Genuneit J, editor. Pediatr Allergy Immunol 2022; 33(1): e13690.
7. Oxford Learner's Dictionaries. 2023. Available from: https://www.oxfordlearnersdictionaries.com/ [cited 2023 Jan 17]
8. Self-care in the context of primary health care. World Health Organization. Regional Office for South-East Asia. 2009 [cited 2023 Jan 17]
9. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med 2006; 119(2): 166.e7-166.e16.
10. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A-R. Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. J caring Sci 2015; 4(2): 165-78.
11. Devon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. J Nurs Scholarsh an Off Publ Sigma Theta Tau Int Honor Soc Nurs 2007; 39(2): 155-64.
12. Saiful M, Yusoff B. ABC of Content Validation and Content Validity Index Calculation. undefined. 2019; 11(2): 49-54.
13. Yusoff MSB. ABC of Response Process Validation and Face Validity Index Calculation. Educ Med J. 2019; 11(3): 55-61.
14. Rizopoulos D. ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses. Vol. 17, JSS Journal of Statistical Software. 2006. [cited 2023 Jan 17]
15. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. Int J Nurs Stud 2014; 51(3): 511-21.
16. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis 7th ed. Upper Saddle River (NJ): Pearson Prentice Hall; 2009.
17. Yusoff MSB, Arifin WN, Hadie SNH. ABC of Questionnaire Development and Validation for Survey Research. undefined. 2021; 13(1): 97-108.
18. Brožek JL, Bousquet J, Agache I, Agarwal A, Bachert C, Bosnic-Anticevich S, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines-2016 revision. J Allergy Clin Immunol 2017; 140(4): 950-8.
19. Bousquet J, Khaltaev N, Cruz AA, Denburg J, Fokkens WJ, Togias A, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). Allergy. 2008; 63 Suppl 8(SUPPL. 86): 8-160.
20. Nevis IF, Binkley K, Kabali C. Diagnostic accuracy of skin-prick testing for allergic rhinitis: a systematic review and meta-analysis. Allergy, Asthma Clin Immunol 2016; 12(1): 20.
21. Tran NP, Vickery J, Blaiss MS. Management of Rhinitis: allergic and non-allergic. Allergy Asthma Immunol Res 2011; 3(3): 148-156.
22. Abdullah B, Snidvongs K, Recto M, Poerbonegoro NL, Wang DY. Primary care management of allergic rhinitis: a cross-sectional study in four ASEAN countries. Multidiscip Respir Med 2020; 15(1): 726.
23. Gupta RS, Kim JS, Springston EE, Smith B, Pongracic JA, Wang X, et al. Food allergy knowledge, attitudes, and beliefs in the United States. Ann Allergy Asthma Immunol 2009; 103(1): 43-50.
24. Nolte H, Nepper-Christensen S, Backer V. Unawareness and undertreatment of asthma and allergic rhinitis in a general population. Respir Med 2006; 100(2): 354-62.
25. Alreshidi FM, Alrashidi AS, Alshammari FNM, Qadi ABA, Alrashidi AGM, Alghaythi SMS, et al. Knowledge, Attitude and Practice About Allergic Rhinitis in Saudi Arabia, 2017. Egypt J Hosp Med 2017; 69(4): 2199-203.
26. Martínez N, Connelly CD, Pérez A, Calero P. Self-care: A concept analysis. Int J Nurs Sci 2021; 8(4): 418-25.
27. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger. Postgrad. Med J 2015; 22(4): 195.

Appendix 1
Questionnaire copyright under MyIPO: LY2024P02979

QUESTIONNAIRE ASSESSING DISEASE KNOWLEDGE AND SELF-CARE LEVEL AMONG ALLERGIC RHINITIS PATIENTS

DEMOGRAPHIC SECTION
BAHAGIAN DEMOGRAFIK

Please tick (V) or write the appropriate answer to each question.
Sila tandakan (V) atau isikan jawapan yang berkenaan bagi setiap soalan.

1. **Age :** .....
   Umur : .....

2. **Gender :** Male ☐ Female ☐
   Jantina : *Lelaki* *Perempuan*

3. **Ethnicity :** Malay ☐ Chinese ☐ Indian ☐ Others ☐
   Etnik : *Melayu* *Cina* *India* *Lain-lain*

4. **Education level :** Primary ☐ Secondary ☐ Tertiary ☐ Others ☐
   Tahap pelajaran : *Rendah* *Menengah* *Tinggi* *Lain-lain*

5. **Occupation :** .....
   Pekerjaan : .....

6. **Family history of atopy :** Bronchial asthma ☐ atopic eczema ☐ allergic rhinitis ☐
   Sejarah penyakit alergi di dalam keluarga : *Asma* *Ekzema* *Alahan rinitis*

7. **Smoking :** Yes ☐ No ☐ Ex-smoker ☐
   Merokok : *Ya* *Tidak* *Bekas perokok*

8. **Time since diagnosis (years) :** .....
   Jangka masa selepas diagnosis dilakukan *(tahun) :* .....

SECTION 2 : SELF-CARE OF ALLERGIC RHINITIS
BAHAGIAN 2 : PENGURUSAN KENDIRI ALAHAN RINITIS

Please tick (√) the appropriate answer for each question.
Sila tandakan (√) bagi jawapan yang berkenaan untuk setiap soalan.

QUESTIONS *SOALAN*	ANSWER *JAWAPAN*			
	ALWAYS *SENTIASA*	OFTEN *KERAP*	SOMETIMES *KADANG-KADANG*	NEVER *TIDAK PERNAH*
1) I adhere to my scheduled clinic appointments. *Saya hadir sesi temujanji klinik yang ditetapkan.*				
2) Avoiding exposure to dust is part of my daily routine. *Mengelakkan diri daripada terdedah kepada habuk adalah amalan saya.*				
3) I use the intranasal steroid spray as instructed by my doctor. *Saya menggunakan ubat semburan hidung steroid mengikut arahan doktor.*				
4) My allergic rhinitis self care is good. *Saya mengamalkan pengurusan kendiri terhadap alahan rinitis yang baik.*				

SECTION 1 : KNOWLEDGE OF ALLERGIC RHINITIS

BAHAGIAN 1 : PENGETAHUAN BERKENAAN ALAHAN RINITIS

Please tick (√) the appropriate answer for each question.

Sila tandakan (√) bagi jawapan yang berkenaan untuk setiap soalan.

KNOWLEDGE QUESTION *SOALAN PENGETAHUAN*	ANSWER *JAWAPAN*		
	YES *YA*	**NO** *TIDAK*	**NOT SURE** *TIDAK PASTI*
1. Typical symptoms of allergic rhinitis are : *Gejala-gejala tipikal alahan rinitis adalah seperti :*			
1) Runny nose *Hidung berair*			
2) Sneezing *Bersin*			
3) Blocked nose *Hidung tersumbat*			
4) Nasal itchiness *Gatal hidung*			
2. These symptoms are also related to allergic rhinitis : *Gejala-gejala berikut juga berkait rapat dengan alahan rinitis :*			
1) Yellow coloured nasal discharge *Hingus berwarna kuning*			
2) Nose bleeding *Pendarahan hidung*			
3) Facial pain *Sakit di bahagian muka*			
4) Loss of smell *Hilang deria bau*			
3. Intranasal corticosteroid spray causes long term side effects. *Ubat semburan hidung steroid menyebabkan kesan sampingan dalam jangka masa panjang.*			
4. Medications to treat allergic rhinitis can cause drug dependence. *Penggunaan ubat-ubatan bagi merawat alahan rinitis boleh menyebabkan kebergantungan kepada ubat.*			
5. Allergic rhinitis is a curable disease. *Penyakit alahan rinitis boleh disembuhkan.*			

ORIGINAL ARTICLE

Credentialing & privileging an important factor for diagnostic accuracy of diabetic retinopathy screening in the health clinics of Penang

Nasim Abdul Kuthoose[1], Ahmad Taufik Jamil[1], Khalid Ibrahim[1], Rohana Taharin[2], Ang Ee Ling[3], Farzaana Adam[4]

[1]Department of Public Health Medicine, UiTM, [2]Department of Ophthalmology, Hospital Bukit Mertajam, [3]Department of Ophthalmology, Hospital Pulau Pinang, [4]Public Health Division, Penang State Health Department

ABSTRACT

Introduction: Early detection and management of diabetic retinopathy (DR) is crucial in preventing blindness. Screening is recommended at diagnosis and yearly for Type 2 diabetes patients. DR screening using non-mydriatic fundus cameras (NMFC) has been extended to Health Clinics since 1997, but competency and experience of the medical officers (MOs) remain an issue in Primary Care. This study aims to assess the accuracy of the eye examination using NMFC for DR screening done by MOs and identify the factors associated with the accuracy.

Materials and Methods: This is a cross-sectional study using secondary data obtained from the Penang State Health Department, which conducted clinical audits of fundus images from health clinics in 2019. The audit involved two consultant ophthalmologists to comment on the accuracy of the interpretation of retinal images and the quality of the images sampled from all health clinics with NFMC. Sampling was performed on the audited data set to include diagnosis by the MOs and diabetic retinopathy spectrum of disease. The subject of the study was the images with the corresponding reports. The outcome of this study was the accurate interpretation of the images, as commented by the ophthalmologists. The independent variables studied were the demographic of the MOs who interpreted the images, their training background and the quality of the images.

Results: The Universal sampling method was used, and the final 1129 images fulfilled the eligible criteria. The sensitivity, specificity, PPV and NPV were found to be 80.6%, 92.7%, 76.4% and 94.2%, respectively. Overall accuracy was 83.8%. After missing values were managed, 997 samples were analysed using logistic regression. The final model shows that significant factors associated with accuracy are foreign graduates MOs (Adjusted OR 1.98, 95% CI: 1.35-3.07), MOs with Credentialing & Privileging (Adjusted OR 2.32, 95% CI:1.32-2.88) and Good Image Quality (Adjusted OR 3.62,95% CI:2.37-5.71).

Conclusion: MO with C&P showed better accuracy than MO without C&P. This study suggests that MOH should emphasise the C&P when performing this procedure in health clinics. As image quality showed the highest association with accuracy, strengthening the C&P among the paramedics who perform the procedure using NMFC to get the retinal images is also necessary. This study also indicates that evaluating DR screening programs in health clinics is necessary nationwide.

KEYWORDS:
Non-Mydriatic Fundus Camera, Diabetic Retinopathy Screening, Health Clinics, Credentialing, Privileging, Training, Retinal Images, Fundus Images

INTRODUCTION

Diabetic Retinopathy (DR) is a condition with progressive retinal damage that occurs due to microvascular complications of diabetes mellitus. The global prevalence of DR among diabetic patients in 2021 was 22.27%,[1] and the National Diabetic Registry in 2020 shows that DR prevalence in Malaysia was 11.52%.[2] Furthermore, Diabetes Mellitus (DM) prevalence among adults in Malaysia has increased steadily over the past decade from 11.2% in 2011 and 13.4% in 2015 to 18.3% in 2019.[3] With the increasing trend in diabetic patients, DR prevalence and eye blindness from DR are projected to increase if no proper intervention is done.

Early detection of DR among DM patients with prompt management will significantly prevent blindness. Hence, it is crucial to have the screening at diagnosis and yearly for Type 2 DM patients.[4] Undiagnosed DR or late-stage diagnosis of DR is one of the problems identified. As reported in a study by Nor Farizah Ngah in 2020, 9% of the 3532 sampled patients had undiagnosed DR.[5] In another study by the University of Malaya in 2005, 28% of 217 diabetic patients sampled had undiagnosed PDR, a late-stage DR.[6] This is most likely due to the screening initiatives not being widely available then or not being detected by the treating physician for DM.

Malaysian Guidelines on DR Screening recommend using the Non-Mydriatic Fundus Camera (NMFC) as a screening tool for DR as it is the best modality of the screening tool for DR with 92% Sensitivity and 97% Specificity. This can ensure a high detection rate during screening and prevent delayed diagnosis.[7] However, NMFC was only made available at the ophthalmology clinic at the tertiary centre and operated by

This article was accepted: 21 April 2025
Corresponding Author: Ahmad Taufik Jamil
Email: atjamil@gmail.com

well-trained medical staff in the presence of an ophthalmologist.8 Given the increasing number of diabetic patients and the need for yearly screening for many patients using the best screening modality, DR screening using NMFC has been extended to Health Clinics since 1997.[8] The initiative is necessary as Health Clinics are the gatekeepers and the place where most DM patients in Malaysia are managed. The number of NMFCs made available in Health Clinics with diabetic management services continues to grow.

Two concerns raise doubts about whether extending NMFC to health clinics will achieve its intended purpose. One of them is the competency of the medical officers (MOs) in grading the fundus images. Fundus image grading was not part of the checklist of compulsory procedures doctors must perform during their horsemanship program in Malaysia.[9] Furthermore, MOs in health clinics do not have enough exposure and systematic training on par with those in the Ophthalmology Department in grading the fundus images.

Addressing this issue, MOH has developed a training module for medical staff to operate the machine and interpret the images since 2016 and incorporated it into the National Standard Program for Credentialing & Privileging (C&P).[8] However, the C&P for this procedure is not compulsory but highly recommended to MOs involved in grading the fundus images. Currently, no research has been conducted to assess whether using NMFC for DR screening in health clinics effectively detects DR in its early stages with an acceptable detection rate.

Realising this gap, the Penang State Health Department conducted an Audit in 2019 with more than 1600 images sampled and the reports by the MOs were commented on by the Ophthalmologists. These images are from about 800 patients. This is about 3% of the total screenings done in 2019.[10] However, the audit conducted was not specific to DR and the outcome was not measured using inferential statistics. Furthermore, factors associated with accuracy were not studied.[10] Therefore, this study examined the accuracy of interpretation of retinal images from NMFC for DR screening conducted by MOs and identified the factors associated with the accuracy, utilising the audit data from the Penang State Health Department.

MATERIALS AND METHODS

Study Design, Location, Period and Study Population

This is a cross-sectional study using secondary data. The secondary data refers to granular data from the audited image reports and ophthalmologist comments. Hence, the study population was Retinal Images, which were included in the clinical audits conducted in 2019 by the Penang State Health Department. The study period was from the 20th of March 2023 – 30th June 2023.

Process of Audit of Fundus Images in Penang (2019)

Penang State Health Department conducted audits of retinal images from health clinics in 2019 to evaluate the DR screening program in Health Clinics. A random sample of retinal images with reports was obtained from all health clinics with NMFC. Overall, 1632 images from 816 patients were included as samples, regardless of whether there were findings of the eye diseases (DR, Hypertensive Retinopathy, Glaucoma, etc.) or normal findings. Each image was given a code as a reference. The details of the reports, such as diagnosis, care plan and name of the medical officers who graded the images, were transferred to Form A. The images (in softcopy with codes) were sent together with Form A to two Ophthalmologist Consultants, each had half of the total reports to evaluate the diagnosis and plan of care. They commented based on the report (Form A) in Form B. Details in Form B are the Quality of images (satisfactory or poor), the diagnosis outcome (correct, partially correct and incorrect) and the appropriateness of the management (correct, partially correct and incorrect). Both Forms A & B were returned to the Penang State Health Department.

Sample Size and Sampling Method

Previous studies on the sensitivity and specificity of non-ophthalmologist graders of retinal images from Thailand and factors associated with Family Physicians' Performance on Competency assessment were cited for sample size calculations.[11-12] The calculation was done using Unmatched Case-Control (Fleiss with continuity correction Statistical Methods for Rates and Proportions) sampling calculation with 95% CI and 80% Power of Study. Ratio case over control 1.0. From the previous study, variable outcomes related to less percentage of unsafe competency are graded as the control group. The biggest sample size was selected, which was 816 samples of images with their respective audit records (Form A and Form B). The sampling frame is the list of Image Codes, and the sampling method was universal sampling methods in view of the strict eligible criteria of DR diagnosis.

Eligible Criteria from the Audit Data

Only DR eye diseases such as NPDR, PDR and Maculopathy and normal images reported by MOs were included in this study. The images that were not gradable by MOs or ophthalmologists were excluded.

Variables Studied

The subject of this study is the retinal images. The outcome variables studied were the diagnosis status (correct or incorrect) based on the audit data. As for Sensitivity, Specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV), the outcome was defined as True Positive, True Negative, False Positive and False Negative.

The independent variables studied are illustrated in Figure 1.

Data Management

The request was made to the Penang State Health Department to excess all reports of Form A and Form B. After the reports were screened and eligible images identified, data from the reports were transcribed into tables. MOs' identities were concealed, and the table was only made available for the desired variables related to the MOs, such as age, gender, years of experience, etc. This file is saved in a protected Word file with a password once the variables on the characteristics and training background of the medical officers are obtained. There is no possible bias that can be addressed in this secondary data collection and analysis process.

Table I: Sensitivity, Specificity, PPV & NPV of the NMFC Images Interpretation among MOs

Type of Analysis	Type of Scenario		
	Best Case (n=1129)	Worst Case (n=1129)	Absolute Case (n=1046)
True Positive	274	191	191
True Negative	750	750	750
False Positive	59	142	59
False Negative	46	46	46
Sensitivity	85.6%	80.6%	80.6%
Specificity	92.7%	84.1%	92.7%
PPV	82.3%	57.4%	76.4%
NPV	94.2%	94.2%	94.2%

Table II: Multiple Logistic Regression on the Factor Related to the Accuracy

Variables	B	S.E	Wald	df	Sig.	Exp(B)	95% C.I. for EXP (B)	
							Lower	Upper
Undergraduate Institution of the MO								
Local University (Reference)								
Foreign University	0.69	0.19	13.03	1	0.00	1.98	1.35	3.07
C&P Status of the MO								
No (Reference)								
Yes	0.84	0.18	21.63	1	0.00	2.32	1.23	2.88
Image Quality								
Poor (Reference)								
Satisfactory	1.29	0.21	36.58	1	0.00	3.62	2.37	5.71

Ethical Consideration

As explained above, this secondary data contains no names or personally identifiable variables. The database that contains the respective input of concern for variables is given study IDs, and the analysis is kept in a protected file. The excess will only be granted upon request for study validation purposes and improvement of service purposes.

This study was ethically cleared by the UiTM Ethical Committee and the Medical Research & Ethics Committee, Ministry of Health Malaysia (NMRR ID-24-00966-VFS).

Statistical Analysis Plan

IBM SPSS version 26 was used for statistical analyses. All the data are categorised and expressed as frequencies and percentages.

The operational definition for the accuracy analyses are as follows:

a) True positive (TP) = number of retinal images with positive findings correctly diagnosed by the MOs
b) True Negative (TN) = number of normal images correctly reported by the MOs
c) False Positive (FP) = number of retinal images with normal findings but incorrectly diagnosed by MOs.
d) False Negative (FN) = number of retinal images with positive findings but incorrectly graded as normal by MOs.

These definitions were extended to include the calculations of Sensitivity, Specificity, PPV and NPV:[13]

Sensitivity = TP/ (TP + FN)
Specificity= TN/ (TN + FP)
PPV= TP/ (TP + FP)
NPV= TN/ (TN + FN)
Overall Accuracy: (TP + TN)/ (TP + TN + FP + FN)

To elucidate the factors associated with image interpretation accuracy, simple logistic regression (SLR) was conducted. Subsequently, Multiple Logistic Regression (MLR) analysis was performed, incorporating adjustments for all variables. Statistical significance was determined at a threshold of $p<0.05$.

RESULTS

Sample Outcomes and Missing Data Management

From the 1632 images, 1129 samples were included in the descriptive analysis for Sensitivity, Specificity, PPV and NPV. The details of the process are shown in Figure 2 and the findings were presented in Table I. The secondary data for the dependent variable has three categories (correct, partially correct and incorrect), so the findings were classified according to three different scenarios. The Best Case Scenarios defined partially correct as correct, the Worst Case Scenarios defined partially correct as incorrect and in the Absolute Case Scenarios, partially correct was omitted.

