ANNALS

OFFICIAL JOURNAL OF THE ACADEMY OF MEDICINE, SINGAPORE

VOLUME 53
NUMBER 3
MAR 2024

Clinical performance of automated machine learning: A systematic review

A study reviewed clinical applications of automated machine learning, with exemplar use cases including identifying pathology on common imaging modalities. (See full article, p.187)

Illustration by Ngiam Li Yi

Also in this issue

Mitigating adverse social determinants of health in the vulnerable population: Insights from a home visitation programme (p.132)

Health practices, behaviours and quality of life of low-income preschoolers: A community-based cross-sectional comparison study in Singapore (p.142)

Singapore tuberculosis (TB) clinical management guidelines 2024: A modified Delphi adaptation of international guidelines for drug-susceptible TB infection and pulmonary disease (p.170)

ISSN 2972-4066

Annals, Academy of Medicine, Singapore
Volume 53 | Number 3 | March 2024

EDITORIALS

Asian media reporting on suicide: Concerning trends
Keith Harris 127

Bridging expertise with machine learning and automated machine learning in clinical medicine
Chien-Chang Lee, James Yeongjun Park, Wan-Ting Hsu 129

ORIGINAL ARTICLES

Mitigating adverse social determinants of health in the vulnerable population: Insights from a home visitation programme
Kevin Wei Hao Liang, Keith Wei Han Liang, Qing Rong Chan, Winnie Hwee Suat Goh 132

Health practices, behaviours and quality of life of low-income preschoolers: A community-based cross-sectional comparison study in Singapore
Shang Chee Chong, Ramkumar Aishworiya, Woon Li Seo, Yee Keow Chiong, Gerald Choon-Huat Koh, Jeremy Bingyuan Lin, Lynn Heng, Thahirah Habib Mohd, Young Ern Saw, Yiong Huak Chan, Jing Shi Chua, Shefaly Shorey 142

REVIEW ARTICLES

The impact of media reporting of suicides on subsequent suicides in Asia: A systematic review
PV Asharani, Yen Sin Koh, Rachel Hsiao Shen Tan, Yoke Boon Tan, Savita Gunasekaran, Benedict Lim, Lorainne Tudor Car, Mythily Subramaniam 152

Singapore tuberculosis (TB) clinical management guidelines 2024: A modified Delphi adaptation of international guidelines for drug-susceptible TB infection and pulmonary disease
Clinical Tuberculosis Guidelines Development Team:
Clinical Tuberculosis Guidelines Panel: Ang Michelle Lay Teng, Chan Si Min, Cheng Lionel Tim-Ee, Cheong Hau Yiang, Chew Ka Lip, Chlebicki Piotr Maciej, Hsu Li Yang, Kaw Gregory Jon Leng, Kee Adrian Chin Leong, Ng Mark Chung Wai, Ong Rick Twee Hee, Ong Catherine Wei Min, Quah Jessica Lishan, Selvamani D/O Balasubramaniam, Sng Li Hwei, Tan Jamie Bee Xian, Tan Cher Heng, Tay Jun Yang, Teo Lynette Li San, Thoon Koh Cheng, Yan Gabriel Zherong
Public Health Translational Team: Chen Jacinta I-Pei, Hud bin Mohammed Helmi, Khoo Benjamin Bing Jie, Lee Dawn Yi Xin, Ng Bob Xian Yi, Park Jia Ying, Tan Belinda Ying Ting, Yang Qian 170

Clinical performance of automated machine learning: A systematic review
Arun James Thirunavukarasu, Kabilan Elangovan, Laura Gutierrez, Refaat Hassan, Yong Li, Ting Fang Tan, Haoran Cheng, Zhen Ling Teo, Gilbert Lim, Daniel Shu Wei Ting 187

Annals, Academy of Medicine, Singapore
Volume 53 | Number 3 | March 2024

LETTERS TO THE EDITOR

Validating two international warfarin pharmacogenetic dosing algorithms for estimating the maintenance dose for patients in Singapore
Stephanie Pei Yun Soh, Wei Yann See Toh, Wei Qing Ten, Khai Pang Leong, Liuh Ling Goh 208