Missing values were found in the variable of MO's age during the procedure (19.6%), years in service (2%) and years of experience in Primary Care during the procedure (5%). The missing values are also found in the variables for C&P Status (3%), Post Graduate Training Status (8%), Undergraduate Institutions Status (2%) and Image Quality (only one image). These resulted in 225 images (19.9%) with at least one missing value. MCAR (Missing Completely At Random) test was then performed, which resulted in the chi-square value of 0.525 with 1 degree of freedom (df), and the associated p-value was 0.469. This suggests that the missing data are missing completely at random. The imputation method was used to fill in missing MO age values based on common ages (mode/mean) corresponding to Years in Service. Listwise deletion was applied to the other variables' missing values,

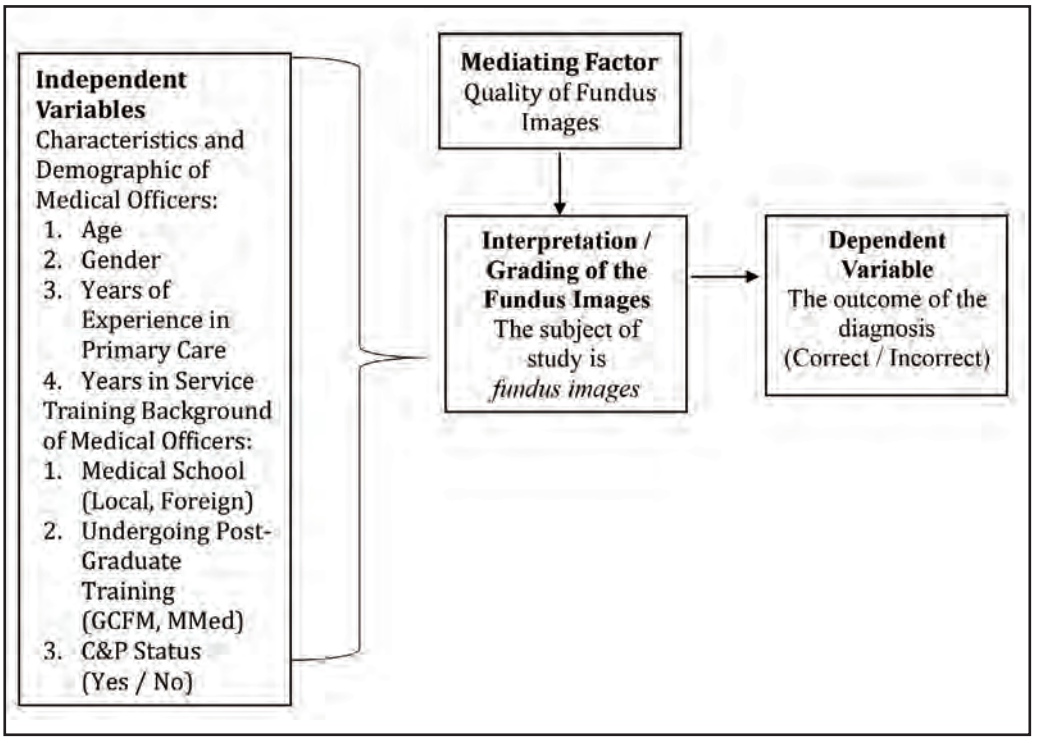


Fig. 1: Independent and Dependent Variables Studied

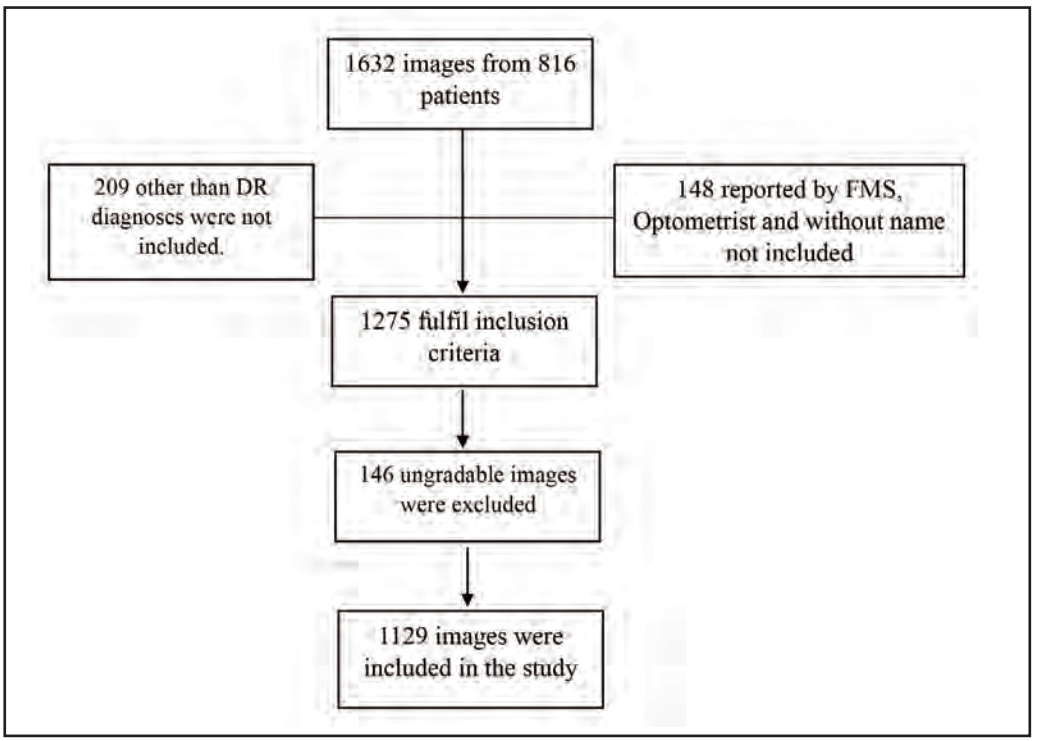


Fig. 2: The outcome of the data sampling process

which led to the total final data included in the analytical analysis, which was 997 image reports. In this analysis, 'partially correct' findings were regrouped as 'incorrect' to maintain dichotomous outcomes.

Logistic Regression

Overall, the accuracy of the examination was 83.8%. All the variables were then entered into MLR using backward and forward LR methods for the adjustments. This process resulted in only 3 significant variables, which are significant and fit well with the data (Hosmer-Lemeshow Test p-value; 0.815).

The statistically significant variables were Undergraduate Institution Status, C&P Status and Image Quality. Multicolinearity was checked, and all VIF values ranged between 1 and 2, indicating no significant collinearity between the variables. No interaction was found between these three variables. The outcome of the MLR is shown in Table II. High adjusted OR for correct diagnosis was found in images which were graded satisfactory by the Ophthalmologists (Adjusted OR 3.62, 95% CI:2.37-5.71) followed by C&P status (Adjusted OR 2.32, 95% CI: 1.23-2.88). The above result also showed that MOs who graduated from foreign universities have better accuracy of DR using NMFC (Adjusted OR 1.98, 95%CI:1.35-3.07). The logistic regression model yielded a Cox & Snell R-Square value of 0.65, indicating that approximately 65% of the variation in the outcome of the accuracy of DR screening using NMFC is explained by the Undergraduate Institution and C&P status of the MOs as well as the Image Quality.

DISCUSSION

This study shows that in absolute case scenarios, the sensitivity and specificity of DR screening using NMFC by MOs in Health Clinics of Penang were 80.6% and 92.7%, respectively. Multiple studies have been conducted to evaluate the accuracy of DR using NMFC with reference to Ophthalmologists.

Somanguan conducted a diagnostic study on non-expert retinal image graders in Thailand health clinics in July 2021 and found that the sensitivity for DR was only 67%.[14] Another study conducted in 2021 among the non-expert medical staff at the Peripheral Health Facilities of Bangladesh found that the sensitivity was 87% and the specificity was 93%.[15] A Pilot Study was conducted in 2018 to look for concordance of Optometrist findings with Ophthalmologists, and the result shows a high concordance of 87.0% with Ophthalmologists.[16] Hence, this study demonstrates results similar to non-expert retinal image grades in other studies.

Even though SLR showed no significant association between C&P status and accuracy, after MLR was performed to adjust with all variables, C&P was shown to have a significant association with accuracy. This shows the importance of C&P in increasing retinal image interpretation skills among MOs. Increasing the number of MOs with C&P or making it compulsory can improve the overall accuracy of this screening procedure.

So far, no study has been done to associate any factors towards the outcome of retinal image grading/interpretation accuracy among doctors nationally or internationally. However, multiple studies have been conducted to associate factors with certain competencies and practices among doctors and other medical professions. A study of 683 physicians referred to the Center for Personalized Education for Physicians in Denver between 2000 and 2010 found that board certification and matched training were associated with safe assessment outcomes.[12]

A study in 2014 regarding factors influencing diagnostic accuracy and management in acute surgical patients in a hospital in the UK found that the consultant was most likely to record a correct diagnosis (75%), followed by SHO (61.3%) and SROC (61.1%).[17] Another study in Shanghai in 2019 found that educational background and job training were among the factors affecting family doctor competency.[18] In all of these studies, job or specific training has consistently been associated with competency, which is similar to this study regarding C&P.

Furthermore, according to this study, foreign graduate Medical Officers (MOs) have higher accuracy in interpreting retinal images. This is likely due to the curriculum provided by foreign institutions, which emphasises this skill more. The Ministry of Health (MOH) should inform local universities about this matter so that they can improve their curriculum and place more emphasis on retinal image interpretation.

Image quality plays an utmost important role in affecting the judgment of the MOs in the interpretation. Good image quality has been shown to have an adjusted OR of a correct diagnosis of 3.62 (95% CI:2.37-5.71), the highest OR in this study. A study was conducted on the assessment of image quality on colour fundus retinal images using automatic retinal image analysis, and the result showed that even software has a significant difference in sensitivity and specificity with regard to image quality.[19] Retinal images in health clinics are taken by the paramedics while the medical officer interprets them.[8] Hence, it is important for the MOH to train the paramedics concurrently and make the C&P on operating the NMFC compulsory for them to perform this procedure.[8]

The latest technological advancement has brought artificial intelligence (AI) to diagnose fundus images. Datuk Dr Nor Farizah Ngah presented MOH's latest initiative, DR.MATA, at a National Institute of Health session on 25th May 2023. She stated that DR.MATA is an AI developed by MOH that is to be implemented in health clinics to interpret fundus images. A study of more than 14000 images showed the sensitivity of this AI is 87.17%, the specificity is 97.17%, and the accuracy is 93.3%.[20] These values are slightly superior to those of the findings from this study, but the difference is minimal. Utilising AI is a timely initiative by MOH. However, if there are technical or financial problems maintaining the system, MOH should consider training and providing more MOs with C&P. This alternative can be achieved with a relatively smaller budget and have a more sustainable impact. A cost-effective analysis can be done to compare the efficacy of both initiatives.

This study has some limitations, mainly due to the use of secondary data. Even though the missing values are not more than 20% of the overall data, conducting the study primarily will improve the availability of the data set to be analysed. This study may serve as baseline findings that can be generalised to other states in Malaysia as there is no significant geographical impact on the accuracy of the eye examination using NMFC.

CONCLUSION

NMFC is an important modality of the DR screening programme in Malaysia. It has been extended to Primary Care to increase the screening of DR among DM patients,

which is growing in numbers as timely diagnosis and prompt treatment can prevent blindness from DR. This diagnostic accuracy study suggested that the MOs, which play the most crucial role in interpreting the images have adequate interpretation capabilities given that they are provided with C&P. Foreign graduates and good image quality shows significant association towards correct interpretation apart from C&P status. The MOH should improve C&P procedures for both MOs and paramedics and ensure more of them receive training. It is recommended that the MOH conduct multicentered studies throughout Malaysia to evaluate the effectiveness of the screening program, as millions of ringgit have been invested in this program, and the prognosis of patients with DR greatly depends on it.

ACKNOWLEDGMENT

We thank the Director General of Health Malaysia for his permission to publish this article.

DECLARATION OF INTEREST

All authors have declared no conflicts of interest in this study, and this research was not funded by any institution.

REFERENCES

1. Teo ZL, Tham YC, Yu M, Chee ML, Rim TH, Cheung N, et al. Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045: Systematic Review and Meta-analysis. Ophthalmology 2021; 128(11): 1580-91
2. Ministry of Health Malaysia. National Diabetes Registry Report 2020. Putrajaya: Disease Control Division, Ministry of Health Malaysia; 2021. Available from: https://leadncdmalaysia.com/wp-content/uploads/2021/08/NDR-Report-2020-FINAL.pdf
3. Institute for Public Health, National Institutes of Health, Ministry of Health Malaysia. National Health and Morbidity Survey (NHMS) 2019: Non-communicable diseases, healthcare demand, and health literacy—Key Findings. Shah Alam: Institute for Public Health; 2020. Available from: https://iptk.moh.gov.my/images/technical_report/2020/4_Infographic_Booklet_NHMS_2019_-_English.pdf
4. Ministry of Health Malaysia, Academy of Medicine Malaysia, Endocrine & Metabolic Society Malaysia, Family Medicine Specialist Association Malaysia. Management of type 2 diabetes mellitus. 6th ed. Putrajaya: Ministry of Health Malaysia; 2021
5. Ngah NF, Muhamad NA, Asnir ZZ, Abdul Aziz RA, Mhad Kassim Z, Sahar SA, et al. Descriptive assessment on diabetic retinopathy screening in an awareness programme in Malaysia. Int J Ophthalmol 2020; 13(11): 1808-13.
6. Tajunisah I, Nabilah H, Reddy SC. Prevalence and risk factors for diabetic retinopathy--a study of 217 patients from University of Malaya Medical Centre. The Medical journal of Malaysia 2006; 61(4): 451-6.
7. Ministry of Health Malaysia, Malaysian Society of Ophthalmologists, Academy of Medicine Malaysia. Clinical practice guideline on screening of diabetic retinopathy. 1st ed. Putrajaya: Ministry of Health Malaysia; 2011.
8. Ministry of Health Malaysia, Diabetic Retinopathy Screening Team. Training module for healthcare providers. 2nd ed. Putrajaya: Ministry of Health Malaysia; 2017.
9. Zai A. HO logbook checklist [Internet]. Dr Zai's Medical Notes. 2019 [cited 2025 Feb 16]. Available from: https://mednote.home.blog/2019/04/09/ho-logbook-checklist/
10. Kamaruddin M. Pelaksanaan audit kualiti saringan diabetik retinopati menggunakan kamera fundus. Georgetown (MY): Public Health Division, Penang State Health Department; 2019 Aug 20.
11. Srihatrai P, Hlowchitsieng T. The diagnostic accuracy of single- and five-field fundus photography in diabetic retinopathy screening by primary care physicians. Indian Journal of Ophthalmology. 2018 Jan 1;66(1):94-7.
12. Grace ES, Wenghofer EF, Korinek EJ. Predictors of physician performance on competence assessment: Findings from CPEP, the Center for Personalized Education for Physicians. Academic Medicine. 2014; 89(6): 912-9.
13. Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol 2008; 56(1): 45-50.
14. Ausayakhun S, Snyder BM, Nanegrungsunk O, Apivatthakakul A, Narongchai C, Melo JS, Keenan JD. Clinic-based eye disease screening using non-expert fundus photo graders at the point of screening: diagnostic validity and yield. American Journal of Ophthalmology 2021; 227: 245-53.
15. Begum T, Rahman A, Nomani D, Mamun A, Adams A, Islam S, Khair Z, Khair Z, Anwar I. Diagnostic accuracy of detecting diabetic retinopathy by using digital fundus photographs in the peripheral health facilities of Bangladesh: validation study. JMIR Public Health and Surveillance. 2021; 7(3): e23538.
16. Hussin DA, Hendicott P, Carkeet A, Baker P, Chen AH, Asnir ZZ, Abidin ZZ, Hussain R, Mokhtar A. Feasibility, safety and clinical efficiency of optometric service pathways at primary and tertiary care level in Ampang, Malaysia. Asian Journal of Ophthalmology 2019 ; 16(3): 193-210.
17. Sajid MS, Hollingsworth T, McGlue M, Miles WF. Factors influencing the diagnostic accuracy and management in acute surgical patients. World journal of gastrointestinal surgery. 2014; 6(11): 229.
18. Liu S, Wang L, Zhang T, Liu C, Liang H, Zhang Y, Guo D. Factors affecting the work competency and stability of family doctors in Shanghai: a tracking study. BMC family practice 2019; 20: 1-1.
19. Shi C, Lee J, Wang G, Dou X, Yuan F, Zee B. Assessment of image quality on color fundus retinal images using the automatic retinal image analysis. Scientific Reports. 2022 Jun 21;12(1):10455.
20. Ngah NF. CME NIH 2023: DR. MATA system for the detection and intervention of diabetic retinopathy [Internet]. 2023 [cited 2023 May 25]. Available from: https://www.youtube.com/watch?v=pvhV_UBeBQA

ORIGINAL ARTICLE

Mental state classification based on electroencephalogram (EEG) using multiclass support vector machine

Santi Wulan Purnami[1], Saffanah Karimah[1], Shofi Andari[1], Diah P Wulandari[2], Yoyok Setyo Hadiwidodo[3], Wardah R. Islamiyah[4], Margarita M Maramis[5], Jasni Muhamad Zain[6]

[1]Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, [2]Department of Computer Enginering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, [3]Department of Ocean Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, [4]Department of Neurology, Universitas Airlangga, Surabaya, Indonesia, [5]Department of Psychiatry, Universitas Airlangga, Surabaya, Indonesia, [6]Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Universiti Teknologi MARA, Malaysia

ABSTRACT

Introduction: Mental state refers to a person's state of mind from various perspectives, including consciousness, intention, and functionalism. Mental states closely related to everyday life include the concentrating state, neutral state, and relaxation state. Concentration is vital for cognitive tasks, while relaxation is crucial for comfort. However, individuals with mental disorders or neurological conditions often struggle to achieve these states, requiring effective detection and intervention. One method for detecting mental states is by using brainwave signals obtained through electroencephalogram (EEG). EEG has been widely used in neuroscience and clinical settings to objectively assess mental states by analyzing brainwave signals. Previous studies have demonstrated the potential of EEG-based mental state classification in stress detection, cognitive workload analysis, or depression detection.

Materials and Methods: The data used in this research is secondary data in the form of recorded brainwave signals using EEG from 2018. and utilises self-reported data obtained from locally validated personal stress inventory questionnaires. The data used was obtained from four participants, including two females and two males. For pre-processing, this study uses the Hamming Windows Finite Impulse Response filtering method to extract features from each wave band. Additionally, feature selection methods are applied to choose the most relevant predictor features. Multiclass Support Vector Machine (SVM) with One-Against-One (OAO) and One-Against-All (OAA) approaches are used for classification.

Results: The feature selection process reduced the number of predictor variables from 160 to 40, focusing on minimum and maximum feature values. Multiclass SVM classification using 40 predictor variables achieved an AUC range of 0.907–0.922 (OAA) and 0.854–0.927 (OAO), while classification using all predictor variables yielded an AUC range of 0.898–0.927 (OAA) and 0.917–0.941 (OAO). Comparative performance analysis indicates that the OAA approach is superior to the OAO approach.

Conclusion: This study highlights the effectiveness of EEG-based classification of mental states using the Multiclass SVM method. The findings reinforce the role of EEG as an objective tool for mental state assessment, supporting its potential application in clinical and cognitive research for early detection of mental health disorders.

KEYWORDS:
Mental States, Electroencephalogram (EEG), Brainwaves, Multiclass SVM

INTRODUCTION

Mental state refers to the state of the mind that can be observed from various perspectives, such as consciousness-based, intentionality-based, and functionalism-based.[1] Mental states that are closely related to everyday life are concentrating states, relaxed states, and neutral states. Mental states, particularly concentration and relaxation, play an important role in everyday life. Concentration is particularly needed in cognitively demanding activities, such as studying and working. The inability to achieve a state of concentration at the right time can have a negative impact on information processing, memory retrieval and problem-solving decision-making. On the other hand, relaxation also plays a crucial role in maintaining psychophysiological balance. In a relaxed state, the body is calm, not tense, and recovers from fatigue, which is important for overall mental and physical health.

Along with increasing demands in education and work, the complexity of modern life has led to the emergence of various mental health problems. Disorders such as anxiety, depression, bipolar, post-traumatic stress disorder (PTSD), and insomnia can prevent a person from achieving an ideal state of concentration or relaxation.[2] For example, individuals with anxiety disorders often have difficulty relaxing, sleeping, and maintaining focus, which in turn impairs their productivity and well-being. In addition to mental disorders, neurological diseases such as Alzheimer's, stroke, vascular dementia and Parkinson's disease can also impact one's ability to concentrate or achieve a relaxed state.[3-5] The inability to manage both conditions is often an indication of a mental or neurological health disorder that needs to be addressed.

This article was accepted: 21 April 2025
Corresponding Author: Santi Wulan Purnami
Email: santi_wulan@its.ac.id

A person's condition, both neurologically and mentally, can influence the frequency and type of brain waves that appear, this is the basis for detecting mental conditions based on brain waves.[6] Electroencephalography (EEG) is a widely used tool to objectively analyze brain activity, as the frequency and type of brain waves that appear can reflect a person's mental state.[7,8] The diagnosis of EEG recordings is often done in the time domain by examining the waveforms, wave sharpness, and wave complexity9. Conventional methods in EEG analysis still rely on manual observation of brain wave patterns, which can be time-consuming and subjective. Therefore, faster and more accurate automated techniques are needed to evaluate a patient's mental state.

Machine learning methods have been developed to identify brain activity patterns related to a person's mental state. This study focuses on applying machine learning techniques to classify EEG signals into concentration, relaxation and neutral categories. Using this method, this study aims to support medical personnel in understanding patients' mental states more objectively, which in turn can improve the accuracy of diagnosis and the effectiveness of clinical interventions.

Furthermore, this EEG-based modeling of mental states can be a preliminary study for further research in understanding a person's psychological or clinical condition. The inability to achieve a situationally appropriate state of concentration, relaxation or neutrality could be an early indication of a more serious psychological or neurological disorder. With this modeling, the potential for developing early screening methods for mental health disorders becomes more open, which in turn can help in prevention efforts and the design of more targeted interventions.

In the organization of this review, section 2 the adopted research method is presented. The results of the study are shown in section 3. In section 4, a discussion and comparison between studied papers are provided. Section 5 at the end is devoted to the conclusion.

MATERIALS AND METHODS

Data Source

The data used in this research is secondary data in the form of recorded brainwave signals using Electroencephalography (EEG), which was obtained through GitHub using the following link https://github.com/jordan-bird/eeg-feature-generation/tree/master/dataset/original_data.[10] The data used in this study was obtained from four participants, consisting of two females and two males who performed the mental state recording process (relaxed, neutral, and concentrated) several times. The conditions imposed when recording are as follows: In the relaxed mental state, subjects were listened to low-tempo music and sound effects designed for meditation while being instructed to relax their muscles and rest. In the neutral mental state, a similar test was conducted but without any stimulus. This test was conducted before the other mental state tests to prevent the lasting effects of relaxation and concentration. As for the concentration mental state, subjects were instructed to run a "shell game" in which a ball was hidden under one of three cups which were then randomized, then the subject's task was to determine which cup had the ball in it.

Each participant made two recordings for each mental state resulting in a total of 24 data recordings. However, in this study only data with a minimum duration of 40 seconds was used so that 22 raw data were used and the recording duration was cut to 40 seconds. From these 22 data, the signal was cut into several signal pieces within a time span of 10 seconds and overlapping 5 seconds so that the data used in the classification analysis was 154 sample data. The recording data used in this study came from four channels: AF7, AF8, TP9, and TP10.[10]

Research Steps

The research steps (shown by Figure 1) are as follows:

1. Brain Activity Analysis: Examines brain wave patterns to understand the characteristics of EEG signals related to mental states.
2. EEG Data Processing: Filtering the signals based on relevant frequency ranges to extract meaningful information according to the type of wave subband as each wave subband is associated with a specific mental state.
3. Data Segmentation: Dividing the EEG signal into time segments to observe changes in brain activity within a certain period.
4. Feature Extraction: Retrieving characteristic features i.e. key information from the EEG signal that can reflect a person's mental state.
5. Feature Selection: Selecting the aspects of the signal characteristics that are most influential in distinguishing mental states.
6. Data Normalization: Scaling the data to make it more consistent in the classification process.
7. Mental State Classification: Analyzing brain activity patterns with machine learning models to classify mental states based on EEG data.
8. Model Evaluation: Comparing multiple analysis approaches to determine the best method for detecting mental states based on EEG.