HLA-B*5801 testing: Is it time to consider mandatory testing prior to prescribing allopurinol in Singapore?
Felicia Li Ling Ong, Shiu Ming Pang 211

Patient's degree of adherence, challenges & preferences towards medicine taking (PACT) in Singapore
Vivien Teo, Benny Kaihui Li, Suat Fern Low, Jing Wen Oh, Kwok Foong Lam, Eng Sing Lee, Cheryl Wei Yan Tan, Poh Ching Tan, Hui Shan Chia, Su Yin Pang, Fu Xun Marcus Heng, Sin Hui Neo, Tat Ming Ng 213

Ann Acad Med Singap 2024;53:127-8
https://doi.org/10.47102/annals-acadmedsg.202458

EDITORIAL

Asian media reporting on suicide: Concerning trends

Keith M Harris[1] *PhD*

Asharani et al. present an enlightening study of media influences on suicidality and suicides from multinational data, all within Asia.[1] This is important, as knowledge based on media and suicide has been dominated by Western cultures, and English and other European languages. Pulling together various independent studies, as Asharani et al. have done, represents a long-overdue contribution to our understanding of how our Asian communities respond to reports of celebrity and other significant suicide events. In addition, these findings highlight the need for improved research methods and data collection, and making data public for meta-analyses and follow-up studies.

Around Asia, we have seen suicide copycat effects, such as from Lee Eun-ju's suicide.[4] Asharani et al.'s analysis revealed that these events share similar patterns around Asia, and associations between media coverage of suicide and suicide rates are also similar to reviews from decades ago in Western media.[7] We see indications that younger people, who may identify with a particular celebrity, may be prompted via media reporting towards increased suicidality and behaviours. When our institution examined newspaper reporting of suicides in Cambodia,[2] we found numerous examples of reports on "how to commit suicide", that is, explicit methods used by the suicide victim. This is an example of reporting that is strongly recommended against.[10] Study findings indicating specific reporting issues (e.g. glamorisation of suicide) generally lead to increased pressures on mass media organisations to adhere to suicide media reporting standards. Media guidelines are necessary, and not a new idea within the Asian region.[8] However, many media guideline recommendations are not strongly evidence-based, and there is always a need for reanalysis and revision.

Suicide reporting guidelines cannot be effective at reducing suicide rates if they are not firmly rooted in reality, which can only be determined through rigorous research. Unfortunately, this new meta-analysis also identified a considerable need for improving the quality of studies by reducing biases and improving research methods. Several improvements in research methods are imperative to ascertain the true effects of media reporting on suicide within the already complex set of known suicide risk and protective factors. Improvements should include better categorisation and measurement of study variables. For example, several studies use a simple yes-or-no variable and other binary-type variables when in fact, the true factor comprises a spectrum. A media report could be described as either a "yes" or "no", on whether it provides information on the suicide method. However, a "yes" answer could include an ordinal scale describing a news report as "moderately" or "substantially" providing information on suicide methods. Such detailed information can allow for more advanced and precise statistical analyses.

Similar to the yes-or-no approach used for many suicide risk and protective factors, scarce attention has been given to describing gender, sexuality and language minorities. If studies do not include information on non-binary gender identities, sexual identities and native language, we cannot determine if such factors are important or not in the reporting of copycat suicide.

Further problems with the literature that make these meta-analyses challenging, lie with inadequate treatment of the data, sometimes resulting in garbage-in garbage-out situations. Pre-registering study methods and analyses, and providing open access to data could help reduce current issues with poor missing data treatment, lack of data cleansing, insufficient statistical analyses and lack of method details. The current meta-analysis provides a wealth of data and should serve as an inspiration to future researchers. Open and robust methods will help us solidify the findings of these studies and lead to much more effective media reporting guidelines. Another area of concern in suicide reporting is the lack of public data and lack of data sharing. Some improvements must come from the public sector, as some governments in Asia still provide suicide data of low quality.[9] Researchers may not be able to impact government procedures, but they can do a great deal more in terms of making data publicly available, encouraging collaboration and developing large databases for future analyses.

[1] School of Psychology, Charles Sturt University, Port Macquarie, New South Wales, Australia
Correspondence: Dr Keith M Harris, Senior Lecturer, School of Psychology, Charles Sturt University, Locked Bag 5000, Port Macquarie, New South Wales, Australia 2444.
Email: esensei@yahoo.com, kmh.psyc@gmail.com

Our institution found great value in making data and methods public for developing the Suicidality Scale in English, Chinese and Spanish.[3]

When looking at suicide studies that focus on external factors (e.g. media reports, social disruption and economy variations), we should not forget the individual. Suicidality is extremely complex but always comes down to the individual and their personal factors that might contribute towards increased or decreased suicide risk. We need to carefully look at the true reasons for an individual's suicide, not only at simple bivariate correlations. Other factors are likely to have more meaningful effects on suicidal behaviours, even within the context of media reporting of celebrity suicides.[5] It is not only demographic factors (e.g. age, sex and ethnicity) that will contribute to increased risk of copycat suicide; personality and mental health factors will also play a much stronger role. Future work can consider integrating online media, including social media, into these research databases. Data mining and other approaches have shown that analyses of depression and suicide-related conversations on social media can provide unique data in suicide research.[6]

When thinking of copycat suicides, the underlying causes of suicidality and vulnerability may be the idolisation of celebrities and strong self-identity with celebrities. Therefore, it is not so much the media reporting of a celebrity's suicide, but the vulnerable recipient's learning of their idol's suicide that is the "real" reason for increased suicidality. If true, that might mean some suicide reporting guidelines may be overreaching. A better approach to suicide prevention could be to address youth and adult over-infatuation with celebrities. Clinicians might be wary of clients who are over-infatuated with celebrities. Perhaps, there are other personal factors that make some much more vulnerable to media reports of suicide.

Improving suicide reporting by governments and improving suicide-related research will require additional resources. There are already shortages of qualified technicians, mental health professionals and supervisors to guide new generations of scientists and medical staff. Research culture is another primary concern, which is often ignored or taken lightly. The Asian region is in need of more research and medical professional staff who can effectively disseminate rigorous methods and protocols. This all requires additional funding and more human resources. Some advancements may be made by better connecting independent researchers and developing networks of scientists working in important areas like suicide prevention.

Asharani et al. have done an important service by providing a large and rich meta-analysis of suicide reporting data in Asia. Future researchers need to move things up a notch by improving the quality, transparency and openness of their methods and data. Clinicians can contribute by providing more qualitative and small-scale studies on celebrity infatuation and suicidality. Governments can help by collecting and disseminating suicide data quickly and more transparently. Media organisations can help by collaborating with researchers and also holding them to account on the strength of the evidence. Media suicide reporting guidelines should be based on facts not myths. For example, there is no positive value in supporting the myth that talking about suicide makes people suicidal. With this meta-analysis, and others to come, we can bring sound science to the tension between the need to report the news and the need to avoid distressing the news consumer.