Pre-processing data

The signal processing process begins with signal filtering. Signal filtering is used to separate signal waves based on frequency ranges, namely delta waves (0.1-3.5 Hz), Theta (4-7.5 Hz), Alpha (8-13 Hz), Beta (13.5-30 Hz), Gamma (30.5-100 Hz). The filtering method used is a Finite Impulse Response (FIR) Filter which has a limited impulse response because there is no feedback in the filter, giving results that tend to be stable compared to Infinite Impulse Response (IIR). Signal data tends to be stationary for a short duration, so the time epoch method will be more effective in reducing the variation of signal.[10] Time epoching is the segmentation of data into several epochs based on time intervals determined by the researcher. Each epoch segment can overlap with other segments, which is known as overlap. The common overlapping used in EEG analysis is 50% of the epoch size.[11]

Feature Extraction and Feature Selection

Features are unique characteristics of an object. Feature extraction is a step to obtain features that will be used in the

classification process as predictor variables. The following features can be extracted from EEG recording signals[10-16]:

1. Minimum: The smallest sample value.
2. Maximum: The largest sample value.
3. Variance: A measure of how spread out the data is.
4. Skewness: A measure of the asymmetrical of the data distribution.
5. Kurtosis: A measure of the peakedness of the data distribution.
6. Energy: A quantitative measure of electrical activity in the brain.
7. Entropy: A nonlinear measure that quantifies the complexity and randomness of the data.
8. Zero Crossing Band: The number of sign changes (from positive to negative and vice versa) in the signal data fluctuations.

In general, a variable is said to be good if it has a high correlation with its class but has a low correlation with other variables. In general, a variable is said to be good if it has a high correlation with class variables but has a low correlation with other variables. Therefore, the Fast Correlation Based Filter (FCBF) feature selection method is used[17] to select several features that have a high correlation with class variables.

Multiclass SVM

Support Vector Machine (SVM) is a machine learning method based on the principle of Structural Risk Minimization and can be used for solving classification problems in high-dimensional feature spaces.[18] In nonlinear cases, a kernel function is used to map the input space to a higher-dimensional feature space. Parameter selection is a step to obtaining the most optimal hyperparameters for the SVM model, which results in the best classification performance. One of the methods that provide satisfactory results is Grid Search[19], The determination of parameters C and γ can be done using an exponential pattern.[20] The method for determining parameter is based on research conducted by Liu, Zhang, Qu, and Bell.[21] Multiclass SVM is an extension of the SVM method, which was originally designed for binary data classification. Multiclass SVM is applied by decomposing multiclass data into a series of binary classification problems, making it possible to apply standard SVM. There are two approaches to Multiclass SVM namely One-Against-All and One-Against-One. In the OAA approach, each class p is compared with all other classes22. To classify data into e classes/categories, e binary SVM models need to be built. However, In the OAO approach, each class is compared with every other class.[13] To classify data into e classes/categories, e(e-1)/2 binary SVM models need to be built.

Classification Performance Evaluation

Classification performance can be evaluated using a confusion matrix. The confusion matrix is a table that presents the results of classification for measuring the classification performance. Measures that can be used to evaluate classification performance include accuracy, sensitivity, and specificity. Besides using the confusion matrix, classification performance can also use the area under curve (AUC). AUC is a method that measures the area under the Receiver Operating Curve (ROC) graph to assess prediction performance.

RESULTS

The total number of recordings is 22, consisting of seven Concentrated mental conditions, seven neutral mental conditions, and eight relaxed mental conditions. Power Spectral Density (PSD) plots display the spectral power density of brainwave signals and depict the frequency energy distribution in brainwave signals. An example of one of the PSD plots is shown in Figure 2. The PSD plots containing signal frequencies from 0 to 120 Hz indicate that brainwave signals for Concentrated, neutral, and relaxed conditions have delta, theta, alpha, beta, and gamma wave sub bands. Therefore, the preprocessing stage will start with filtering, where the data will be separated into five wave sub bands.

Filtering

The Finite Impulse Response (FIR) Hamming Windows filtering method is used to separate the signal into five wave sub bands. Each mental condition contains five sub bands: delta, theta, alpha, beta, and gamma. The higher the frequency range, the higher the signal density. Therefore, delta will have the lowest signal density, while gamma will have the highest signal density.

Segmentation

Brainwave signals tend to be stationary over short periods. Therefore, time epoching or data segmentation based on time duration is performed to reduce signal waveform variation. The data is divided into ten-second epochs with a five-second overlap, resulting in seven epochs for each recording data.

Signal Feature Extraction

After data segmentation, the next step is feature extraction to capture the characteristics of the signal data. Feature extraction is performed on the five wave sub bands, with each sub band having eight features. Therefore, for one channel, there will be a total of 40 features. This study uses four channels, resulting in a total of 160 features entering the classification stage. The feature extraction results show that there are striking differences in the range of signal features. Therefore, Z-Score Normalization is used to balance the impact of variables in the data modeling process.

Feature Selection

Feature selection using Fast Correlation Based Filter (FCBF) results in 40 selected features. These features include the minimum and maximum features from the delta, theta, alpha, beta, and gamma waves in channels AF7, AF8, TP9, and TP10. These 40 selected features are used as predictor variables in the classification analysis.

Classification Mental State Using the OAA Multiclass SVM

In the One-Against-All (OAA) approach, the class p data is compared with all other data except class p. As a result, three binary models are formed. The Multiclass SVM modeling is performed with optimal parameters obtained from Grid Search. The best model is selected based on the highest Area Under the Curve (AUC) value. The evaluation of the OAA approach with all variables is shown in Table I. Table I shows

Table I: Model Evaluation of All Variables

Kernel	AUC	Accuracy	Sensitivity	Specificity
One Against All Approach Model				
Linear	0.927	0.902	0.901	0.952
Polynomial d=2	0.922	0.896	0.895	0.949
Polynomial d=3	0.926	0.902	0.900	0.952
Polynomial d=4	0.912	0.883	0.882	0.942
RBF	0.907	0.876	0.875	0.939
Sigmoid	0.898	0.864	0.862	0.933
One Against One Approach Model				
Linear	0.941	0.922	0.920	0.962
Polynomial d=2	0.927	0.902	0.901	0.952
Polynomial d=3	0.922	0.896	0.895	0.949
Polynomial d=4	0.931	0.909	0.907	0.955
RBF	0.927	0.902	0.902	0.952
Sigmoid	0.917	0.889	0.889	0.946

Table II: Prediction results using the best model

Recording Data Number	Actual Class	Prediction Class
5	Concentrated	Concentrated
	Concentrated	Concentrated
	Concentrated	Concentrated
	Concentrated	Concentrated
	Concentrated	Concentrated
	Concentrated	Concentrated
	Concentrated	Concentrated
12	Neutral	Concentrated
	Neutral	Concentrated
	Neutral	Neutral
	Neutral	Neutral
	Neutral	Neutral
	Neutral	Neutral
	Neutral	Neutral
19	Relaxed	Relaxed
	Relaxed	Relaxed
	Relaxed	Relaxed
	Relaxed	Relaxed
	Relaxed	Neutral
	Relaxed	Relaxed
	Relaxed	Relaxed

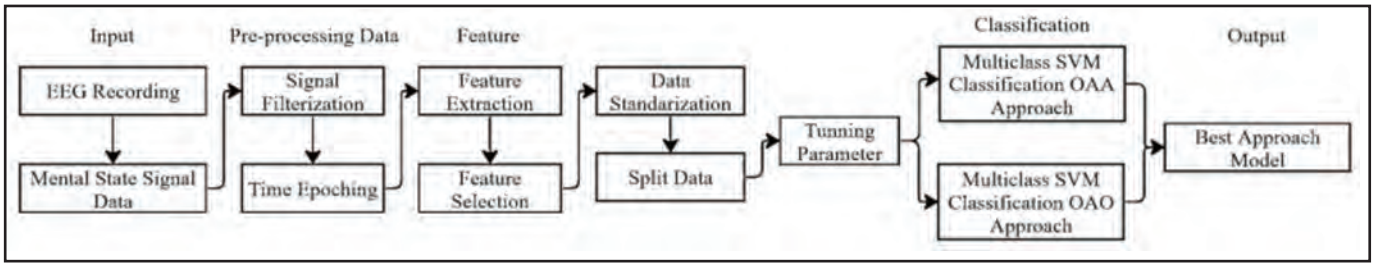


Fig. 1: Research Step

that the linear kernel has the highest AUC compared to other kernels. Therefore, the best model in the OAA approach with all predictor variables is the model with the linear kernel with a parameter C of 0.0625. The AUC value of 0.927 indicates that the classification performance falls into the Excellent Classification category. The accuracy value shows that the model can correctly classify 90.2% of the data into their respective classes.

The OAA approach with predictor variables from feature selection has the best kernel performance, which is Polynomial degree 4 with an AUC of 0.922, indicating that the classification performance is in the Very Good Classification category. The accuracy value of 0.896 indicates that the model can correctly classify 89.6% of the test data using the best kernel and parameters.

The best OAA model is obtained by comparing the AUC between the OAA model with all variables and the OAA

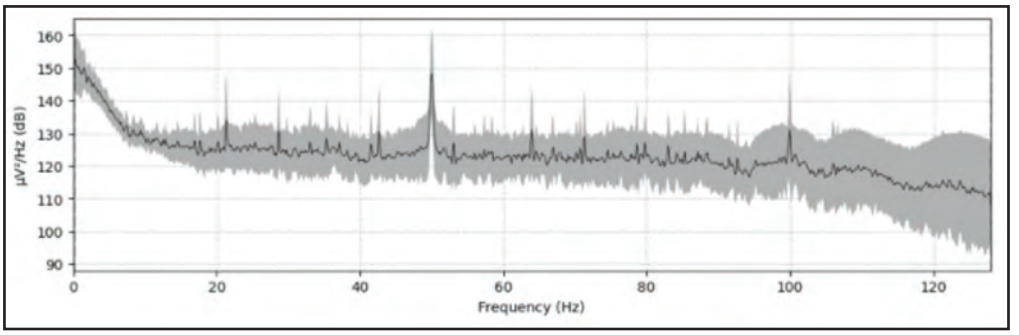


Fig. 2: PSD Plot of Signal Data for Concentrated Mental Condition

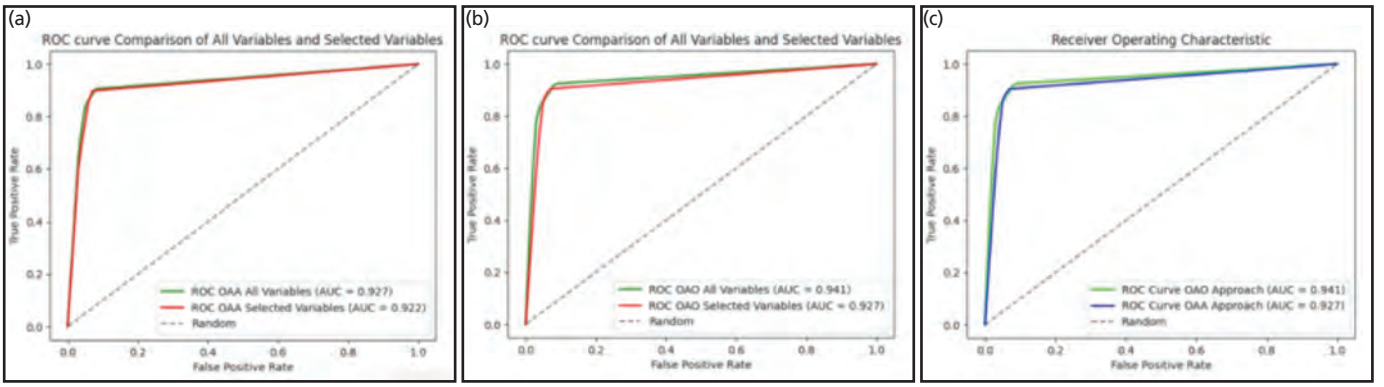


Fig. 3: Comparison ROC curve (a) OAA Model with All Variables and Selected Variables, (b) OAO Model with All Variables and Selected Variables, (c) OAA Approach and OAO Approach

model with selected variables. The OAA model for all variables has an AUC of 0.927, higher than the kernel AUC value of the OAA model for selected variables, which is 0.922. This shows that the model with all variables is better than the model with selected variables. A comparison of the ROC curve between all variables and selected variables in the OAA model is shown in Figure 3 (a). The difference in AUC values of 0.005 means that visually there is no significant difference between the two graphs.

Classification Mental State Using the OAO SVM

In the One-Against-One (OAO) approach, one class is compared with every other class. Like the OAA approach, modeling was performed on linear, polynomial, RBF, and sigmoid kernels using the best parameters according to Grid Search. The model evaluation results of the OAO approach using all feature predictor variables are shown in Table I.

The best kernel is the Linear kernel with the C parameter of 0.031256. The AUC value of 0.941 also implies that the classification ability is in the Excellent Classification category. The accuracy value on the optimal kernel of 0.922 indicates that the model can classify 92.2% of the testing data correctly.

The OAO approach with selected predictor variables has the best kernel performance, which is polynomial degree 3 with an AUC value of 0.922 shows that the classification level reaches the Excellent Classification category. The accuracy value of 0.902 shows that the model can classify 90.2% of the testing data correctly.

The highest AUC in the model with All Variables was 0.941 but the highest AUC in the model with Selected Variables was 0.927. So, it can be concluded that in the OAO approach, the best model is the model with all predictor variables. The ROC curve showing the comparison between the ROC of all variables and the selected variables is shown in Figure 3(b). Visually, the ROC of all features is higher than the ROC of the selected features, this is in accordance with the computational results which state that the area under the curve for all variables is larger than the area under the curve of the selected variables.

Selection of the Best Classifier

The selection of the best classifier between the OAA approach and the OAO approach in the Multiclass SVM method is done by comparing the performance of the best model of each approach. The comparison of the two models is shown in Table I. Table I shows that the AUC, accuracy, sensitivity, and specificity of the best model multiclass SVM OAO approach are higher than the best model multiclass SVM OAA approach. The comparison of the ROC curves between the two approaches is shown in Figure 3(c). Figure 3(c) shows that the area under the OAO curve is larger than the area under the OAA curve. This is in accordance with the results of mathematical calculations.

The higher AUC value of the OAO approach indicates that the Multiclass SVM OAO approach is better at classifying mental state data from electroencephalogram than the Multiclass SVM OAA approach

An example of predicting subjects using the best model is carried out on the 5th recording data for the Concentrated class, the 12th recording data for the neutral class, and the 19th recording data for the relax class. Each recording data has 7 epochs and 160 extracted features.

The comparison between the actual class and the predicted class using the best model on the 5th recording data, 12th recording data, and the 19th is shown in Table II. Table II shows that there are three predicted incorrect data, namely two neutral class data but predicted as a concentrated class, and one relaxed class data but predicted as a neutral class.

DISCUSSION

The series of signal processing stages in this study demonstrates an improved accuracy compared to previous research using the same dataset10. The study by Bird, Manso, Ribeiro, Ekart, & Faria[10], which did not specify the employed filtering method, achieved the highest accuracy of 87.16% using the random forest method, whereas the accuracy of the SVM method was 75.24%. In this study, it achieved the highest accuracy value of 92.2%. This improvement confirms that the series of EEG waveform signal processing steps and the multiclass SVM approach used in this study can improve accuracy compared to processing from previous research with the same dataset.

The high classification performance indicates that EEG signals can effectively represent a person's mental condition. An accuracy of 92.2% implies that EEG-based models can distinguish between concentrated, neutral, and relaxed mental states with high reliability. Therefore, EEG analysis serves not only as a recording of brain activity but also as a practical tool for interpreting psychological states objectively. This capability opens opportunities for EEG-assisted applications in clinical environments, especially in supporting medical personnel to monitor, assess, and tailor treatments based on a patient's cognitive state.

Furthermore, the ability of the model to detect discrepancies between a patient's expected and actual mental state highlights its potential clinical relevance. For instance, if a subject is expected to concentrate but the model detects a relaxed or neutral state, it may suggest signs of cognitive decline, mental fatigue, or stress. Similarly, persistent tension detected during a supposed relaxation phase may indicate underlying conditions such as anxiety, insomnia, or chronic stress. These findings underscore the potential of EEG-based classification models as early screening tools for psychological or neurological disorders.

This study, therefore, lays the groundwork for future research aimed at integrating EEG-based monitoring into mental health diagnostics and personalized interventions, ultimately contributing to improved psychological well-being.

Despite the favorable results, several limitations must be acknowledged. First, this study employed publicly available data, thus future research is advised to utilize real data from EEG recordings in hospital settings to assess the generalizability of this method across different cases. Secondly, we solely utilized data sourced from four EEG channels—AF7, AF8, TP9, and TP10—while there remains a possibility that other EEG channels might be more optimal for classifying mental states such as concentration, neutrality, and relaxation. The signal filtering method in this study only focuses on segmenting the signal waveform into five sub bands, while there is no handling of noise generated when the recording is in progress such as eye blinks, muscle movements, or noise from the recording machine. Future studies may explore various filtering combinations to produce cleaner signals.

Furthermore, feature selection in this study was focused on the level of correlation between predictor variables and class variables without overcoming multicollinearity. Meanwhile, signal data is susceptible to multicollinearity. Therefore, it is suggested that further research apply feature extraction methods such as Principal Component Analysis (PCA) and factor analysis which are more effective in managing multicollinearity. In addition, in this study we use the Grid Search method to select the hyperparameters of the SVM model. However, this method has limitations, including requiring a long computation time, so future research needs to consider the parameter tuning method which has a shorter computation time.

CONCLUSION

This paper presented a study on mental state classification based on EEG signals. The signals go through the data pre-processing stages before finally several features from each wave sub band are extracted to be able to classify mental states of Concentrated, neutral, and relaxed. Through the multiclass SVM approach, the highest AUC value is 0.941 with an accuracy of 92.2% which indicates that the accuracy of the model in classifying data is an excellent classification. These good results indicate that signal processing using computation can support the medical team in classifying the patient's mental condition based on the electroencephalogram accurately, so that it can support medical personnel in clinical monitoring and decision-making.

An important implication of this study is its ability to detect mismatches between the expected and actual state of the patient. If a person is expected to be in a state of concentration but the model shows otherwise, this may indicate cognitive impairment, stress, or mental fatigue that requires further evaluation. Conversely, if the patient is supposed to be in a relaxed state but the model shows that he or she is still in a tense state or has difficulty achieving relaxation, this could be an indication of an anxiety disorder, insomnia, or chronic stress that requires further intervention.

As such, this method can serve as a preliminary study for further research in detecting indications of psychological or clinical disorders, as well as assist in the development of more targeted interventions to improve the quality of patients' mental health and well-being.

ACKNOWLEDGMENTS

The research was funded by Directorate of Research and Community Service, Institut Teknologi Sepuluh Nopember

(ITS), Ministry of Education and Culture, Research and Technology, Indonesia. In accordance with the ITS Fund Partnership Research Implementation Agreement Batch 2 Year 2023. Master Contract Number: 2247/PKS/ITS/2023, on 26 June 2023. Researcher Contract Number: 2252/PKS/ITS/2023, on 26 June 2023.

DECLARATION OF INTEREST

The authors would like to disclose that they have no conflict of interests to declare and have no competing interests in this study.

REFERENCES

1. S Baron-Cohen, H Ring, J Moriarty, B Schmitz, D Costa, P Ell. Recognition of mental state terms. Br J Psychiatry 1994; 165: 640-9.
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (Vol. 5). American Psychiatric Publishing, Arlington (2013).
3. IN Khikmatin, DR Desiningrum. Sebuah Interpretative Phenomenological Analysis Tentang Pengalaman Family Caregiver dalam Merawat Pasien Alzheimer. Empati 2018; 1-14.
4. SH Pratiwi, EA Sari. Faktor Resiko Stroke pada Masyarakat Desa Pangandaran. Jurnal Pengabdian Kepada Masyarakat, 412-5 (2017).
5. AN Kumalasari, F Rahmayani, S Hamidi. Diagnosis dan Pencegahan Perburukan Demensia Vaskular pada Pasien Pasca Stroke. Medula, 25 (2018).
6. N Boutros, S Galderisi, O Silvana, S Riggio. Standard Electroencephalography in Clinical Psychiatry. John Wiley & Sons, Ltd, UK (2011).
7. H Bhavsar, A Ganatra. A Comparative Study of Training Algorithms for Supervised Machine Learning. International Journal of Soft Computing and Engineering 2012; 2: 74-81.
8. M Mazher, I Faye, A Qayyum, AS Malik, Classification of Resting and Cognitive States using EEG-based Feature Extraction and Connectivity Approach, in 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Sarawak, Malaysia (2018)
9. R Djemili, H Bourouba, MCA Korba. Biocybernetics and Biomedical Engineering 2016; 36: 285-91.
10. JJ Bird, LJ Manso, EP Ribeiro, A Ekart, DR Faria, A Study on Mental State Classification using EEG-based Brain-Machine Interface, In IEEE 2018 International Conference on Intelligent Systems (IS), 25-27 September 2018, Funchal, Portugal (2018)
11. KD Tzimourta, LG Astrakas, AM Gianni, AT Tzallas, N Giannakeas, I Paliokas, DG Tsalikakis, MG Tsipouras. Evaluation of Window Size in Classification of Epileptic Short-Term EEG Signals Using Epileptic Short-Term EEG Signals Using a Brain Computer Interface software. Engineering, Technology & Applied Science Research, 8, 3093-7 (2018)
12. NH Saputra, AD Wibawa, MH Purnomo, . Pamungkas, EEG-based Statistical Analysis on Determining the Stress Mental State on Police Personnel, In IEEE I2022 1st International Conference on Information System & Information Technology (ICISIT), Yogyakarta, Indonesia (2022)
13. Avasiloaie, BA Zîrnă, M Oniga, AE Sultana, MC Frunzete, Classification of Mental Effort vs. Relaxation States Using EEG Signals in IEEE The 10th IEEE International Conference on E-Health and Bioengineering, Iasi, Romania (2022)
14. MS Dehnavi, VS Dehnavi, M Shafiee. Classification of Mental States of Human Concentration Based on EEG Signal in IEEE International Conference on Information and Knowledge Technology (IKT), Babol, Iran (2020)
15. M Ali, Abdulghani AM. Development of Efficient Algorithm to Detect and Classify Alpha Activity in Real-time EEG Signal using LabVIEW for Mental State Detection in 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada (2017)
16. M Alirezaei, SH Sardouie. Detection of Human Attention Using EEG Signals, In IEEE 2017 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME), 9 August 2017, Tehran, Iran (2017)
17. L Yu, H Liu. Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution in Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003), 21-24 August 2003, Washington DC, USA (2003)
18. Abe S. Support Vector Machines for Pattern Classification (Springer, New York, 2010)
19. Z Chen, N Tang, X Xia. A Parallel Genetic Algorithm Based Feature Selection and Parameter Optimization for Support Vector Machine. Scientific Programming 2016; 1-10.
20. C.-W. Hsu, C.-C. Chang, C.-J. Lin, A Practical Guide to Support Vector Classification (National Taiwan University, Taipei, 2016)
21. L Liu, Z Zhang, Z Qu, A Bell, Utilizing Bayesian Optimization of Stacking. Electronics 2023,; 12: 1-18.
22. X Yang, Q Yu, L He, T Guo. The one-against-all partition based binary tree support vector machine algorithms for multi-class classification. Neurocomputing 2013; 113: 1-7.

Challenges in setting up the first cyto-reductive surgery (CRS) and hyperthermic intra-peritoneal chemotherapy (HIPEC) service in Malaysia

Mohamed Rezal Abdul Aziz, MSurg, Nora Abdul Aziz, MSurg, Ruben Gregory Xavier, MSurg, Tak-Loon Khong, FRCS, April Camilla Roslani, FRCS

Colorectal Unit, Department of Surgery, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

ABSTRACT

Introduction: Cyto-Reductive Surgery (CRS) with Hyperthermic Intra-Peritoneal Chemotherapy (HIPeC) improves survival in selected patients with peritoneal surface malignancies (PSM) of various tumour entities. This treatment was not available in the Malaysian public health sector prior to 2018, due to lack of expertise, funding and accessibility. We report our challenges in establishing the first such service.

Materials and Methods: Patients referred for CRS-HIPeC between February 2018 and November 2023 were discussed in a multidisciplinary team meeting. Prospectively collected data, including patient demographics, extent of disease (radiological stage, peritoneal carcinomatosis index - PCI), pre-operative workup, completeness of cytoreduction (CCR) score and surgical outcomes, were analysed.

Results: Of 162 patients referred, 47 (30.0%) underwent CRS-HIPeC. The median age was 59 years (range 20-76 years). Median PCI was 11 (range 1-39). CCR distribution was as follows: CCR 0 - 40 (85.1%), CCR 1 - 2 (4.3%), CCR 2 - 2 (4.3%) and CCR 3 – 3 (6.3%). Median operative time was 645 minutes (range 360 –1575 minutes) with a median length of in-hospital stay of 11 days (range 6–146 days). All patients were initially managed in the intensive care unit. Sixteen (34.0%) patients developed complications of Clavien-Dindo Class 3 and above, with three operative mortalities (6.3%).

Conclusion: CRS-HIPeC requires adequate clinical expertise, facilities and volume. Its labour and resource-intensive nature mean that centralization of services is necessary for sustainability. Further evaluation of its cost-benefit in our setting will be required.

KEYWORDS:
Cyto-reductive Surgery, Hospital service, Hyperthermic Intra-Peritoneal Chemotherapy, Malaysia, Peritoneal Surface Malignancy

INTRODUCTION

Peritoneal surface malignancies (PSM) are rare, and difficult to diagnose. Primary subtypes are primary peritoneal carcinoma and malignant mesothelioma. Secondary subtypes are metastases from other primaries, such as gynaecologic, urologic and gastrointestinal cancers. Overall, PSM is associated with poor prognosis, as evident from the EVOCAPE 1 study.[1]

However, there is mounting evidence that Cytoreductive Surgery with Hyperthermic Intra-Peritoneal Chemotherapy (CRS-HIPeC) significantly improves overall survival, particularly for those with secondary PSM. In this technique, the primary disease and affected organs are resected together with the peritoneum followed by heated intraperitoneal chemotherapy.

Malaysia is a middle-income country with a population of 32.7 million.[2] Public healthcare in Malaysia is subsidized by local taxes and government general revenue, whereas private healthcare is funded by medical insurers and/or out-of-pocket expenditure. Diseases that carry a large socio-economic impact, such as cancers and cardiovascular diseases, are understandably prioritized to receive public funding. Unfortunately, lack of awareness often results in cancers presenting at an advanced stage.

As a public teaching hospital and tertiary referral centre, University Malaya Medical Centre (UMMC) receives significant numbers of patients with PSM. Until recently, they would primarily be referred to oncologists for palliative chemotherapy. Nevertheless, it was apparent that a subset of these patients could benefit from CRS-HIPeC. CRS-HIPeC is complex, and requires multi-disciplinary involvement, comprising specialized surgeons, oncologists, anaesthetists, intensivists, and radiologists.[3–5] Consensus on patient selection is of paramount importance, and establishing a service is associated with a significant learning curve, particularly with limited resources.[6,7] We report our initial experience with CRS-HIPeC, including the challenges and obstacles.

MATERIALS AND METHODS

PHASE 1: ESTABLISHMENT OF SERVICE

We engaged with established centres and reviewed existing guidelines for the development of our CRS-HIPeC services. Hospital administration was involved at an early stage of planning for logistics and funding. Multidisciplinary team members were identified and underwent training. Subsequently, we created our protocol, including patient

This article was accepted: 14 April 2025
Corresponding Author: April Camilla Roslani
Email: april@ummc.edu.my; aprilroslani@um.edu.my

selection criteria. Following establishment of protocols, we were proctored by an established regional centre, in both patient selection and conduct of the procedure.

PHASE 2: EVALUATION OF OUTCOMES

Data on all patients was prospectively collected. Variables analysed included demographics, operative time, disease burden, completeness of cytoreduction (CCR), primary pathology, morbidity, mortality and length of stay. Frequency was evaluated using median (range). Number of surgeons and types of subspeciality were also included. Performance status was classified according to the Eastern Cooperative Oncology Group (ECOG). The disease burden was defined by the Peritoneal Cancer Index (PCI).8 The CCR score was determined at the completion of cytoreduction and prior to HIPeC.[9,10] Cytoreduction was performed as described by P. Bao and O. Glehen et al.[9,11] Following CRS, HIPeC was performed via a closed technique using the Hyperthemia Pump™ (Belmont Medical Technologies, Billerica, Massachusetts, USA) to infuse cytotoxic chemotherapeutic drugs at 42 degree Celsius for 60 minutes. Anastomosis, stoma and drain placement were performed after copious washout upon the completion of HIPEC. Patients were monitored in the intensive care unit post-operatively and transferred to the surgical ward when appropriate. Post-operative morbidity events were graded using the Clavien-Dindo Classification.[12] Follow-up was performed at 3 weeks, 6 weeks, 3 months and 6 months, at which time physical examination, tumour markers (CEA, CA 125 and CA 19-9) and computed tomography (CT) scan were performed.

RESULTS

PHASE 1

The pioneer members included surgeons, anaesthetists, oncologists, and operating theatre nursing staff. The proposal for establishing the service was approved by the hospital administration. Pioneer members underwent training in an established regional centre (National Cancer Centre, Singapore). Funding was through a novel public-private partnership. The CRS-HIPEC team was established by a colorectal surgeon with special interest in the management of PSM. Following careful case selection through a multi-disciplinary team discussion, the first case was performed on 8 February 2018. Additional colorectal surgeons joined the team from 2019 onwards, with one assigned to lead the management of peritoneal malignancies from December 2022 onwards (Fig. 1). The CRS-HIPeC team expanded to 13 people within a year, including supporting sub-specialities such as gynaecologists, hepatobiliary surgeons, urologists and anaesthetists. This allowed a doubling of case volume from the first to the second year (Fig. 2).

A clinical database was constructed and maintained prospectively. Agreed selection criteria included every patient with incidental or symptomatic peritoneal disease seen on imaging or intra-operatively. Referred patients were initially discussed in our institutional tumour board. Shortlisted patients were then further discussed in an international inter-institutional multidisciplinary tumour board, which allowed for proctoring on patient selection. Fifteen cases were discussed in the international MDT board, over a period of six months, which allowed the UMMC team to reach a comfort level on independent decision-making. On-site proctoring by an experienced visiting surgeon was provided for the conduct of the operations in our initial cases.

A Temporary Practicing Certificate (TPC) was obtained for the proctor. This required submission of numerous documents to the Malaysian Medical Council, and took several months to complete. In addition, a trained perfusionist was needed to operate the HIPeC infusor. The infusor required prior compatibility and safety assessment by our institutional Biomedical Engineering Department. Occupational, Safety, Health and Environmental (OSHE) standards were followed to safeguard patients, staff and the environment, in handling chemotherapy peri-operatively.

Patients with good Eastern Cooperative Oncology Group (ECOG) performance status (ECOG 0 or 1) and potentially resectable disease, based on pre-operative imaging and diagnostic laparoscopy, were considered for CRS-HIPeC. In some patients, the performance status was assessed objectively with cardiopulmonary exercise testing (CPET) to provide a thorough integrative assessment of multi-organ physiological function to exercise.[13]

PHASE 2

A total of 162 patients with PSM were assessed from February 2018 to November 2023 for consideration of CRS-HIPEC. Of these, 17.5% were from other centres around Peninsular and East Malaysia, while two patients (1.2%) were referred from Indonesia. Eighty percent were referrals from oncologists, gynaecologists or other surgical units within our centre. Thirty-one percent of patients were receiving systemic chemotherapy at the time of referral, and were referred due to disease progression. Fifty (30.9%) were PSM with colorectal primaries. Following assessment, fifty-one patients (31.5%) underwent surgery but three were found unresectable. An additional patient developed intraoperative pneumothorax preventing completion of the surgery.

Demographics of patients who had CRS-HIPeC are summarized in Table I.

The median age was 59 years, and the majority were female (68%). The Chinese were the predominant ethnic group (68.1%), in keeping with the national ethnic distribution of overall cancer incidence as reported in the Malaysian National Cancer Report 2016. Eighty percent were ECOG status 0. Of 47 patients who successfully completed CRS-HIPeC (Table II), 28 (59.6%) patients had multi-visceral resection. The median PCI score was 11 (range 1 - 39). The median operative time was 645 minutes with a range of 360 – 1575 minutes. Mean total blood loss was 2 litres with a range of 0.2 – 4.0 L . The majority of patients had two organs resected (42.6%).

All patients with colorectal primaries and low grade mucinous neoplasia (LAMN) received mitomycin for HIPeC component. During the initial set-up, the majority of the ovarian-PM (12.6%) received cisplatin and doxorubin for the HIPeC. The sarcomatosis peritonei patient received doxorubicin.

Table I: Patient demographic and clinical characteristics (n=47)

Variable	Results
Age [years; median (range)]	59 (20-76)
Body Mass Index [kg/m^2; median (range)]	23.9 (15.7-33.7)
Race	
Malay	12 (25.5%)
Chinese	32 (68.1%)
Indian	2 (4.3%)
Other	1 (2.1%)
Gender	
Male	15 (31.9%)
Female	32 (68.1%)
ECOG Status	
0	38 (80.9%)
1	8 (17.0%)
2	1 (2.1%)
Primary Pathology	
Appendix carcinoma	5 (10.6%)
Ovary	11 (23.4%)
Colon	13 (27.7%)
Primary peritoneal cancer	3 (6.4%)
Sarcoma	1 (2.4%)
LAMN	14 (29.8%)
Pre-operative tumour markers [U/ml; median (range)]	
CEA	6.3 (0.3-202)
CA 19-9	26.0 (1-2600)
CA-125	55.5 (4-9939)

Values are presented as n (%) unless otherwise stated.

Table II: Operative characteristics

Variables	Results
Operative time [minutes; median (range)]	645 (360-1575)
PCI score [median (range)]	11 (1-39)
Cytoreductive score (CC)	
CC – 0	n=40(85.1%)
CC – 1	n=2 (4.3%)
CC – 2	n=2 (4.3%)
CC – 3	n=3 (6.3%)
Chemo drugs	
Mitomycin	34 (72.4%)
Cisplatin and doxorubicin	6 (12.8%)
Cisplatin and paclitaxel	1 (2.1%)
Cisplatin	5 (10.6%)
Doxorubicin	1 (2.1%)
Estimated blood loss [L; median (range)]	2.0 (0.2 – 4.0)
No organ of resection	
1	19 (40.4%)
2	20 (42.6%)
3	8 (17.0%)

Values are presented as n (%) unless otherwise stated.

Morbidity requiring intervention (Grade III – IV) was 31.9%, while overall operative mortality was 6.3%. There was one early operative mortality, with death occurring on the fifth post-operative day, due to neutropenic sepsis where cisplatin was used in the HIPeC regimen. Another two patients died at four and five months post-operatively, within the index admission, due to complications of pneumonia and pulmonary embolism respectively (Table III).

Median follow up was 33 months. At last follow up, 17 patients (36.2%) had no evidence of disease recurrence or progression. Of the 42 patients who had CC0 or CC1 clearance, 13 (31%) had local recurrence. Six of these recurred within six months, while the others recurred after more than a year. There were seven (16.7%) distant recurrences in these 42 patients, five occurring within a year (Table III). The five patients with CC2 or CC3 clearance all progressed within a year.

DISCUSSION

CRS-HIPeC has been in practice for many decades in some countries.[14] Peritoneal surface malignancy generally has a poor prognosis, but selected patients do benefit from this

Table III: Post-operative outcomes

Variables	Results
Hospital stay [days; median (range)]	11 (6-146)
ICU stay [days; median (range)]	2 (1-29)
30-day morbidity	
Clavien-Dindo	
I	13 (27.7%)
II	18 (38.3%)
IIIa	10 (21.3%)
IIIb	3 (6.4%)
IVa	1 (2.1%)
IVb	1 (2.1%)
V	1 (2.1%)
Operative Mortality	3(6.3%)
Oncologic outcomes	
1-year disease free survival	27 (57.4%)
1-year local recurrence	6 (14.3%)
1-year distant recurrence	5 (11.9%)
1-year overall survival	43 (91.5%)

Values are presented as n (%) unless otherwise stated.

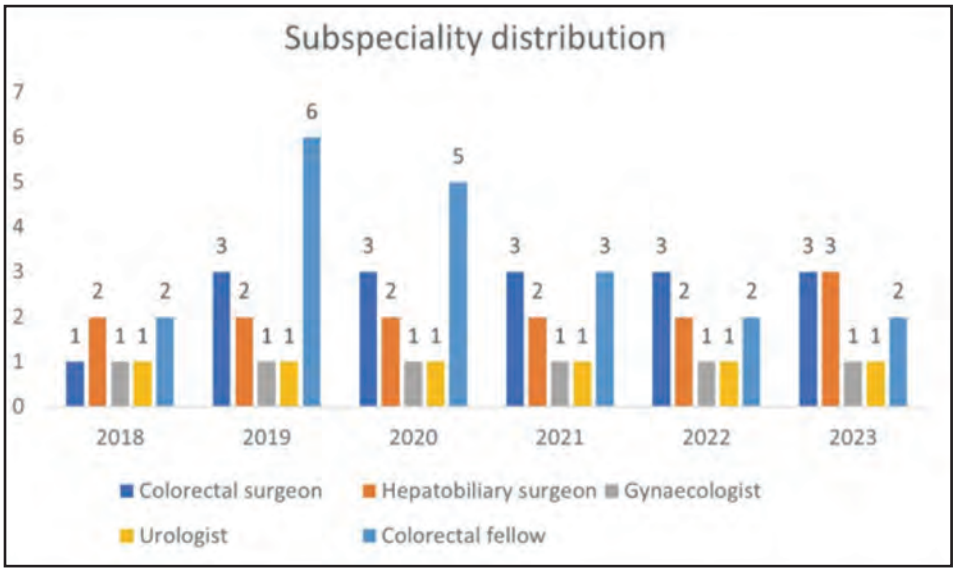


Fig. 1: Multidisciplinary team composition

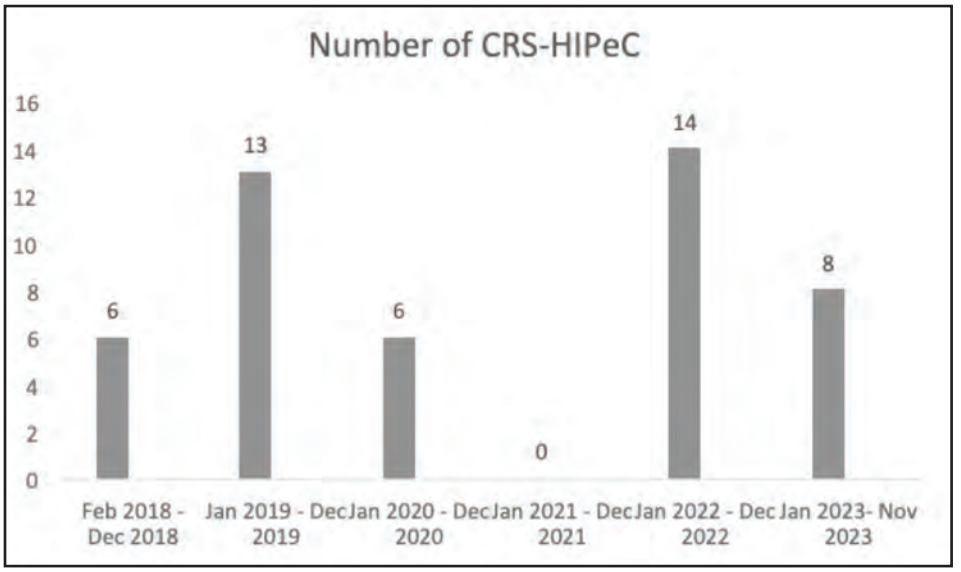


Fig. 2: Annual cases of CRS-HIPeC

treatment modality even though it is time consuming, associated with significant morbidity and mortality, and is an expensive operation to conduct.[15,16]

Tumour biology impacts outcomes, and influences selection. Simkins GA et al reported that the median survival for colorectal cancer peritoneal metastasis (CRC-PM) is 36 months, with one year mortality rate and recurrence rate post-CRS-HIPeC procedure of 13% and 35% respectively.[17] Nevertheless, while there are recommended PCI ceilings for CRC and gastric carcinomas (15 and 9 respectively), there are no stipulated PCI ceilings for primary PSMs or sarcomas, as there is a survival benefit for the latter two even with very high PCIs.

Oncologic clearance is the primary goal, and affects survival. Adherence to adjuvant therapies is also critical for optimal survival. Following discharge, three of our patients (6.4%) died due to disease progression post CRS-HIPeC. Our 13th patient had a dedifferentiated liposarcoma arising from the retroperitoneum, with gastrointestinal symptoms from extrinsic compression; his PCI was 21 and we achieved CC-1 clearance. His symptoms improved significantly, but he recurred seven months after surgery. He commenced palliative chemotherapy with overall survival of 15 months. Another patient had appendicular mucinous adenocarcinoma with PCI of 39, CC-2 clearance, was symptom free for nine months. He developed obstructive symptoms and died 18 months after surgery. A third patient had sigmoid carcinoma recurrence which was KRAS mutated. Her PCI was 10 and we achieved CC-0, all good prognostic features, but she refused adjuvant chemotherapy. She was disease free for six months and died from liver and lung metastasis at 15 months post-operatively.

While CRS-HIPeC can improve survival, morbidity is significant. The National Cancer Centre Singapore analysis of morbidity post-CRS-HIPeC over a 10-year period showed that for every additional resection performed, there was a 53% increase in the odds of experiencing post-operative complications.[18] The number of resections performed and intraoperative blood loss is directly proportion to the morbidity.[18]

Our outcomes also reflect the increase in morbidity with more extensive resections. Sixteen of our patients (34.0%) had severe complications (Clavien-Dindo III and above). One patient developed abdominal compartment syndrome that required an emergency laparostomy. Two developed pancreatic fistula post-distal pancreatectomy for appendicular carcinoma with peritoneal metastasis and sarcomatosis peritonei. These two patients had three organs resected with blood loss of 500 ml and 1500 ml respectively. It is clear that as centres become more experienced and attempt more complex cases, managing the associated morbidities will add to the overall cost of CRS-HIPeC.

Malaysian healthcare has a complex funding framework. While much of public healthcare is subsidized by the government, specialized equipment and pharmaceuticals often require out-of-pocket funding. In addition, only 14.2% of the population has personal medical insurance coverage.[19]

There are also differences in public hospitals within the Ministry of Health compared with teaching hospitals under the Ministry of Higher Education. In the latter, civil servants undergoing CRS-HIPeC would need to pay out-of-pocket for consumables, amounting to, on average, RM9888.00 (USD2088.35). Non-government staff would, in addition, have to pay the costs of medication and hospital stay. Three of our patients were unable to proceed with the operation due to financial constraint. One of them struggled financially to travel to the Peninsular of Malaysia for pre-operative assessment. While patients may avail of social welfare services, funding is limited. Given that the average household income in Malaysia is RM5,228 (USD1173.25) per month, public hospitals, which have a higher proportion of low income patients, are often unable to fully subsidize expensive treatments.[20]

There are also logistic challenges. Workup for, and subsequent review of CRS-HIPeC patients in the outpatient setting often requires long consultation times. Clinical decision-making is frequently shared with the family. CRS-HIPeC patients in our hospital are reviewed in the general colorectal clinic, comprising 100 to 120 patients with variable colorectal conditions. This limits the duration of each consultation, and some patients may require additional consultations with the presence of other family members to discuss high morbidity and financing. In addition, follow-up telephone calls are needed to ensure there is no miscommunication regarding the subsequent investigations and surgery.

A further obstacle is the limited availability of operating time. This means that the interval to surgery generally may be as long as three months. Although CRC-PM patients are prioritized, this is sometimes at the expense of deferring surgery for other patients.

CRS-HIPeC draws intensive resources: lengthy operation which requires extra anaesthetic and theatre support, complexity of surgical procedures which require multi-disciplinary surgical teams, and several days of ICU stays and ward stays requires specialist nurse and allied healthcare professional supports. Therefore, it becomes an expensive treatment modality, which also carries a high morbidity and mortality rate with the overall 5-year survival being considerably low. At the moment, the cost-effectiveness of CRS-HIPeC is yet to be determined in our setting, in order to draw full financial support from the government fund.

The recently published results from the PRODIGE 7 phase III multicenter randomized control trial showed no significant difference in the overall survival (OS) and relapse free survival (RFS) comparing patients undergoing CRS with or without HIPeC for peritoneal metastases from colorectal cancer.[21] The study also demonstrated that despite no significant difference in 30-day mortality, the 60 day morbidity was higher in the HIPeC group. These results have further raised the conundrum, if HIPeC should really be offered to patients with CRC-PM, especially in our setting with resource limitations.

The result of PRODIGE 7 has had an impact in the treatment of CRC-PM around the world. A web-based survey was conducted among the countries that registered under Peritoneal Surface Oncology Group International (PSOGI) to achieve expert opinion and consensus on the study. Among the several critiques of PRODIGE 7 were the use of oxaliplatin for a duration of 30 minutes during the HIPeC, and PCI score of less than 25 as a criterion for patient selection. Since cytotoxic activity of the chemotherapy is dependent on the duration of exposure and temperature, the 30 minutes exposure used in the PRODIGE 7 was sub-optimal to achieve the optimal oncological activity.