Keywords: *media, mental health, psychology, suicide, suicide reporting*

REFERENCES

1. Asharani PV, Koh YS, Tan RSH et al. The impact of media reporting of suicides on subsequent suicides in Asia: A systematic review. Ann Acad Med Singap 2024;53:152-69.
2. Harris KM, Thandrayen J, Samphoas C, et al. Estimating suicide rates in developing nations: A low-cost newspaper capture-recapture approach in Cambodia. Asia Pac J Public Health 2016;28:262-70.
3. Harris KM, Wang L, Mu GM, et al. Measuring the suicidal mind: The 'open source' Suicidality Scale, for adolescents and adults. PloS One 2023;18:e0282009.
4. Ju Ji N, Young Lee W, Seok Noh M, et al. The impact of indiscriminate media coverage of a celebrity suicide on a society with a high suicide rate: Epidemiological findings on copycat suicides from South Korea. J Affect Disord 2014;156:56-61.
5. Scherr S, Reinemann C. Belief in a Werther Effect: Third-person effects in the perceptions of suicide risk for others and the moderating role of depression. Suicide Life Threat Behav 2011;41:624-34.
6. Seah, Jane HK, Shim KJ. Data mining approach to the detection of suicide in social media: A case study of Singapore. (2018). 2018 IEEE International Conference on Big Data (Big Data): Seattle, WA, 10-13: Proceedings: 5442-4.
7. Stack, S. Media coverage as a risk factor in suicide. J Epidemiol Community Health 2003;57:238-40.
8. Tor PC, Beng YN, Yong GA. The media and suicide. Ann Acad Med Singap 2008;37:797-9.
9. World Health Organization. Global Health Estimates: Life expectancy and leading causes of death and disability, 2015. http://www.who.int/healthinfo/mortality_data/en/. Accessed 9 November 2023.
10. World Health Organization. Preventing suicide: a resource for media professionals, 2023. https://www.who.int/publications/i/item/9789240076846. Accessed 9 November 2023.

Ann Acad Med Singap 2024;53:129-31
https://doi.org/10.47102/annals-acadmedsg.202481

EDITORIAL

Bridging expertise with machine learning and automated machine learning in clinical medicine

Chien-Chang Lee[1,2,3] *MD*, James Yeongjun Park[4] *MSc*, Wan-Ting Hsu[5] *MSc*

In this issue of the *Annals*, Thirunavukarasu et al.'s systematic review on the clinical performance of automated machine learning (AutoML) highlights its extensive applicability across 22 clinical specialties, showcasing its potential to redefine healthcare by making artificial intelligence (AI) technologies accessible to those without advanced computational skills.[1] This enables the development of effective AI models that could rival or exceed the accuracy of traditional machine learning (ML) approaches and human diagnostic methods.

This editorial discusses the critical role of ML as the foundation of AutoML, and addresses the complexities involved in model development that AutoML aims to simplify. It also examines the challenges associated with integrating AutoML into healthcare, such as performance variability and ethical concerns, emphasising the necessity of a principled approach to its deployment. We also advocate ongoing dialogue among healthcare stakeholders, increased investment in education, and a steadfast commitment to ethical standards, to fully leverage AutoML's transformative potential in patient care.

In recent years, the expansive potential of ML to transform medical care has gained widespread recognition. ML's capabilities, ranging from diagnosing diseases to predicting treatment outcomes, present an unprecedented opportunity to revolutionise healthcare delivery.[2] The advent of ML in medicine promises a seismic shift in patient care, offering more accurate diagnostics, personalised treatment plans and predictive health outcomes. However, the gap between the expertise required and the available resources in clinical settings has slowed down the adoption of ML in healthcare. This gap is largely caused by intricate programming skills required to develop and implement ML models, which poses a considerable barrier to its broad application in clinical practice. Many healthcare professionals, despite acknowledging the immense benefits of ML, are deterred by the daunting technical complexities involved in programming and deploying ML algorithms. Watson et al. conducted interviews with academic medical leaders to understand these challenges and recommended solutions, such as building partnership with vendors to overcome them.[3]

Recognising this challenge, researchers and developers have been working tirelessly to develop solutions that bridge this gap. One such solution gaining traction is AutoML, which is designed to streamline the ML analysis process, allowing users with limited programming skills to harness the power of ML for various applications in healthcare. AutoML software packages offer intuitive interfaces that guide users through the entire ML workflow, from data preprocessing to model selection and evaluation. By automating complex tasks, such as feature engineering, hyperparameter tuning and model selection, AutoML empowers clinicians and researchers to focus on interpreting results and making informed decisions rather than grappling with technical complexities. The beauty of AutoML lies in its democratising effect on ML, making it accessible to a broader audience within the healthcare community. Clinicians, who may not have a background in computer science or statistics, can now leverage ML techniques to derive insights from medical data with relative ease. This democratisation of ML holds the potential to accelerate innovation and improve patient outcomes across various medical domains.