Authors also stated that there was a reduction of using adjuvant HIPeC in 2 out of 18 countries that participated, perhaps due to the PRODIGE 7 results showing no advantage on overall survival.[22] In addition, there was a shift towards mitomycin-based regimens, and increased duration of exposure from 30 minutes to either 60 or 90 minutes. Perhaps, patient selection is the key here, given that subgroup analysis from PRODIGE 7 suggests that a subset of patients could benefit from HIPeC. Furthermore, CRS-HIPeC is used to treat peritoneal metastases from a wide variety of primary malignancies, thus there would still be a need for this service, even if indications in CRC diminish.

Two of our patients with PMP and a patient with peritoneal mesothelioma had CCR-3 because of extensive disease especially on the diaphragm, with PCI scores of 39. In a study by Verwaal et al, prognostic factors depend on the gross residual disease. Residual disease of more than 2.5 cm had a median survival of just 5 months as compared to 17 months in patients with residual disease between 2.5mm – 2.5cm and 39 months in patient with CCR-0.[23] In addition, Konstantinos et al reported that a repeat CRS-HIPeC should not be due to previous CCR-2. Tumour biology plays a crucial role in selecting patients for repeat CRS-HIPeC. A repeat CRS-HIPeC can be undertaken if it can improve survival and control symptomatic disease with good quality of life.[24] Our median DFS is 13 months and median survival is 31 months. Due to the small distribution from the primary cancer in our initial experience, we are unable to analyse individual primary cancer with peritoneal malignancies.

Moving forward, the provision of CRS-HIPeC services in Malaysia must evolve. Given the complexity of resources needed, and considering the logistic and financial challenges to both healthcare providers and patients, we believe that it is essential to centralize such services for each region: Peninsular Malaysia – north, central, south and east coast; East Malaysia – Sabah & Sarawak.

There is limited literature on learning curves for CRS-HIPEC, which can vary widely, depending on the baseline expertise of the team.[25] For example, Kusamura estimated that approximately 140 cases are necessary to ensure surgical proficiency in CRS and HIPEC.[26] Our team members were already experienced general surgeons in independent practice, and were regularly managing other complex colorectal surgeries prior to their fellowships in CRS-HIPEC, so independence was achieved rapidly. These were high volume centres, thus each team member was able to complete 60-100 independently conducted cases each before returning to Malaysia. This exceeds the European School of Peritoneal Surface Oncology (ESPSO) recommendation of a minimum of 20 cases performed independently to overcome learning curves.

Since establishing our service, we have also facilitated the development of public hospital services in Peninsular Malaysia (North – Penang; Central – Seremban). We face obstacles to developing services in other regions, especially in East Malaysia, primarily due to lack of human resource. Thus, we anticipate that patients will still need to travel in the short-term to avail of CRS-HIPeC services. Consideration needs to be given for government funding to support such logistics.

CONCLUSION

We have initiated and provided CRS-HIPEC treatment safely, and the practice should be continued, but quality control, collaboration work and support are required in order to meet international standards. We were able to rapidly set up the CRS-HIPeC multidisciplinary team due to ready availability of experienced specialists in colorectal, hepatobiliary, gynae-oncology, radiology, oncology, anaesthesia and critical care. Coupled with experienced proctoring, our learning curve in establishing our CRS-HIPEC service was relatively short. At the time of establishment, we were the only public hospital offering this service, but since then, we have facilitated the establishment of two more centres, illustrating the demand for PSM management in this country. This demand would only be expected to increase. Therefore, more funding and resources is needed in order to sustain and improve the management of PSM in Malaysia.

REFERENCES

1. Sadeghi B, Arvieux C, Glehen O, Beaujard AC, Rivoire M, Baulieux J, et al. Peritoneal carcinomatosis from non-gynecologic malignancies: results of the EVOCAPE 1 multicentric prospective study. Cancer 2000; 88(2): 358-63.
2. United Nations Development Programme. Human development report 2020: the next frontier – human development and the Anthropocene. http://hdr.undp.org/en/data. Accessed date: 11th October 2024.
3. Esquivel J, Sticca R, Sugarbaker P, Levine E, Yan TD, Alexander R, et al. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in the management of peritoneal surface malignancies of colonic origin: a consensus statement. Ann Surg Oncol 2007; 14(1): 128-33.
4. Yan TD, Deraco M, Baratti D, Kusamura S, Elias D, Glehen O, et al. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy for malignant peritoneal mesothelioma: multi-institutional experience. J Clin Oncol 2009; 27(36): 6237-42.
5. Elias D, Gilly F, Quenet F, Bereder JM, Sidéris L, Mansvelt B, et al. Pseudomyxoma peritonei: a French multicentric study of 301 patients treated with cytoreductive surgery and intraperitoneal chemotherapy. Eur J Surg Oncol 2010; 36(5): 456-62.
6. Esquivel J, Elias D, Baratti D, Kusamura S, Deraco M. Consensus statement on the loco regional treatment of colorectal cancer with peritoneal dissemination. J Surg Oncol 2008; 98(4): 263–7.
7. Turaga K, Levine E, Barone R, Sticca R, Petrelli N, Lambert L, et al. Consensus guidelines from the American Society of Peritoneal Surface Malignancies on standardizing the delivery of hyperthermic intraperitoneal chemotherapy (HIPEC) in colorectal cancer patients in the United States. Ann Surg Oncol 2014; 21(5): 1501-5.

8. Jacquet P, Sugarbaker PH. Clinical research methodologies in diagnosis and staging of patients with peritoneal carcinomatosis. Cancer Treat Res 1996; 82: 359-74.
9. Mehta SS, Bhatt A, Glehen O. Cytoreductive surgery and peritonectomy procedures. Indian J Surg Oncol 2016; 7(2): 139-51.
10. Hirano M, Yonemura Y, Canbay E, Ichinose M, Togawa T, Matsuda T, et al. Laparoscopic diagnosis and laparoscopic hyperthermic intraoperative intraperitoneal chemotherapy for pseudomyxoma peritonei detected by CT examination. Gastroenterol Res Pract 2012; 2012: 1-4.
11. Glehen O, Kwiatkowski F, Sugarbaker PH, Elias D, Levine EA, de Simone M, et al. Cytoreductive surgery combined with perioperative intraperitoneal chemotherapy for the management of peritoneal carcinomatosis from colorectal cancer: a multi-institutional study. J Clin Oncol 2004; 22(16): 3284-92.
12. Dindo D, Demartines N, Clavien PA. Classification of surgical complications. Ann Surg 2004; 240(2): 205-13.
13. Guazzi M, Bandera F, Ozemek C, Systrom D, Arena R. Cardiopulmonary exercise testing: what is its value? J Am Coll Cardiol 2017; 70(13): 1618-36.
14. Neuwirth MG, Alexander HR, Karakousis GC. Then and now: cytoreductive surgery with hyperthermic intraperitoneal chemotherapy (HIPEC), a historical perspective. J Gastrointest Oncol 2016; 7(1): 18-28.
15. Glockzin G, Schlitt HJ, Piso P. Peritoneal carcinomatosis: patient selection, perioperative complications and quality of life related to cytoreductive surgery and hyperthermic intraperitoneal chemotherapy. World J Surg Oncol 2009; 7: 5.
16. García M, Abadía MA, Solórzano F. Segunda citorreducción más quimipterapia hipertérmica intraperitoneal por recaída de seudomixoma peritoneal. Rev Colomb Cir 2017; 32(3): 193–204.
17. Simkens GA, van Oudheusden TR, Braam HJ, Luyer MD, Wiezer MJ, van Ramshorst B, et al. Treatment-related mortality after cytoreductive surgery and HIPEC in patients with colorectal peritoneal carcinomatosis is underestimated by conventional parameters. Ann Surg Oncol 2016; 23(1): 99-105.
18. Teo MCC, Tan GHC, Tham CK, Lim C, Soo KC. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in Asian patients: 100 consecutive patients in a single institution. Ann Surg Oncol 2013; 20(9): 2968-74.
19. Life Insurance Association of Malaysia. Annual report 2019. www.liam.org.my. Accessed date: 25th October 2024.
20. Coefficient G. Household equipment ownership, Malaysia, 2016. https://www.dosm.gov.my/portal-main/release-content/report-on-household-expenditure-survey-2016. Accessed date: 1st June 2024.
21. Roussy G, Elias FD, Goéré D, Quénet F, Elias D, Roca L, et al. Cytoreductive surgery plus hyperthermic intraperitoneal chemotherapy versus cytoreductive surgery alone for colorectal peritoneal metastases (PRODIGE 7): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol 2021; 22: 256-66.
22. van de Vlasakker VCJ, Lurvink RJ, Cashin PH, Ceelen W, Deraco M, Goéré D, et al. The impact of PRODIGE 7 on the current worldwide practice of CRS-HIPEC for colorectal peritoneal metastases: a web-based survey and 2021 statement by Peritoneal Surface Oncology Group International (PSOGI). Eur J Surg Oncol 2021; 47(11): 2888-92.
23. Verwaal VJ, van Tinteren H, van Ruth S, Zoetmulder FAN. Predicting the survival of patients with peritoneal carcinomatosis of colorectal origin treated by aggressive cytoreduction and hyperthermic intraperitoneal chemotherapy. Br J Surg 2004; 91(6): 739-46.
24. Votanopoulos KI, Ihemelandu C, Shen P, Stewart JH, Russell GB, Levine EA. Outcomes of repeat cytoreductive surgery with hyperthermic intraperitoneal chemotherapy for the treatment of peritoneal surface malignancy. J Am Coll Surg 2012; 215(3): 412-7.
25. Chidambarasamy ES, Chia CS, Ong CAJ, Soo KC, Teo MC, Tan GH. Effect of the learning curve of cytoreductive surgery (CRS) and hyperthermic intraperitoneal chemotherapy (HIPEC) on the treatment of colorectal peritoneal metastasis. Asian J Surg 2022; 45(1): 339-45.
26. Kusamura S, Baratti D, Deraco M. Multidimensional analysis of the learning curve for cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in peritoneal surface malignancies. Ann Surg 2012; 255(2): 348-56.

ORIGINAL ARTICLE

High concordance between Immunohistochemistry and RT-PCR in diagnosing ALK rearrangement in lung adenocarcinoma cytologic samples

Didik Setyo Heriyanto, MD[1,2], Lisnawati Rachmadi, MD[3], Ika Trisnawati, MD[4], Jeffry Beta Tenggara, MD[5], Rita Cempaka, MD[1], Haryo Aribowo, MD[2], Yunanto Kurnia, MD[2], Vincent Lau, MD[1], Andrew Nobiantoro Gunawan, MD[1], Brigitta Natasya Halim, MD[1], Fara Silvia Yuliani, DMD[6], Vincent Laiman, MD[7], Hsiao-Chi Chuang, Prof[8,9,10,11]

[1]Department of Anatomical Pathology, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada/Dr. Sardjito, General Hospital Yogyakarta, Indonesia, [2]Division of Cardiac, Thoracic, and Vascular Surgery, Department of Surgery, Faculty of Medicine, Public Health, and Nursing/Dr. Sardjito General Hospital, Yogyakarta, Indonesia, [3]Department of Anatomical Pathology, Faculty of Medicine Universitas Indonesia-Cipto Mangunkusumo Hospital, Jakarta, Indonesia, [4]Division of Pulmonary Medicine, Department of Internal Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada/Dr. Sardjito General Hospital, Yogyakarta, Indonesia, [5]Division of Hematology and Medical Oncology, Department of Internal Medicine, MRCCC Siloam Hospital Jakarta, Jakarta, Indonesia, [6]Department of Pharmacology and Therapy, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia, [7]Department of Radiology, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada/Dr. Sardjito General Hospital Yogyakarta, Indonesia, [8]School of Respiratory Therapy, College of Medicine, Taipei Medical University, Taipei, Taiwan, [9]Division of Pulmonary Medicine, Department of Internal Medicine, Shuang Ho Hospital, Taipei Medical University, Taipei, Taiwan, [10]Cell Physiology and Molecular Image Research Center, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan, [11]National Heart and Lung Institute, Imperial College London, London SW7 2AZ, UK

ABSTRACT

Introduction: Lung cancer, predominantly lung adenocarcinoma, remains a major health challenge in Indonesia, with late-stage detection being common. This study explores the use of quantitative real-time polymerase chain reaction (qRT-PCR) for assessing ALK rearrangement from smear samples, a significant shift towards less invasive diagnostic methods, by assessing its concordance with immunohistochemistry (IHC) in detecting Anaplastic Lymphoma Kinase (ALK) rearrangements in lung adenocarcinoma patients.

Materials and Methods: This observational cross-sectional study analyzed 175 lung adenocarcinoma samples lacking EGFR mutations collected between 2018 and 2022. IHC was performed with the Ventana ALK D5F3 clone antibody on cell blocks or core needle biopsy specimens. The EML4-ALK fusion rearrangement status was determined using quantitative RNA qRT-PCR analysis on the smear specimen from transthoracic needle aspiration (TTNA) from the same sample. Only specimens with viable tumor cells were included, ensuring the exclusion of metastatic or necrotic samples.

Results: ALK rearrangements were identified in 16.2% (23/142) of samples via IHC and 14.8% (21/142) via qRT-PCR. Prevalence did not significantly differ by age and sex. The study found a 98.5% concordance rate between the two methods, with a κ coefficient of 0.95 (95% CI, 0.91-0.98), indicating almost perfect agreement.

Conclusion: The high concordance between IHC and qRT-PCR underscores their reliability in detecting ALK rearrangements, crucial for the precise diagnosis and treatment of lung adenocarcinoma in Indonesia. These findings support the use of either method, depending on available resources and expertise, to enhance lung cancer management.

KEYWORDS:
Lung cancer; Adenocarcinoma; Anaplastic Lymphoma Kinase (ALK); Immunohistochemistry (IHC); Quantitative Real Time-Polymerase Chain Reaction (qRT-PCR)

INTRODUCTION

Lung cancer remains the leading cause of cancer-related morbidity and mortality worldwide. In 2020, Indonesia experienced 34,783 new lung cancer diagnoses, and 30,843 fatalities attributed to lung cancer, representing 13.2% of the nation's total cancer death toll.[1] These numbers underscore the urgent need for effective management strategies.

Lung cancer is primarily classified into two types: small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC), with NSCLC making up approximately 80–85% of all lung cancers.[2,3] NSCLC is further subdivided into squamous-cell carcinoma, large cell carcinoma, and adenocarcinoma. Often, adenocarcinoma is the most common and usually presents at an advanced stage with limited treatment modalities.[4] The National Comprehensive Cancer Network (NCCN) for NSCLC emphasizes the importance of molecular testing in advanced and metastatic NSCLC.[5] The emergence of targeted therapies using these molecular markers has transformed the field of precision medicine. The rapid advancement in molecular biology has

This article was accepted: 21 April 2025
Corresponding Author: Didik Setyo Heriyanto
Email: didik_setyoheriyanto@mail.ugm.ac.id

led to the identification and confirmation of lung cancer driver genes, paving the way for the development of molecularly targeted medications and the onset of the era of targeted drug therapy.[5]

The Anaplastic Lymphoma Kinase (ALK) rearrangement has emerged as a distinct molecular subtype in NSCLC since 2007, with the prevalence approximately 5%.[6,7] Patients typically present with adenocarcinoma histology and are often light smokers or have never smoked.[8] Research has found that echinoderm microtubule-associated protein-like 4 (EML4)-ALK fusion gene is typically exclusive of other carcinogenic genes such as EGFR, ROS1, KRAS, and others.[9] EML4 represents the most frequent partner for ALK in lung cancer.[10,11] Thus, detection of the EML4-ALK fusion gene is crucial for administering targeted therapy.[9] The NCCN recommends testing the ALK rearrangements in patients with metastatic nonsquamous NSCLC. This recommendation is based on the efficacy of ALK inhibitors like alectinib (which is covered by national insurance of Indonesia), brigatinib, ceritinib, crizotinib, or lorlatinib. Moreover, since 2023, the detection of ALK rearrangements in early-stage NSCLC (stages IB-IIIA, and IIIB [only T3, N2]) may evaluate the possibility of adjuvant therapy with atezolizumab or pembrolizumab.[5]

Three primary techniques exist for ALK testing: immunohistochemistry (IHC), fluorescence in situ hybridization (FISH), and real-time polymerase chain reaction (RT-PCR).[10,12] The FISH method of testing was considered as the gold standard of fusion gene detection including ALK fusions, but this method is notably labor-intensive and its implementation remains limited, particularly in developing countries.[12,13] In regions such as Indonesia, IHC and RT-PCR are widely accessible and become feasible alternatives for biomolecular testing, especially ALK rearrangements. Previous studies have consistently demonstrated a substantial agreement between the results obtained through IHC or RT-PCR when compared with those derived from FISH.[10,12,14–17]

The increasing accessibility of molecular testing methods in Indonesia presents a unique opportunity to study the prevalence of oncogenic mutations in lung adenocarcinoma patients, especially important in advanced stages where extensive tissue resection is limited. The role of minimally invasive diagnostic methods like transthoracic needle aspiration (TTNA) becomes crucial, although they yield only smear samples. In Indonesia, biomarker testing has traditionally relied on protein-based techniques such as IHC, which necessitate tissue or cell block preparation.[18–20] This study aims to determine the feasibility of using qRT-PCR to accurately evaluate ALK rearrangement from cytology smear samples, marking a notable advancement in the validation of less invasive sample types for molecular diagnostics, while investigating the concordance between IHC and RT-PCR analyses within a patient group. To our knowledge, no study in Indonesia has compared RT-PCR with IHC for ALK testing. However, research from other countries has demonstrated a high concordance between RT-PCR and IHC in ALK testing.[21,22] Achieving this could facilitate more precise and early interventions, ultimately improving treatment outcomes by leveraging the adaptability of diagnostic tests to various sample types, thus enhancing the prospects for effective management of lung adenocarcinoma.

MATERIALS AND METHODS

Study Design

This study was an observational cross-sectional study employing consecutive sampling methods. The research was conducted using NSCLC samples collected between 2018 and 2022 from the Department of Anatomical Pathology, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada and Cito Clinical Laboratory, Yogyakarta, Indonesia. Ethical approval for specimen collection was granted by the Medical and Health Research Ethics Committee (MHREC) of the Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Universitas Gadjah Mada - Dr. Sardjito General Hospital, Yogyakarta, Indonesia (Ref. No.: KE/FK/0532/EC/2020). The study adhered to the principles outlined in the Declaration of Helsinki. Given its retrospective and non-interventional design, the Ethics Committee granted a waiver for the implementation of a free informed consent.

Sample Collection, Preparation, and Selection

All specimens were primarily obtained via TTNA cytology. In cases where clinical conditions permitted, a core needle biopsy was also performed on the same patient, resulting in two types of samples (i.e., cytology and histology) from a single individual. For TTNA cytology samples, a portion of the aspirated material was prepared as a Diff-QuickTM stained smear, while the remaining material (residual clotting material from the syringe) was processed into a formalin-fixed, paraffin-embedded (FFPE) cell block, particularly when a core needle biopsy was not feasible. The cytology smear was reserved for RNA extraction and subsequent RT-PCR analysis, whereas the FFPE cell block and any available core needle biopsy specimens were utilized for IHC examination.

A total of 175 specimens of native Indonesian, that initially tested negative for EGFR mutations were evaluated for inclusion in this study. Only the specimens that were identified as lung adenocarcinoma were selected for additional analysis. Each sample was required to exhibit at least 30% tumor cell content (according to AmoyDx® EML4-ALK Fusion Gene Detection Kit [Amoy Diagnostics, Xiamen, China] and Lindeman et al., 2018),[23] following post-EGFR mutation testing to ensure sufficient cellular material for subsequent analyses. Furthermore, the availability of both a cytology smear (for qRT-PCR) and FFPE cell block or core needle biopsy (for IHC) sample from the same patient was necessary. Specimens indicated metastatic lung cancer originating from non-pulmonary primary sites, or predominantly displayed necrosis, which rendered adequate molecular and histological analysis unfeasible are excluded from the analyses.

Immunohistochemistry (IHC) Analysis

ALK rearrangement was evaluated using Ventana® ALK (D5F3 clone antibody) CDx assay (Roche Diagnostics, Indiana, USA) on FFPE cell block tissue sections or core needle biopsy tissue sections with a thickness of 4 µm, following the

manufacturer's standard protocols. This kit is already certified for in vitro diagnostics (IVD). The ALK status interpretation was determined based on specific criteria provided by the clone antibody manufacturer. A positive result required strong granular cytoplasmic staining in tumor cells, irrespective of the percentage of positive tumor cells. Certain staining elements were excluded, including light cytoplasmic stripping in alveolar macrophages, staining in neural cells (e.g., nerve and ganglion cells), glandular epithelial staining, scattered lymphoreticular cells within lymphocytic infiltrates, and background staining in normal mucosa or necrotic tumor areas. Absence of strong granular cytoplasmic staining in tumor cells indicating negative ALK status.

Two independent pathologists, blinded to each other's assessments and to the RT-PCR results, evaluated the stained slides according to the manufacturer's interpretation guide. Results were categorized as positive or negative. In instances of discordant interpretations, a third blinded pathologist reviewed the slides to reach a consensus.

RNA Extraction

RNA extraction was performed on scraped TTNA smear preparations using the RibospinTM II RNA Purification Kit (Cat. No. 314-150, GeneAll Biotechnology Co., Ltd., Seoul, Korea) according to the manufacturer's instructions. This manual process was conducted by trained personnel to maintain RNA integrity and quality. Samples intended for immediate analysis were stored at 4°C, while those designated for long-term storage were kept at -70°C to prevent degradation.

Reverse Transcription Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

The RNA samples in this study were analyzed using the AmoyDx® EML4-ALK Fusion Gene Detection Kit (Amoy Diagnostics, Xiamen, China), which is certified under CE-IVD (https://www.amoydiagnostics.com/products/amoydx-eml4-alk-fusion-gene-detection-kit) standard. This analysis was conducted in strict adherence to the manufacturer's guidelines. Quantitative real-time PCR (qRT-PCR) was performed on the Bioneer Exicycler96TM Real-Time Quantitative Thermal Block (ver. 4, Bioneer inc., California, USA) platform. The PCR conditions followed were as recommended by the manufacturer's protocol. The AmoyDx® EML4-ALK Fusion Gene Detection Kit (Amoy Diagnostics, Xiamen, China) validation process, conducted in accordance with vendor's recommendation, ensured the kit's reliability and the accuracy using Bioneer Exicycler96TM Real-Time Quantitative Thermal Block (ver. 4, Bioneer inc., California, USA) platform. Interpretation of the PCR results was based on the FAM fluorescence channel cycle threshold (Ct) values. Samples with a Ct value of less than 30 were classified as positive for EML4-ALK fusion genes. Conversely, samples with a Ct value of 30 or higher were deemed negative, suggesting no detectable EML4-ALK fusion or values below the kit's limit of detection (LOD).

Statistical Analysis

Statistical evaluations in this study were conducted using Microsoft Excel 2020 and IBM SPSS Statistics 27.0 software. The analysis categorized data into two types: categorical data, presented as frequencies and percentages, and quantitative data, expressed as mean values. For categorical data, the Chi-square test was used to assess associations, while for quantitative data, the independent t-test was employed to compare means between groups. For non-parametric data, the Mann-Whitney U test was used as an alternative to the independent t-test.

A univariate analysis was employed to gain insights into each variable. For binomial variables, the confidence interval was determined using the Clopper-Pearson method, specifically set at a 95% confidence interval (CI). The concordance rate was determined from the contingency table.[24,25] The 95% CI was estimated for κ. κ coefficient of ≤ 0 indicates no agreement, while coefficients between 0.01 and 0.20 reflect slight agreement. κ coefficient between 0.21 and 0.40 indicates fair agreement, and values between 0.41 and 0.60 are considered moderate. Substantial agreement is represented by coefficient between 0.61 and 0.80, and κ coefficient of more than 0.80 will correspond to an almost perfect agreement as per the scale proposed by Landis and Koc.[24]

RESULTS

After applying inclusion and exclusion criteria, 142 samples met all the necessary requirements and were included in the final analysis. The baseline characteristics of the samples is shown in Table I. The mean age of the patients was 60 years old, with 24 (16.9%) samples were 50 years of age or younger and 118 (83.1%) samples were older than 50. Of the 142 samples, a total of 79 samples (55.6%) were from males, and 63 samples (44.4%) were from females.

All the samples were successfully tested for ALK rearrangement status by IHC and qRT-PCR. The immunostaining pattern of ALK in lung adenocarcinoma was assessed using the Ventana anti-ALK (D5F3) antibody. The ALK IHC assay demonstrates significant cytoplasmic protein expression, characterized by a distinct granular pattern, in cases that produced positive results (Figure 1A&1B). For testing with IHC, 23/142 (16.2%) samples were positive for ALK rearrangement, while the remaining 119/142 (83.8%) samples had negative ALK rearrangement status (Table II). The result was similar with qRT-PCR testing. There were 21/142 (14.8%) samples with positive ALK rearrangement and 121/142 (85.3%) samples with negative ALK rearrangement (Table II).

The prevalence of individuals aged > 50 years old was greater than those ≤ 50 years old in both the IHC (8.3% versus 17.8%; p>0.05) and qRT-PCR (8.3% versus 16.1%; p>0.05) results, but these differences were not statistically significant. The prevalence of ALK rearrangement was also more common in males than females, as indicated by both IHC (20.3% versus 11.1%; p>0.05) and qRT-PCR (19.0% versus 9.5%; p>0.05) results. However, these differences also did not reach statistical significance.

In terms of the agreement between the IHC and qRT-PCR result, it was determined that cases could be classified as

Table I: Baseline characteristics of the sample

Characteristics	N (%)
Age	
≤ 50 years old	24 (16.9%)
> 50 years old	118 (83.1%)
Sex	
Male	79 (55.6%)
Female	63 (44.4%)

Table II: Characteristics of IHC for ALK rearrangement and qRT-PCR for EML4-ALK fusion

Characteristics	ALK rearrangement status by IHC		p-value	EML4-ALK fusion status by qRT-PCR		p-value
	ALK rearrangement (+) N (%)	ALK rearrangement (-) N (%)		ALK rearrangement (+) N (%)	ALK rearrangement (-) N (%)	
Prevalence	23 (16.2%)	119 (83.8%)		21 (14.8%)	121 (85.2%)	
Age			0.251			0.328
≤ 50 years old	2 (8.3%)	22 (91.7%)		2 (8.3%)	22 (91.7%)	
> 50 years old	21 (17.8%)	97 (82.2%)		19 (16.1%)	99 (83.9%)	
Sex		0.142			0.115	
Male	16 (20.3%)	63 (79.7%)		15 (19.0%)	64 (81.0%)	
Female	7 (11.1%)	56 (88.9%)		6 (9.5%)	57 (90.5%)	

Table III: Cross-tabulation of IHC for ALK rearrangement and qRT-PCR for EML4-ALK fusion

Examination Methods	EML4-ALK fusion status by RT-PCR		Total
	ALK rearrangement (+) N	ALK rearrangement (-) N	
ALK rearrangement status by IHC			
ALK rearrangement (+)	21	2	23
ALK rearrangement (-)	0	119	119
Total	21	121	142

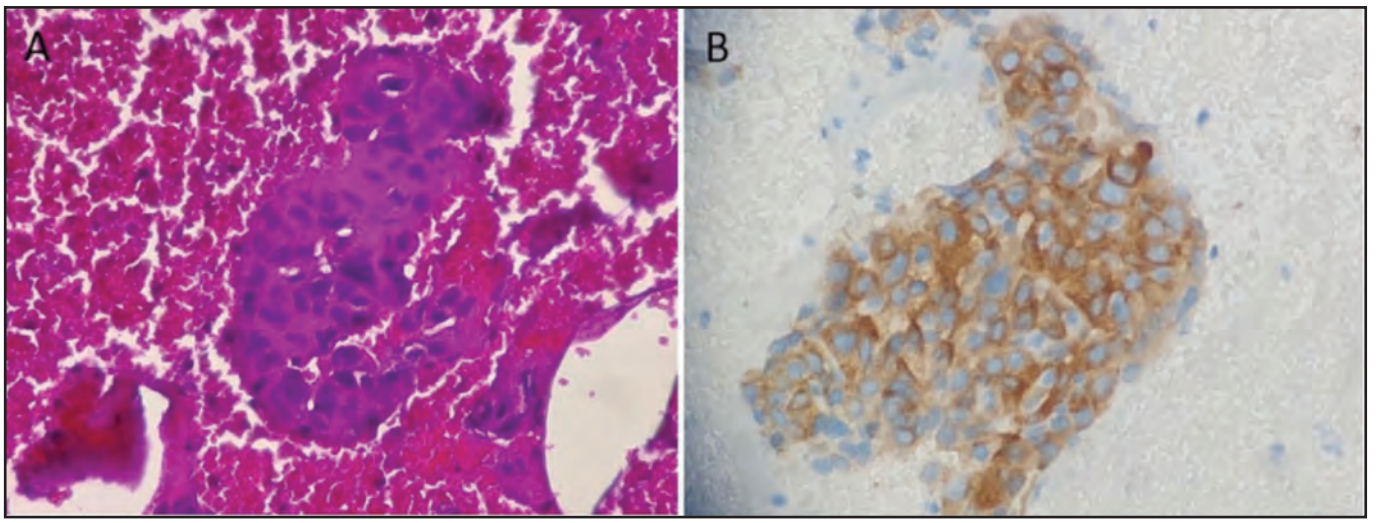


Fig. 1: A. Photomicrograph of the cell block from TTNA sample that revealed a lung adenocarcinoma (Cell block, HE, 400x); B. The immunostaining pattern of ALK in lung adenocarcinoma using the Ventana anti-ALK (D5F3) antibody. ALK IHC demonstrates a significant level of protein expression in the cytoplasm, characterized by a granular pattern, in the cases that tested positive (Cell block, Anti-ALK antibody, 400x)

either IHC negative/qRT-PCR negative or IHC positive/qRT-PCR positive. The cross-tabulation of ALK rearrangement status by IHC with D5F3 antibody in relation to ALK rearrangement status by qRT-PCR is shown (Table III). All 21 (100%) of positive ALK rearrangement status samples tested with qRT-PCR also showed positive results with IHC.

Only two (8.7%) samples that were tested positive with IHC showed negative results with qRT-PCR. Meanwhile, all the negative samples tested with IHC showed the same result with qRT-PCR. The concordance between the two tests was 98.5% with a corresponding κ coefficient of 0.95 (95%CI, 0.91–0.98). The κ coefficient is greater than 0.80, which corresponds to almost perfect agreement.

DISCUSSION

In this study, we reported a prevalence of ALK gene rearrangement between 14.8% (by qRT-PCR) and 16.2% (by IHC) in a sample of 142 Indonesian patients. These observations mark a higher prevalence rates of ALK gene rearrangement in NSCLC patients, which globally reported to be between 2-7% among individuals with adenocarcinoma histology subtype, predominantly comprising of light smokers or non-smokers patients; its presence is usually exclusive to other driver mutations.[5,7,8,13] Comparable neighboring countries, including Thailand, Malaysia, and Vietnam, have also exhibited a reduced incidence of ALK rearrangement ranging from 4.1% to 9.7%.[26–29] Our earlier study also reflected this trend, revealing that 20% samples (seven out of 35 cytology samples) of the Indonesian population exhibiting a positive EML4-ALK rearrangement.[30] Additionally, this finding gains further significance when compared to the more common and well-established data of EGFR mutation in the Indonesian lung cancer patients, where a larger study of 1,874 patients found a 44.5% positive rate of EGFR mutation.[31]

Considering the prevalent EGFR-targeted treatments in Indonesia, our research emphasizes the necessity of also focusing on ALK rearrangements. Although EML4-ALK fusion genes generally do not coexist with other oncogenic mutations, such as EGFR, ROS1, or KRAS, exceptions have been reported in a minority of patients.[32] These rare cases, where co-occurring actionable oncogenic drivers are present, highlight an area of clinical complexity that warrants further exploration.[32]

ALK rearrangements may contribute to osimertinib resistance in NSCLC patients with EGFR mutations, underscoring the importance of identifying and targeting ALK rearrangements even within this subset.[33] The exploration and consideration of various ALK inhibitors are crucial steps towards enhancing treatment outcomes for Indonesian patients. The NCCN guideline strongly recommends testing for both ALK and EGFR mutations, supported by substantial evidence to optimize therapeutic strategies.[5]

A systematic review and meta-analysis showed higher specificity for cell specimens than for tumor specimens for detecting ALK rearrangement with IHC compared to FISH in NSCLC patients.[34] In this study, we utilized the cell block, a smaller form of formalin-fixed paraffin-embedded, obtained from fine-needle aspirate or fluid sediment that can be utilized for IHC study. The retention of cytologic material in the cell block for methods designed for IHC and molecular studies enhances its diagnostic precision.[35]

The IHC has been demonstrated to be a cost-effective, fast, and accurate method for detecting ALK rearrangements in tissue or cell block samples, offering advantages over FISH due to lower requirements for specialized technical resources and expertise.[15,18] Both IHC and RT-PCR have shown high sensitivity and specificity, making them suitable for routine clinical practice.[15,18,36] Previous studies have shown a high concordance between IHC and FISH, some with rates of 98.4% for negative agreement and 98.5% for overall agreement.[14,18,36] This further validates the efficacy of IHC. Uruga et al.,[37] highlighted IHC's performance, with the D5F3 clone achieving 76-100% sensitivity and specificity compared to FISH in studies conducted between 2011-2018. Additionally, RT-PCR has demonstrated a 99.2% concordance with FISH in patients treated with crizotinib, affirming its reliability in detecting EML4-ALK fusion in NSCLC patients, with a specificity of 94% in cases without full-length ALK expression when compared to FISH and sequencing.[38,39] A comparison with NGS also confirmed RT-PCR's reliability.[40] With their validated sensitivity and specificity, both IHC, particularly using the D5F3 test, and RT-PCR are indispensable in the molecular diagnosis of NSCLC, offering efficient, reliable, and cost-effective alternatives for identifying ALK rearrangements and optimizing patient treatment plans.

We reported a strong concordance between the data obtained using IHC with Ventana D5F3 antibody and RT-PCR for EML4-ALK fusion. The observation that two samples were positive for the IHC D5F3 clone but negative according to the EML4-ALK RT-PCR test suggests the possibility of rare ALK fusion partners that are not typically identified by our RT-PCR kit. The observed discordance between testing methods can largely be attributed to the limited scope of EML4-ALK fusions detected by the RT-PCR assay and the dependency on the quality of RNA in the samples, as highlighted by numerous studies.[40–42] Both the methodology employed in this study, specifically IHC for D5F3 ALK rearrangement and RT-PCR for EML4-ALK fusion, have demonstrated a high level of sensitivity and high concordance with FISH in previous research.[15,38,39] While FISH is still included in some guidelines as a means to predict ALK inhibitor sensitivity, IHC and RT-PCR testing can be considered as reliable alternative methods for predicting the use of ALK inhibitors, particularly in Indonesia. Cytology specimens continue to be the most accessible type of sample.

These methods are especially useful for identifying patients who are eligible for treatment with tyrosine kinase inhibitors (alectinib, brigatinib, and crizotinib), particularly in instances where tissue samples are unavailable. Therefore, our findings can be utilized as a reliable reference to improve the ALK rearrangement assessment in Indonesia and another country where cytological samples or small biopsies are commonly used for diagnosing lung cancer. Both the IHC and RT-PCR methods are reliable for accurate detection of ALK rearrangements.

This analysis is particularly important in Indonesia and neighboring nations, where healthcare resources may be constrained, and effective diagnostic strategies are essential for enhancing patient outcomes while controlling expenses. The elevated concordance rates between IHC and RT-PCR indicate that both techniques are dependable for diagnosing ALK rearrangements, which are essential for targeted therapies in lung adenocarcinoma. The cost-effectiveness of these diagnostic methods can greatly impact clinical decision-making. Research indicates that molecular diagnostics, such as IHC and RT-PCR, facilitate earlier and more precise diagnoses, thereby potentially decreasing the overall treatment expenses linked to advanced lung cancer management.[43,44] This is especially crucial in areas with

limited healthcare budgets, as prompt and precise diagnostics may prevent the financial strain of treating advanced diseases.

Furthermore, the incorporation of economical diagnostic methods, such as the integration of IHC and RT-PCR, can enhance resource distribution within healthcare systems. The implementation of rapid on-site evaluation (ROSE) alongside endobronchial ultrasound transbronchial lung biopsy has demonstrated a reduction in diagnostic expenses while enhancing the efficiency of lung cancer diagnosis.[45] This corresponds with findings that highlight the necessity of employing less invasive and more economical diagnostic instruments to improve patient care while avoiding excessive expenses.[46]

In the future, we can improve the generalizability of our study by including samples from other types of NSCLC, such as squamous cell carcinoma. In our current study, we only focused on the samples of adenocarcinoma. In addition, this study does not employ FISH as a definitive benchmark due to limited sample availability and procurement method in this scope of the investigation. Hence, it is recommended as a potential subject for future investigation, to carry out concordance and diagnostic examinations employing other methodologies.

CONCLUSION

With high concordance, the status of ALK rearrangement in lung adenocarcinoma can be determined using either IHC or qRT-PCR. Depending on the availability of samples, the most suitable assay may be employed. In a country where cytological samples or small biopsies are commonly used for diagnosing lung cancer, the IHC or qRT-PCR methods can be relied upon for accurate testing of ALK rearrangements.

ACKNOWLEDGEMENT

We acknowledge that there are no additional acknowledgements to declare.

CONFLICT OF INTEREST

The authors declared that they had no competing interests.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021; 71(3): 209-49.
2. Guo H, Zhang J, Qin C, et al. Biomarker-Targeted Therapies in Non–Small Cell Lung Cancer: Current Status and Perspectives. Cells 2022; 11(20): 3200.
3. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. Transl Lung Cancer Res 2016; 5(3): 288.
4. Doval DC, Prabhash K, Patil S, et al. Clinical and epidemiological study of EGFR mutations and EML4-ALK fusion genes among Indian patients with adenocarcinoma of the lung. Onco Targets Ther 2015; 8: 117-23.
5. Riely GJ, Wood DE, Ettinger DS, et al. Non-Small Cell Lung Cancer, Version 4.2024, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2024; 22(4): 249-74.
6. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. Nature 2007 448: 7153 2007; 448(7153): 561-6.
7. Kwak EL, Bang Y-J, Camidge DR, et al. Anaplastic Lymphoma Kinase Inhibition in Non–Small-Cell Lung Cancer. New England Journal of Medicine 2010; 363(18): 1693-703.
8. Shaw AT, Yeap BY, Mino-Kenudson M, et al. Clinical Features and Outcome of Patients With Non–Small-Cell Lung Cancer Who Harbor EML4-ALK. Journal of Clinical Oncology 2009; 27(26): 4247.
9. Lu S, Lu C, Xiao Y, et al. Comparison of EML4-ALK fusion gene positive rate in different detection methods and samples of non-small cell lung cancer. J Cancer 2020; 11(6): 1525.
10. Hernandez S, Conde E, Alonso M, et al. A narrative review of methods for the identification of ALK fusions in patients with non-small cell lung carcinoma. Transl Lung Cancer Res 2023; 12(7): 1549-62.
11. Morán T, Quiroga V, de los Llanos Gil M, et al. Targeting EML4-ALK driven non-small cell lung cancer (NSCLC). Transl Lung Cancer Res 2013; 2(2): 128-41.
12. Kozuma Y, Toyokawa G, Seto T. ALK testing methods: is there a winner or loser? Expert Rev Anticancer Ther 2019; 19(3): 237-44.
13. Chia PL, Dobrovic A, Dobrovic A, et al. Prevalence and natural history of ALK positive non-small-cell lung cancer and the clinical impact of targeted therapy with ALK inhibitors. Clin Epidemiol 2014; 6: 423-32.
14. Jazieh AR, Gaafar R, Errihani H, et al. Real-World Data on the Prevalence of Anaplastic Lymphoma Kinase–Positive Non–Small-Cell Lung Cancer in the Middle East and North Africa. JCO Glob Oncol 2021; 7(7): GO.21.00067.
15. Li Y, Pan Y, Wang R, et al. ALK-Rearranged Lung Cancer in Chinese: A Comprehensive Assessment of Clinicopathology, IHC, FISH and RT-PCR. PLoS One 2013; 8(7): e69016.
16. Hofman P, Lai R. ALK in Non-Small Cell Lung Cancer (NSCLC) Pathobiology, Epidemiology, Detection from Tumor Tissue and Algorithm Diagnosis in a Daily Practice. Cancers 2017; 9(8): 107.
17. Zhang YG, Jin ML, Li L, et al. Evaluation of ALK Rearrangement in Chinese Non-Small Cell Lung Cancer Using FISH, Immunohistochemistry, and Real-Time Quantitative RT- PCR on Paraffin-Embedded Tissues. PLoS One 2013; 8(5): e64821.
18. Guo HQ, Jia J, Zhao LL, et al. Application of Ventana immunocytochemical analysis on ThinPrep cytology slides for detection of ALK rearrangement in patients with advanced non-small-cell lung cancer. BMC Cancer 2018; 18(1): 1-8.
19. Tanaka H, Tone K, Hayashi A, et al. Clinical application of immunocytochemical detection of ALK rearrangement on cytology slides for detection or screening of lung adenocarcinoma. Lung Cancer 2013; 80(3): 289-92.
20. Zhang C, Randolph ML, Jones KJ, et al. Anaplastic Lymphoma Kinase Immunocytochemistry on Cell-Transferred Cytologic Smears of Lung Adenocarcinoma. Acta Cytol 2015; 59(2): 213-8.
21. Xu C wei, Wang W xian, Chen Y ping, et al. Simultaneous VENTANA IHC and RT-PCR testing of ALK status in Chinese non-small cell lung cancer patients and response to crizotinib. J Transl Med 2018; 16(1): 1-10.
22. Teixidó C, Karachaliou N, Peg V, et al. Concordance of IHC, FISH and RT-PCR for EML4-ALK rearrangements. Transl Lung Cancer Res 2014; 3(2): 70.
23. Lindeman NI, Cagle PT, Aisner DL, et al. Updated Molecular Testing Guideline for the Selection of Lung Cancer Patients for Treatment With Targeted Tyrosine Kinase Inhibitors: Guideline From the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. Arch Pathol Lab Med 2018; 142(3): 321-46.
24. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977; 33(1): 159.