The review by Thirunavukarasu et al., conducted according to a PROSPERO-registered protocol (CRD42022344427), searched the Cochrane Library, Embase, MEDLINE and Scopus databases up to 11 July 2022.[1] Screening of abstracts and full texts, data extraction, and quality assessment were performed by 2 researchers, with disagreements resolved through discussion or third-

[1] Department of Information Management, Ministry of Health and Welfare, Taiwan
[2] Department of Emergency Medicine, National Taiwan University Hospital, Taiwan
[3] Center of Intelligent Healthcare, National Taiwan University Hospital, Taiwan
[4] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, US
[5] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, US
Correspondence: Prof Chien-Chang Lee, Department of Information Management, Ministry of Health and Welfare, No.488, Sec. 6, Zhongxiao E. Rd., Nangang Dist., Taipei City 115, Taiwan.
Email: hit3transparency@gmail.com, cclee100@gmail.com

party arbitration. The review included 82 studies featuring 26 distinct autoML platforms, primarily focused on brain and lung diseases among 22 specialties. Variable performance was observed across autoML platforms, with area under the receiver operator characteristic curve (AUCROC) ranging from 0.35 to 1.00, F1-score from 0.16 to 0.99, and area under the precision-recall curve (AUPRC) from 0.51 to 1.00. AutoML demonstrated strong performance metrics in the majority of trials, with AutoPrognosis and Amazon Rekognition emerging as top performers for unstructured and structured data, respectively. However, the quality of reporting was generally poor, with a median Developmental and Exploratory Clinical Investigations of DEcision support systems driven by Artificial Intelligence (DECIDE-AI) score of 14 out of 27. Despite this, autoML shows promise in various clinical contexts, performing comparably to bespoke computational and clinical benchmarks. Future research should focus on enhancing the quality of validation studies, as the integration of autoML with large language models holds potential for advancing user-defined goals in AI-driven healthcare development. Despite the varied performance of these platforms, a significant number of instances highlighted AutoML's capability to either match or surpass the efficacy of bespoke ML models, with significant advantages over traditional ML in detection and diagnosis.[4] This revelation is groundbreaking, positing AutoML as a potential equaliser in the realm of medical AI, thereby facilitating a more inclusive and efficient approach to patient care that is both personalised and scalable.

Acknowledging the transformative potential of ML in clinical settings, analysis of sepsis patient data from the US National Inpatient Sample by Park et al. serves as a testament to this.[2] Through the development of models employing advanced computational techniques and benchmarking their performance against the Super Learner model, they have illustrated the power of traditional ML approaches in medical research. This experience, while not directly involving AutoML, highlights the significant impact that ML can have on understanding and treating complex conditions such as sepsis. It also underscores the potential for AutoML to further simplify and expedite these analytical processes in future studies. By automating the more technical aspects of model development, AutoML could make it feasible for a wider range of healthcare professionals to engage with ML, thereby broadening the scope of its application in clinical research and patient care. Despite not utilising AutoML,[2] the insights gained lay a foundational understanding of ML's impact, setting the stage for AutoML to potentially streamline similar research endeavours. The variability in AutoML's performance, as noted in the review by Thirunavukarasu et al.,[1] coupled with the need for rigorous validation protocols, presents challenges that must be navigated carefully. Ethical considerations, data privacy and algorithmic bias are additional factors that necessitate a principled approach to AI's[5] and AutoML's deployment in clinical settings.