25. Hackner K, Buder A, Hochmair MJ, et al. Detection of EGFR Activating and Resistance Mutations by Droplet Digital PCR in Sputum of EGFR-Mutated NSCLC Patients. Clin Med Insights Oncol 2021; 15.
26. Zhou X, Shou J, Sheng J, et al. Molecular and clinical analysis of Chinese patients with anaplastic lymphoma kinase (ALK)-rearranged non-small cell lung cancer. Cancer Sci 2019; 110(10): 3382-90.
27. Poh ME, How SH, Ho GF, et al. Real-World Treatment and Outcomes of ALK-Positive Metastatic Non–Small Cell Lung Cancer in a Southeast Asian Country. Cancer Manag Res 2023; 15: 31.
28. Incharoen P, Reungwetwattana T, Saowapa S, et al. ALK-rearranged pulmonary adenocarcinoma in Thai Patients: From diagnosis to treatment efficacy. World J Surg Oncol 2016; 14(1): 1-9.
29. Dang ATH, Tran VU, Tran TT, et al. Actionable Mutation Profiles of Non-Small Cell Lung Cancer patients from Vietnamese population. Sci Rep 2020; 10(1): 2707.
30. Heriyanto DS, Trisnawati I, Kumara EG, et al. The Prevalence of the EML4-ALK Fusion Gene in Cytology Specimens from Patients with Lung Adenocarcinoma. Pulm Med 2020; 2020(1): 3578748.
31. Syahruddin E, Fachrucha F, Zaini J, et al. Liquid biopsy and tissue biopsy for epidermal growth factor receptor (EGFR) mutation analysis in lung adenocarcinoma. The Indonesian Journal of Cancer Control 2024; 3(1): 9-13.
32. Zhao Y, Wang S, Yang Z, et al. Co-Occurring Potentially Actionable Oncogenic Drivers in Non-Small Cell Lung Cancer. Front Oncol 2021; 11: 665484.
33. Hou H, Sun D, Zhang C, et al. ALK rearrangements as mechanisms of acquired resistance to osimertinib in EGFR mutant non-small cell lung cancer. Thorac Cancer 2021; 12(6): 962.
34. Ma H, Yao WX, Huang L, et al. Efficacy of D5F3 IHC for detecting ALK gene rearrangement in NSCLC patients: a systematic review and meta-analysis. Oncotarget 2016; 7(43): 70128.
35. Gupta N, Sekar A, Rajwanshi A. Role of FNAC, fluid specimens, and cell blocks for cytological diagnosis of lung cancer in the present era. Journal of Cytology / Indian Academy of Cytologists 2015; 32(4): 217.
36. Yatabe Y. ALK FISH and IHC: You cannot have one without the other. Journal of Thoracic Oncology 2015; 10(4): 548-50.
37. Uruga H, Mino-Kenudson M. ALK (D5F3) CDx: an immunohistochemistry assay to identify ALK-positive NSCLC patients. Pharmgenomics Pers Med 2018; 11: 147-55.
38. Hout DR, Schweitzer BL, Lawrence K, et al. Performance of a RT-PCR Assay in Comparison to FISH and Immunohistochemistry for the Detection of ALK in Non-Small Cell Lung Cancer. Cancers 2017; 9(8): 99.
39. Wang Y, Zhang J, Gao G, et al. EML4-ALK Fusion Detected by RT-PCR Confers Similar Response to Crizotinib as Detected by FISH in Patients with Advanced Non-Small-Cell Lung Cancer. Journal of Thoracic Oncology 2015; 10(11): 1546-52.
40. Kuang Y, Xu P, Wang J, et al. Detecting ALK Rearrangement with RT-PCR: A Reliable Approach Compared with Next-Generation Sequencing in Patients with NSCLC. Mol Diagn Ther 2021; 25(4): 487-94.
41. Wang B, Chen R, Wang C, et al. Identification of novel ALK fusions using DNA/RNA sequencing in immunohistochemistry / RT-PCR discordant NSCLC patients. Hum Pathol 2021; 114: 90-8.
42. Letovanec I, Finn S, Zygoura P, et al. Evaluation of NGS and RT-PCR Methods for ALK Rearrangement in European NSCLC Patients: Results from the European Thoracic Oncology Platform Lungscape Project. Journal of Thoracic Oncology 2018; 13(3): 413-25.
43. Nkosi D, George G V., Liu H, et al. Efficient Lung Cancer Molecular Diagnostics by Combining Next Generation Sequencing with Reflex Idylla Genefusion Assay Testing. Genes (Basel) 2023; 14(8): 1551.
44. Wolff HB, Steeghs EMP, Mfumbilwa ZA, et al. Cost-Effectiveness of Parallel Versus Sequential Testing of Genetic Aberrations for Stage IV Non–Small-Cell Lung Cancer in the Netherlands. JCO Precis Oncol 2022; (6).
45. Xu CH, Wang JW, Wang W, et al. The diagnosis value of endobronchial ultrasound transbronchial lung biopsy combined with rapid on-site evaluation in peripheral lung cancer. Clin Respir J 2020; 14(5): 447-52.
46. Simon M, Simon I, Tent PA, et al. Cryobiopsy in Lung Cancer Diagnosis—A Literature Review. Medicina (B Aires) 2021; 57(4): 393.

Exploring the factors leading to tiered referrals of pregnant women until tertiary healthcare facilities: An in-depth analysis

M Masrur Rizal, MD[1,2], Rizki Pranadyan, OBGYN[2], Alifina Izza, MIDWIFE[1,2], Hanifa Erlin Dharmayanti, OBGYN[2], Pandu Hanindito Habibie, OBGYN[2], Aida Musyarrofah, OBGYN[3]

[1]Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia, [2]Division of Social Obstetrics and Gynecology, Department of Obstetrics and Gynecology, General and Academic Hospital Dr. Soetomo, Indonesia, [3]Faculty of Medicine, Universitas Muhammadiyah Malang, Malang, Indonesia

ABSTRACT

Introduction: Maternal mortality in Indonesia remains a significant health issue, with a mortality rate of 305 per 100,000 live births, the highest in Southeast Asia. Prolonged referral processes and delays in receiving medical assistance are major factors contributing to the high maternal mortality rate.

Objective: This study aims to explore the factors influencing tiered referrals of pregnant women to tertiary healthcare facilities.

Materials and Methods: This descriptive analytical study uses secondary data from medical records of patients admitted to the Obstetrics and Gynecology Maternity Room at Dr. Soetomo Hospital, Surabaya, from July to September 2023, with ethical clearance number 2813/104/4/III/2024. Sampling was done using the total sampling method, with the dependent variable being the referral source (hospital and non-hospital) and independent variables including the number of Antenatal Care (ANC) visits, BMI, gravidity, residence, occupation, and education. Data analysis was performed using multiple logistic regression with SPSS Software.

Results: Among the 280 patients studied, the majority were referred from hospitals (196 patients), and most had abnormal BMI (193 patients). Logistic regression analysis showed that women from rural areas had significantly higher delays in referrals compared to urban residents (p=0.004), while other variables such as age, number of ANC visits, BMI, gravidity, occupation, and education did not show significant influence.

Conclusion: Place of origin is a critical factor influencing referral outcomes, highlighting the significant role of geographical and socio-economic determinants in the accessibility and quality of maternal healthcare services. This study underscores the necessity of an effective and timely referral system to mitigate maternal mortality rates in Indonesia, particularly emphasizing the need for enhanced referral infrastructure in rural areas to ensure prompt access to maternal care and ultimately reduce maternal mortality.

Summary: Indonesia's maternal mortality rate remains high (305 per 100,000 live births), influenced by prolonged referral processes. Secondary data from 280 medical records at Dr. Soetomo General Hospital, Surabaya (July–September 2023), revealed that most patients were referred from hospitals (196), predominantly exhibiting abnormal BMI (193). Logistic regression analysis identified residential location (rural vs. urban) as significantly impacting referral outcomes (p=0.004), whereas other factors (age, ANC visits, BMI, gravidity, occupation, education) showed no significant influence. These findings highlight the critical role of geographic and socioeconomic conditions in maternal healthcare accessibility and emphasise the importance of an efficient and timely referral system to reduce maternal mortality in Indonesia.

KEYWORDS:
Tertiary Care Centers, Maternal Mortality, Antenatal Care, Socioeconomic Factors, Maternal Health Services

INTRODUCTION

Maternal mortality remains a critical issue both in Indonesia and globally. Despite the World Health Organization's 2017 estimate of 177 maternal deaths per 100,000 live births[1], Indonesia's national data reports a significantly higher maternal mortality rate of 305 deaths per 100,000 live births in 2023. Furthermore, the country has the second highest neonatal mortality rate in Southeast Asia, with approximately 32 neonatal deaths per 1,000 live births.[2] The target is to reduce this number to 74 per 100,000 live births by 2025.[3] There is a significant disparity in maternal mortality rates between developed and developing countries. In developed countries, the rate is much lower, at 12 per 100,000 live births, but in Sub-Saharan Africa, it can reach 546 per 100,000 live births.[4]

Various factors contribute to maternal mortality, such as prolonged referral processes, delays in decision-making, delays in arrival at healthcare facilities, and delays in receiving medical assistance.[5,6] Research has shown that delayed referrals can lead to poor outcomes for both mothers and newborns, highlighting the importance of a timely and efficient referral system in preventing maternal deaths.[7] Addressing delays in the referral process, such as triage

This article was accepted: 21 April 2025
Corresponding Author: Rizki Pranadyan
Email: rizki-p@fk.unair.ac.id

delays and inadequate access to healthcare facilities, it is crucial in reducing maternal mortality and improving maternal healthcare outcomes.[8] Therefore, an assessment of the factors related to referrals is necessary.

MATERIALS AND METHODS

This study is a descriptive analytical research using secondary data from medical records of patients admitted to the Obstetrics and Gynecology Maternity Room at Dr. Soetomo Hospital, Surabaya. The data collection period was from July to September 2023 with ethical clearance number 2813/104/4/III/2024 after thorough considerations by specialists, residents, and other healthcare professionals. Patients with incomplete medical records or patients transferred from tertiary hospitals were excluded. The sampling method used was total sampling.

The dependent variable in this study is the source of patient referral, categorized into hospital and non-hospital group. Non-hospitals includes Self-Administered Patients, Independent Practicing Midwives, Clinic, and Community Health Centers. Hospital referral is from secondary hospital. The independent variables are as follows: the number of ANC visits (based on the Ministry of Health of the Republic of Indonesia, 2020 [9], categorized as normal if ≥ 6 times and abnormal if < 6 times), BMI (based on the Ministry of Health's P2PTM, 2018 (10), categorized as normal if 18.5 – 25 and abnormal if <18.5 or >25), gravidity (categorized as normal if ≤ 4 and abnormal if >4), residence (categorized into rural and urban), occupation (categorized into unemployed and employed), and education (categorized into out-of-school and in school).

Research data analyze in two method using SPSS software. First, the univariate level, the percentage distribution of the research sample is presented to show the distribution of respondents according to the characteristics mentioned above. Futher, Multiple logistic regression was performed to further test relationship between referral source and its independent variables. The results are presented as adjusted odds ratio (OR) estimates with 95% confidence intervals (CI).

RESULTS

The table shows the number of ANC (Antenatal Care) visits at various healthcare facilities in Indonesia, including Independent Practicing Midwives, Community Health Center, Clinic, Obstetrics and Gynecology Specialist Practices, and Hospital. The data indicate that the majority of ANC visits occur only once, primarily at community health center (189 visits), Hospitals (116 visits), and independent practicing midwives (76 visits). Clinics exhibit a more evenly distributed range of ANC visits. There are no cases where no ANC visits were reported. The highest total number of ANC visits was recorded at hospitals (284 visits), followed by Clinic (271 visits), and Community Health Center (255 visits).

The table also shows the number of patients based on their referral sources to healthcare facilities. Most patients were referred from hospitals (196 patients), followed by self-administered patients (75 patients). Referrals from community health center (4 patients), independent practicing midwives (3 patients), and clinic (2 patients) were relatively few. The total number of referred patients was 280.

The table describes patient characteristics across various categories. A total of 193 patients had an abnormal BMI, while 87 patients had a normal BMI. Regarding gravidity, 22 patients had an abnormal gravidity condition, and 258 patients had a normal gravidity condition. Based on the patients residence, 126 patients were from rural areas and 154 patients were from urban areas. In terms of occupation, the majority of patients, 225 in total, were unemployed, while 55 patients were employed. Regarding education, 85 patients had out-of-school, while 195 patients had in school. Overall, the majority of patients had an abnormal BMI, normal gravidity, were from urban areas, were unemployed, and had formal education.

The results of the logistic regression analysis show the influence of various variables on the observed outcomes. The variables included in the model are age category (age), ANC category (ANC), Body Mass Index category (BMI), gravidity category (gravidity), place of residence (residence), employment category (occupation), and education category (education).

The analysis results indicate that the age category (age), ANC category (ANC), Body Mass Index category (BMI), gravidity category (gravidity), employment category (occupation), and education category (education) are not significant predictors of the outcome. However, one variable stands out: place of residence, which has a significant influence on the outcome. Rural residents were 58% more likely to experience referral delays compared to urban residents (OR 0.42, 95% CI: 0.23-0.76, p=0.004).

Additionally, the constant variable also shows a significant influence on the outcomes, indicating geographical or socio-economic factors of the patient play an important role in maternal outcomes.

DISCUSSION

Based on the research findings, the variables of age category, number of antenatal care (ANC) visits, body mass index (BMI), gravidity, and employment were not significant. The lack of significance for age category may be attributable to the homogeneity of the study population or the presence of unmeasured confounding variables, such as healthcare provider availability and emergency transport accessibility, which could have influenced the outcomes. Similarly, the number of ANC visits might show no significant effect due to the inconsistent quality of services provided in different healthcare settings, creating limited variation in measurable outcomes. Furthermore, BMI, gravidity, and employment could be overshadowed by other unmeasured factors.

Nonetheless, place of residence demonstrated a significant influence, suggesting that socio-economic or geographical aspects of a patient's living environment may play a pivotal role. These findings suggest that residence-based disparities—such as differing accessibility to healthcare services, environmental conditions, and cultural practices—may contribute to variations in maternal health behaviors and outcomes.

Table I: Characteristics of antenatal visits

Number of ANC	Independent Practicing Midwives	Community Health Center	Clinic	Obstetrics and Gynecology Specialist Practices	Hospitals
0	0	0	0	0	0
1	76	189	30	104	116
2	18	26	16	8	26
3	18	12	30	12	33
4	20	4	36	16	16
5	5	5	45	10	25
6	-	12	36	-	24
7	7	7	21	7	35
8	-	-	16	-	-
9	-	-	18	-	9
10	10	-	10	-	-
13	-	-	13	-	-
Total	154	255	271	157	284

Table II: Distribution of Patients Based on Referral Sources

Referrer	Number of Patients
Independent Practicing Midwives	3
Self-Administered Patients	75
Clinic	2
Community Health Center	4
Hospital	196
Total	280

Table III: Demographic and Clinical Characteristics of Patients

Characteristics of Patients	Number of Patients
BMI Classification	
Abnormal	193
Normal	87
Gravidity	
Abnormal	22
Normal	258
Residence	
Rural	126
Urban	154
Occupation	
Unemployed	225
Employed	55
Education	
Out-of-School	85
In School	195

Table IV: Logistic Regression Analysis Results

Variable		p-value	OR	95% CI	
				Lower	Upper
Step 1a	Age	0,123	0,596	0,309	1,150
	Antenatal Care	0,242	0,696	0,380	1,276
	Body Mass Index	0,085	0,613	0,350	1,071
	Gravidity	0,537	1,394	0,487	3,993
	Residence	0,004	0,419	0,232	0,757
	Occupation	0,294	0,705	0,367	1,353
	Education	0,783	1,094	0,578	2,073
	Constant	0,002	5,529	-	-

a. Variable(s) entered on step 1: age, antenatal care, body mass index, gravidity, residence, occupation, education.

An effective referral system remains critical for maternal healthcare because it ensures timely access to specialized services, reduces maternal mortality rates, and supports the reduction of stillbirths and neonatal deaths.[8,11] Improved referral mechanisms also bolster maternal outcomes through adequate communication systems and standardized operational guidelines [12–14], along with the integration of emergency obstetric and neonatal care.[15–17] However, these benefits can be undermined by factors such as transportation limitations and broader socio-economic inequalities, which often impair referral efficiency.[13,18]

Similar access disparities have been observed in healthcare contexts such as kidney transplantation and genetic counseling, where socio-economic variables and physician referral patterns contribute to unequal service utilization.[19,20] Consequently, mitigating these disparities may include targeted educational interventions, quality improvements, and the adoption of electronic referral systems.[21–23]

Understanding how health access disparities affect maternal mortality at tertiary healthcare facilities requires consideration of multiple elements, including distance to care, availability of skilled midwives, and the comprehensiveness of maternal healthcare services.[24,25] Geographic barriers and shortages of trained providers can severely impede the management of obstetric emergencies.[26,27)]

In Indonesia, these obstacles can be categorized into modifiable and unmodifiable factors: the former encompasses inadequate healthcare infrastructures, limited access to skilled professionals, and insufficient emergency transportation systems, while the latter covers distinct cultural and socio-economic conditions that affect the uptake of modern healthcare services. Such factors may manifest in varying attitudes toward antenatal care (ANC), childbirth practices, and acceptance of medical interventions, further highlighting the role of culture and tradition in shaping maternal health-seeking behaviors.

To address both sets of barriers holistically, enhancing the quality and standardization of ANC services across facilities is an essential first step. Evidence-based strategies to reduce maternal mortality should include strengthening healthcare infrastructure—particularly through the implementation of mobile health (mHealth) referral tracking systems, which would improve the coordination of care and ensure timely referrals. Additionally, emergency transportation subsidies for rural women can mitigate the impact of geographical barriers and ensure more equitable access to essential obstetric care. The integration of private midwife clinics into existing referral networks would further enhance the reach of skilled providers and ensure continuity of care, particularly in underserved areas.

Moreover, continuous training and education for healthcare providers, particularly midwives, are vital for ensuring consistent service quality and fostering more efficient care pathways. To overcome social and cultural barriers, community-based education and culturally sensitive initiatives are necessary to encourage positive health-seeking behaviors and bolster the utilization of maternal health services. Implementing these measures comprehensively can help mitigate the combined effects of geographical, socio-economic, and cultural factors, ultimately reducing maternal mortality and improving pregnancy outcomes across diverse local contexts in Indonesia.

CONCLUSION

This study aimed to identify factors influencing the tiered referral of pregnant women to tertiary healthcare facilities. The findings indicated that place of residence was the only variable significantly associated with referral outcomes, highlighting the critical impact of geographic and socio-economic disparities on maternal healthcare accessibility. Conversely, other examined factors—including maternal age, frequency of antenatal care (ANC) visits, body mass index (BMI), gravidity, employment status, and educational background—were not significantly associated with referral outcomes, possibly due to variations in ANC quality or the influence of unmeasured confounding variables.

Given these findings, interventions should focus primarily on strengthening healthcare infrastructure, particularly through enhanced transportation systems, effective referral protocols, and standardized ANC service delivery. Additionally, culturally sensitive community-based educational programs are essential to address social and cultural barriers affecting maternal healthcare utilization. Further research is warranted to validate the proposed interventions, including intervention-based trials such as mHealth referral monitoring and community-based maternal transport systems, which have the potential to improve maternal health outcomes. Additionally, exploring additional context-specific determinants is critical for developing comprehensive strategies to optimize referral pathways and mitigate maternal mortality rates, particularly in rural Indonesia.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to all healthcare providers, staff, and administrative personnel who contributed their valuable time, effort, and support throughout the data collection process. We also extend our appreciation to the collaborating colleagues and institutions whose cooperation and contributions have been integral to the successful completion of this research.

REFERENCES

1. WHO. Trends in maternal mortality 2000 to 2017: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division: executive summary. World Health Organization; 2019.
2. Kurniawati W, Afiyanti Y, Akmelia A, Santi NF. Family support and peer support related to the physical activity of the prospective bride and groom. Gac Med Caracas 2022; 130(Supl 5): S1169-74.
3. Giena VP, Sari RM, Sasmita FN, Rahmawati I. Developing a Mobile Application as a Learning Medium on Maternal Emergencies Related to Postpartum Hemorrhage: 10.55131/jphd/2022/200104. J Public Heal Dev. 2022 Jan 29; 20(1): 38-50.

4. Alkema L, Chou D, Hogan D, Zhang S, Moller AB, Gemmill A, et al. Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: A systematic analysis by the un Maternal Mortality Estimation Inter-Agency Group. Lancet 2016; 387(10017): 462-74.
5. Siswosudarmo R. Effect of Delay in Postpartum Hemorrhage Management on the Rate of Near-Miss and Maternal Death Cases. Indones J Obstet Gynecol 2018; 2(4): 177.
6. Gustiani R, Hakimi M, Suryaningsih EK. Implementation of Referral System in Postpartum Hemorrhage Cases By Midwife. Int J Heal Med Sci 2022; 5(3): 211-20.
7. Owen M, Batakji M, Goodman D, Kim SM, Olufolabi A, Srofen E. Measuring delay in high-risk obstetric referrals in Accra, Ghana: How long does it really take? Research Square; 2020.
8. Bugomola MA, Leshabari S, Mkoka D, Mbekenga CK. Effectiveness of Maternal Referral System in Tanzania: A Mixed Method Study 2022; 1-18.
9. Kemenkes RI. Pedoman Pelayanan Antenatal Terpadu 2020 Ed.3 [Internet]. Vol. III, Kementerian Kesehatan Republik Indonesia. 2020. 1-86 p. Available from: https://repository.kemkes.go.id/book/147
10. P2PTM Kemenkes RI. Kementerian Kesehatan Republik Indonesia. 2018 [cited 2024 May 15]. Klasifikasi obesitas setelah pengukuran IMT. Available from: https://p2ptm.kemkes.go.id/infographic-p2ptm/obesitas/klasifikasi-obesitas-setelah-pengukuran-imt
11. Matsui M, Tsukada M, Masaki Y, Tung R, Iwamoto A. Over-utilisation of maternity in a tertiary referral hospital in Cambodia – a descriptive study with implications for equity in access to health services and official development assistance. 2020; 1-18.
12. Atuoye KN, Dixon J, Rishworth A, Galaa SZ, Boamah SA, Luginaah I. Can she make it? Transportation barriers to accessing maternal and child health care services in rural Ghana. BMC Health Serv Res 2015; 15: 333.
13. Nuamah GB, Agyei-Baffour P, Akohene KM, Boateng D, Dobin D, Addai-Donkor K. Incentives to yield to Obstetric Referrals in deprived areas of Amansie West district in the Ashanti Region, Ghana. Int J Equity Health 2016; 15(1): 117.
14. Jakhar R, Choudhary A. Study of maternal outcome in referral obstetric cases in a tertiary care centre. J Fam Med Prim care 2019; 8(9): 2814–9.
15. Dewi A, Sundari S, Nursetiawan N, Supriyatiningsih S, Sugiyo D, Dewi DTK, et al. Reducing maternal mortality: A qualitative study of health worker's expectation in urban area, indonesia. Open Access Maced J Med Sci 2021; 9: 18-26.
16. Akaba GO, Ekele BA. Maternal and fetal outcomes of emergency obstetric referrals to a Nigerian teaching hospital. Trop Doct. 2018; 48(2): 132-5.
17. Singh S, Doyle P, Campbell OM, Mathew M, Murthy GVS. Referrals between public sector health institutions for women with obstetric high risk, complications, or emergencies in India - A systematic review. PLoS One. 2016; 11(8): 1-23.
18. Hussein J, Kanguru L, Astin M, Munjanja S. The effectiveness of emergency obstetric referral interventions in developing country settings: a systematic review. PLoS Med 2012; 9(7): e1001264.
19. Cragun D, Bonner D, Kim J, Akbari MR, Narod SA, Gomez-Fuego A, et al. Factors associated with genetic counseling and BRCA testing in a population-based sample of young Black women with breast cancer. Breast Cancer Res Treat 2015; 151(1): 169-76.
20. Gander JC, Zhang X, Plantinga L, Paul S, Basu M, Pastan SO, et al. Racial disparities in preemptive referral for kidney transplantation in Georgia. Clin Transplant 2018; 32(9): e13380.
21. Alnajar A, Kareff SA, Razi SS, Rao JS, De Lima Lopes G, Nguyen DM, et al. Disparities in Survival Due to Social Determinants of Health and Access to Treatment in US Patients with Operable Malignant Pleural Mesothelioma. JAMA Netw Open 2023; 6(3): E234261.
22. Harding JL, Perez A, Patzer RE. Nonmedical barriers to early steps in kidney transplantation among underrepresented groups in the United States. Curr Opin Organ Transplant 2021; 26(5): 501-7.
23. Bettano A, Land T, Byrd A, Svencer S, Nasuti L. Using Electronic Referrals to Address Health Disparities and Improve Blood Pressure Control. Prev Chronic Dis 2019; 16: E114.
24. Okonofua F, Ogu R, Agholor K, Okike O, Abdus-Salam R, Gana M, et al. Qualitative assessment of women's satisfaction with maternal health care in referral hospitals in Nigeria. Reprod Health 2017; 14(1): 44.
25. Kapula N, Shiboski S, Dehlendorf C, Ouma L, Afulani PA. Examining socioeconomic status disparities in facility-based childbirth in Kenya: role of perceived need, accessibility, and quality of care. BMC Pregnancy Childbirth 2022; 22(1) :804.
26. Haruzivishe, Clara Opha; Ellen, Chirwa; Margaret, Maimbolwa; Alfred, Maluwa; Patricia, Mukwato-Katowa; Janet, Mambulusa; Concepta, Kwaleyela; Exnevia, Gomo; Midion C. Quality of Antenatal, Labour and Delivery Services in Selected Districts in Malawi, Zambia and Zimbabwe: A Descriptive Cross-Sectional Study. J Obstet Gynecol Surg 2021; 2(1): 1-9.
27. Jombo SE, Enabudoso E, Njoku A, Afekhobe J. Reducing Maternal Mortality: an Assessment of the Availability and Quality of Emergency Obstetric and Newborn Care in Esan Central Local Government Area of Edo State. Research Square; 2019.

COMMENTARY

Artificial intelligence in healthcare: A call for strategic implementation to maximize impact and minimize costs

Keng Sheng Chew, PhD

Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

ABSTRACT

Artificial Intelligence (AI) is transforming healthcare by improving diagnostics, decision-making, and patient outcomes. This commentary emphasizes the need for strategic AI implementation to maximize benefits while minimizing costs. Although AI can automate routine tasks and freeing clinicians to focus more on patient care, challenges like costs and medicolegal concerns must be addressed. A two-dimensional impact-effort framework is proposed to prioritize AI tools implementation based on their impact and costs.

KEYWORDS:

Artificial intelligence in healthcare, AI implementation strategy, Asplin conceptual model, medico-legal challenges, impact-effort matrix

INTRODUCTION

Artificial Intelligence (AI) was first conceptualized by John McCarthy in 1956 as an "electric brain," capable of executing tasks typically associated with human intelligence, including speech recognition, visual perception, learning, and decision-making.[1] A fundamental question in AI development has been: "Can a machine think?" This question was central to the work of Alan Turing, who in 1951 proposed a practical test for assessing machine intelligence, now popularly known as the Turing Test. According to Turing, a machine passes this test if a human, after asking a series of questions, cannot distinguish whether the responses come from a machine or from that of another human. To pass the Turing test, the machine must demonstrate three key capabilities: (1) understanding speech (natural language processing), (2) storing and recalling information (knowledge representation), and (3) using that information to reason, adapt, and learn (automated reasoning and machine learning).[1]

Indeed, the integration of AI into various fields has become so ubiquitous, that it is now likened as a "tsunami". Mustafa Suleyman, in his 2023 book The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma, vividly captures this phenomenon.[2] He describes a "wave" as the global diffusion or proliferation of this generation of new technologies that are of high demand across different industries at affordable cost.

In the field of medicine, the impact of this wave is particularly evident. Over the last decade, the number of publications on AI applications in healthcare has increased exponentially.[3] This surge reflects the growing recognition of the potentials of AI to revolutionize medical practice, from diagnostics and treatment planning to patient care and workflow optimization. To illustrate how AI implementation can be mapped strategically in a real-world clinical setting, a case study from emergency medicine is presented below.

CASE STUDY: THE IMPACT OF AI IN EMERGENCY MEDICINE

In emergency medicine for example, AI applications have shown promising results. A systematic review by Piliuk and Tomforde[4] categorizes AI applications into two broad categories, i.e., (1) diagnostics-specific and (2) triage-specific applications. Diagnostics-specific AI tool such as tools for predicting diseases like stroke, heart disease, and sepsis can enhance decision support by interpreting medical images and test results with greater accuracy and reliability. Triage-specific AI applications, on the other hand, focus on predicting patient outcomes, including mortality risk, which helps prioritize high-risk patients.

Given the wide array of AI tools available, systematically categorizing these tools to facilitate their integration into emergency department (ED) workflows is essential. One useful approach is perhaps the Asplin's input-throughput-output conceptual model.[5] This framework categorizes ED processes into three stages: input, which addresses the demand for ED services; throughput, which involves the processes throughout a patient's stay in the ED; and output, which refers to the processes related to patient disposition, including admission, discharge, or continued care within the ED. Table I provides an example of mapping some AI applications according to the Asplin's conceptual model.

REHUMANIZING HEALTHCARE WITH AI

One of the most intriguing aspects of AI in healthcare is its potential to rehumanize healthcare. This may seem paradoxical, given that AI is often perceived as a depersonalizing force. However, by automating routine tasks such as appointment scheduling and follow-up reminders, AI can free up time for more meaningful interactions with patients. Another example is the use of machine learning models known as the ambient AI scribes to transcribe clinical conversations in real time using speech-to-text technology.[6] These transcripts are then processed by large language models (LLMs) to generate structured summaries (such as the

This article was accepted: 28 April 2025
Corresponding Author: Keng Sheng Chew
Email: kschew@unimas.my

Table I: Some of the potential AI applications as described in previous studies[4,5] mapped according to Asplin's conceptual framework

Input	Throughput	Output
Prehospital Urgent Stroke Triage Score Using Machine Learning	Stroke alert at triage	ICU admission and in-hospital death of trauma patients
ROSC Identification for OHCA	Prediction of in-hospital cardiac arrest	Prediction of mortality following TBI
Triaging of undifferentiated patients	Classification of TBI into mild, moderate and severe according to GCS & metabolic stress profiles	Predicting the Need for Hospitalization for Pediatric Asthma Exacerbation
Shock decision during load distributing device mechanical CPR	Differentiating COPD from congestive heart failure	Neurological outcome at 90-day after ROSC in OHCA Prediction of need for hospitalization for urgent care patients
Prediction of out-of-hospital cardiac arrest	Detecting pneumothorax from ECG	Prediction of need for emergency neurosurgery within 24 h after moderate to severe TBI
Pre-hospital prediction of sepsis	Identification of anaphylaxis	
E-triage using machine learning	Early diagnosis of hypovolemic in trauma	
Prediction of neurological outcome following OHCA	Chest X-ray Interpretation	
Crash scene data to identify adults with moderate and severe vehicular injuries	Early recognition of signs of shock	

ICU = intensive care unit; ROSC = return of spontaneous circulation; OHCA = out-of-hospital cardiac arrest; TBI = traumatic brain injuries; CPR = cardiopulmonary resuscitation; COPD = chronic obstructive pulmonary disease; ECG = electrocardiogram

Table II: Examples of Challenges in AI Implementation and Applications in Healthcare[10]

Types of challenges	Examples in the Malaysian Context
P = policy	Lack of national strategies or clear regulatory direction for AI implementation in the Malaysian healthcare setting Absence of AI-specific legislation in the Malaysian healthcare setting
E = economic	Insufficient investment in infrastructure to support AI deployment
S = socio-cultural	Public mistrust and low awareness of AI capabilities and limitations
T = technological	Low technological awareness and literacy among the Malaysian public Lack of technological integration and interoperability in the healthcare system
E = environmental	Lack of AI ethics and governance expertise within the institutional review boards A lot of bureaucratic red tape slowing down innovation and AI trial approvals
L = legal	No dedicated legal framework for AI-specific liability and malpractice Unclear responsibility when AI-generated decisions lead to harm

commonly used "Subjective, Objective, Assessment, Plan", or SOAP format).[7] In this regard, AI scribes can significantly ease the documentation burden for clinicians. By eliminating this burden, AI scribes enable clinicians to focus more fully on their patients and families in fostering a more compassionate and attentive human communication. Indeed, as Eric Topol[8], in his book *Deep Medicine: How Artificial Intelligence can Make Healthcare Human Again,* suggests, one of most significant contributions of AI tools to healthcare could in fact, be the gift of time, i.e., time that could be spent on patient engagement, education, understanding, and empathy. Topol quoted several past studies that have shown that even a slight increase in the time clinicians spend with patients can lead to better outcomes. For instance, by adding just one additional minute to a home visit, the risk of readmission could be reduced by 8%, and by spending more time with patients, hospitalization needs could be reduced by 20%.[8] Topol also referenced a report by the Institute for Public Policy Research titled "Better Health and Care for All" which suggests that should the potential of AI to automate healthcare processes can be fully realized, this could lead to productivity improvements valued at £12.5 billion annually, or approximately 9.9% of the NHS budget in England.[9]

STRATEGIC IMPLEMENTATION OF AI: THE IMPACT-EFFORT MATRIX APPROACH

Unfortunately, with the multitude of AI tools available, selecting the right ones for initial implementation can be daunting. In this regard, a practical approach to navigate through this maze is to evaluate each of these AI tools using a two-dimensional impact-effort matrix. In this regard, "impact" refers to the degree of automation that could be achieved, while "effort" could refer to the cost of implementation (Figure 1). By using this matrix, those tools that can yield a high degree of automation at a lower cost can be prioritized for implementation.

While this commentary uses the ED as a case study, similar systematic categorization of clinical flow processes may be adopted for other clinical specialties. For example, in surgical disciplines, AI applications could be categorized along the continuum of pre-operative, intra-operative, and post-operative phases. In primary care setting, alternative categorization may involve screening, diagnosis, treatment, and follow-up.

Although the potential of AI in healthcare is vast, its implementation and application face many challenges. The

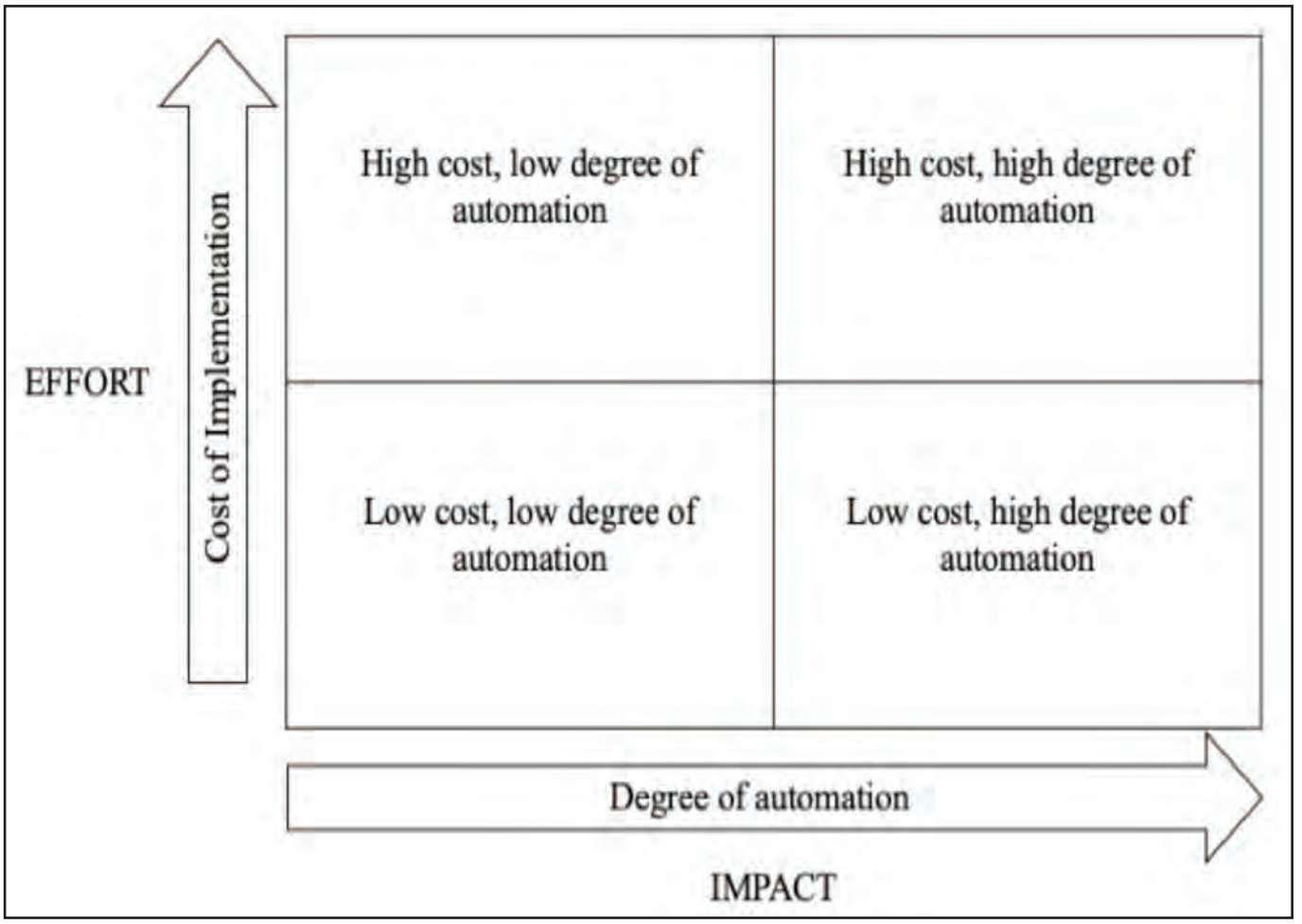


Fig. 1: Impact-Effort Matrix for AI Implementation in Clinical Practices

PESTEL framework, where P = Policy, E = Economic, S = Socio-cultural, T = Technological, E = Environmental, and L = Legal dimensions, offers a structured way to examine these challenges. Some specific examples of these challenges within the Malaysian context, mapped and adapted from Phang et al[10], are as outlined in Table II. Interestingly, Vearrier et al.[11] have highlighted another potential medico-legal concern in near future when AI tools become a standard clinical practice. Under the current Bolam test, clinicians may be held liable for not meeting the expected standard. But in near future, the use of these AI tools may redefine what constitutes acceptable care. As a result, the failure to adopt AI when it becomes the norm could be viewed as substandard care.

To support effective AI integration in healthcare, several policy actions are recommended. These include developing a national framework to guide ethical and legal standards, investing in infrastructure and digital training for healthcare staff, and establishing multidisciplinary regulatory bodies to oversee AI governance. Policymakers should also support pilot testing of AI tools in clinical settings to evaluate its impact and its cost of implementation, as well as to promote public engagement to build trust. These steps can help create a supportive environment for safe and meaningful AI adoption.

CONCLUSION

In conclusion, AI is poised to play an increasingly prominent role in healthcare, offering tools that can enhance diagnostic accuracy and optimize patient management. However, the successful integration of AI into clinical practices requires a careful balance between leveraging technology and preserving the human touch in patient care. The medico-legal landscape will similarly need to evolve in tandem with AI integration. As Hippocrates wisely observed, "It is more important to know what sort of person has a disease than to know what sort of disease a person has." AI has the potential to provide the additional time for clinicians to know the person with the disease rather than the disease of the person by automating processes to reduce wastages and inefficiencies. In the current setting of escalating healthcare cost, patient overload, and workforce burnout, AI offers a means to enhance patient-doctor communication and foster deeper and more meaningful engagement. By doing so, AI can fulfil its potential to make healthcare not only more efficient but also more humane again.

ACKNOWLEDGEMENT

The author acknowledges Universiti Malaysia Sarawak and the Scholarship of Teaching and Learning Research Grant (SoTL) grant (no: SoTL/FPSK/2023(22)/022) for the support rendered of this publication.

REFERENCES

1. Howard J. Artificial intelligence: Implications for the future of work. Am J Ind Med 2019; 62(11): 917-26.
2. Suleyman M. The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma. London: Bodley Head; 2023.
3. Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. npj Digital Medicine 2020; 3(1): 126.
4. Piliuk K, Tomforde S. Artificial intelligence in emergency medicine. A systematic literature review. Int J Med Inform 2023; 180: 105274.
5. Asplin BR, Magid DJ, Rhodes KV, Solberg LI, Lurie N, Camargo CA, Jr. A conceptual model of emergency department crowding. Ann Emerg Med 2003; 42(2): 173-80.
6. Tierney AA, Gayre G, Hoberman B, Mattern B, Ballesca M, Kipnis P, et al. Ambient Artificial Intelligence Scribes to Alleviate the Burden of Clinical Documentation. NEJM Catalyst 2024; 5(3): CAT.23.0404.
7. Mess SA, Mackey AJ, Yarowsky DE. Artificial Intelligence Scribe and Large Language Model Technology in Healthcare Documentation: Advantages, Limitations, and Recommendations. Plastic and reconstructive surgery Global open 2025; 13(1): e6450.
8. Topol EJ. Deep Medicine: How Artificial Intelligence can Make Healthcare Human Again. New York: Basic Books; 2019.
9. Darzi A. Better health and care for all: A 10-point plan for the 2020s. 2018. Available from: https://www.ippr.org/articles/better-health-and-care-for-all.
10. Phang KC, Ng TC, Singh SKG, Voo TC, Alvis WA. Navigating Artificial Intelligence in Malaysian Healthcare: Research Developments, Ethical Dilemmas, and Governance Strategies. Asian Bioethics Review 2024.
11. Vearrier L, Derse AR, Basford JB, Larkin GL, Moskop JC. Artificial Intelligence in Emergency Medicine: Benefits, Risks, and Recommendations. J Emerg Med 2022; 62(4): 492-9.

ERRATUM

Reference: Lim LJ, Kit AMF, Wong PS. Clinical outcomes of ovarian stimulation with follitropin delta in a mixed regimen with HP-hMG: a real-world retrospective analysise. Med J Malaysia. 2024 May; 79(3): 275-280. PMID: 38817059.

Table II: Demographics and baseline characteristics

Characteristic	
Total number of patients	20
Age (years)	35.2 ± 4.47
Women < 35 years (n, %)	9 (45.0%)
Women 35-40 years (n, %)	8 (40.0%)
Women > 40 years (n, %)	3 (15%)
Race:	
Malay (n, %)	2 (10.0%)
Chinese (n, %)	16 (80.0%)
Indian (n, %)	1 (5.0%)
Others (n, %)	1 (5.0%)
Body weight (kg)	58.3 ± 8.22
BMI (kg/m^2)	23.4 ± 3.67
AMH (pmol/L)	26.0 ± 14.71
AMH (pmol/L)	20.1 (16.0 - 32.7)
AFC - for both ovaries (n)	13.0 ± 4.35
Infertility history	
Duration of infertility (mo)	54.0 ± 32.31
Primary infertility	11/20 (55.0%)
Primary etiology:	
Tubal infertility (%)	3/20 (15.0%)
Male infertility (%)	3/20 (15.0%)
Unexplained (%)	3/20 (15.0%)
PCOS (%)	5/20 (25.0%)
Endometriosis (%)	1/20 (5.0%)
Others (%)	4/20 (20.0%)
Smoking (n, %)	1/20 (5.0%)
Number of first IVF/ICSI cycles	16/20 (80.0%)
Number of repeat IVF/ICSI cycles (non-naïve)	4/20 (20.0%)

Values are mean ± SD, median (interquartile range), or number (percentage), unless stated otherwise.

Table III: Ovarian response and pregnancy outcomes

Total patients	20
Duration of stimulation (days)	11.0 ± 1.16
Average daily dose of follitropin delta (mcg)	9.0 ± 2.50
Average total dose of follitropin delta (mcg)	96.6 ± 28.18
Women dosed according to follitropin delta algorithm	15/20 (75%)
Average daily dose of HP-hMG (IU)	65.2 ± 10.94
Average total dose of HP-hMG (IU)	714.3 ± 137.52
Percentage starting HP-hMG from D1 of OS	19/20 (95.0%)
No of cancelled cycles (n, %)	0/20 (0%)
Triggering of final oocytes maturation:	
hCG	11/20 (52.4%)
GnRHa	10/20 (47.6%)
No of oocytes retrieved (n)	13.2 ± 6.43
Poor responders (< 4 oocytes) (n, %)	0/20 (0%)
Excessive responders (≥ 20 oocytes) (n, %)	3/20 (15.0%)
Target ovarian response (8-14 oocytes) (n, %)	14/20 (70.0%)
No of MII oocytes (n)	10.8 ± 5.23
Type of fertilization:	
IVF	0
ICSI	21
Fertilization rate (%)	67.9 ± 19.93
Blastulation rate (%)	62.6 ± 25.42
Blastocysts:	
Total (n)	5.3 ± 3.52
Top quality (n)	2.4 ± 1.75
≥ 2 cryopreserved blastocysts per cycle start (n, %)	18/20 (90.0%)
No of patients who have undergone embryo transfers (n):	10
Fresh	0
Frozen	10
Average number of embryos per-transfer (n)	1.1 ± 0.32
Implantation rate (%)	72.7%
Clinical pregnancy rate per transferred cycle (n, %)	7/10 (70.0%)
OHSS - any grade (n)	0/20 (0%)

Values are mean ± SD, median (interquartile range), or number (percentage), unless stated otherwise.

Acknowledgement
May Issue 2025

The Editorial Board of The Medical Journal of Malaysia gratefully acknowledge the following individuals for reviewing the papers submitted for publication:

1. Dr Azlan Husin
2. Dr Evelyn Tai
3. Dr Gagandeep Mann
4. Dr Hayani Abdul Wahid
5. Dr Iskandar Firzada Osman
6. Dr James Khaw
7. Dr Jeyasakthy Saniasiaya
8. Dr Jun Kit Koong
9. Dr Kavita Jetly
10. Dr Krishna Kumar G
11. Dr Ling Yien Hii
12. Dr Mazlin Baseri
13. Dr Md Mizanur Rahman
14. Dr Mohan Arumugam
15. Dr Nor Anita Che Omar
16. Dr Norfarizan Azmi
17. Dr Norlelawati binti Ab Latip
18. Dr Qi Zhe Ngoo
19. Dr Riyadh Shiltagh Alrudaini
20. Dr Rozita Abdul Malik
21. Dr Salwana Abu Bakar
22. Dr Sanihah Abdul Halim
23. Dr Sharif Ahmad
24. Dr Subasri Armon
25. Dr Sufian Adenan
26. Dr Suresh Subramaniam
27. Dr Syadwa Abdul Shukor
28. Dr Thameen Ansari
29. Dr Tim Aung
30. Dr Vasu Pillai Letchumanan
31. Dr Wan Hazabbah Wan Hitam
32. Dr Zainab Abdul Majeed