It is important to acknowledge that AutoML comes with several inherent limitations. First, it lacks the flexibility to adapt ML algorithms, which may limit its applicability to specific use cases or evolving data landscapes. Second, the rapid pace of ML algorithm development means that AutoML may struggle to incorporate the latest advancements in a timely manner, potentially lagging behind state-of-the-art approaches. Additionally, AutoML may not offer several advanced training methods, such as pre-training, which could restrict its effectiveness in certain complex scenarios. Last, the rapid evolution of large language models has significantly reduced the barrier to entry for ML programming, potentially overshadowing the utility of AutoML in certain contexts. These limitations underscore the importance of considering the specific needs and constraints of each project when deciding whether to employ AutoML or traditional ML approaches.

The journey towards fully integrating AutoML into healthcare is complex and multifaceted, fraught with challenges but brimming with potential. To bridge this gap, we recommend establishing interdisciplinary teams that combine clinical knowledge with ML expertise to oversee the implementation of AutoML solutions, ensuring they are clinically relevant and effectively integrated into patient care workflows. As we stand on the cusp of this new era in healthcare, the promise of AutoML to revolutionise patient care is palpable, contingent upon our collective commitment to harnessing this technology responsibly and effectively. The integration of AutoML into clinical practice, especially within specialties such as neurology and pulmonology, heralds a new epoch of medical treatment, where the confluence of human expertise and AI paves the way for unprecedented advancements in patient care and health outcomes.

Keywords: *AI, artificial intelligence, automated machine learning, autoML, machine learning, medical AI*

REFERENCES

1. Thirunavukarasu AJ, Elangovan K, Gutierrez L, et al. Clinical performance of automated machine learning: A systematic review. Ann Acad Med Singap 2024;53:187-207.
2. Park JY, Hsu TC, Hu JR, et al. Predicting sepsis mortality in a population-based national database: Machine learning approach. J Med Internet Res 2022;24:e29982.
3. Watson J, Hutyra CA, Clancy SM, et al. Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers? JAMIA Open 2020; 10:167-72.
4. Bai A, Si M, Xue P, et al. Artificial intelligence performance in detecting lymphoma from medical imaging: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2024;24:13.
5. Busch F, Adams LC, Bressem KK. Biomedical Ethical Aspects Towards the Implementation of Artificial Intelligence in Medical Education. Med Sci Educ 2023;33:1007-12.

Ann Acad Med Singap 2024;53:132-41
https://doi.org/10.47102/annals-acadmedsg.2023119

ORIGINAL ARTICLE

Mitigating adverse social determinants of health in the vulnerable population: Insights from a home visitation programme

Kevin Wei Hao Liang[1,2,3] *MMed*, Keith Wei Han Liang[1,3] *MMed*, Qing Rong Chan[2] *BA*, Winnie Hwee Suat Goh[2,3,4] *FAMS*

ABSTRACT

Introduction: Low-income families are exposed to adverse childhood experiences and psychosocial risks that impact child development. At the KK Women's and Children's Hospital in Singapore, Kids Integrated Development Service (KIDS0-3) is a home visitation programme that aims to optimise the development of children from low-income families.

Method: Data comprising family demographics, maternal psychosocial risks and outcomes of child development were collated through a chart review of 469 mother-child dyads enrolled from June 2014 to October 2022.

Results: Based on the Family and Adult Support Tool, 312 families (67%) were identified as moderate or high-risk. Children from moderate and high-risk families had poorer Bayley cognitive (mean 95.88 [SD 8.25] versus [vs] 98.44 [SD 8.72], *P*=0.014) and language scores (mean 87.38 [SD 10.35] vs 90.43 [SD 9.61], *P*=0.016] at 24 months of age, compared to the low-risk group. Children of teenage mothers had lower Bayley cognitive scores (mean 95.16 [SD 8.42] vs 97.76 [SD 8.55], *P*=0.037), and children of mothers who experienced sexual abuse had lower Bayley cognitive scores (mean 93.1 [SD 5.68] vs 99.7 [SD 8.17], *P*=0.013) and language scores (mean 82.3 [SD 12.87] vs 91.3 [SD 10.86], *P*=0.021]. Antenatal enrolment yielded better child language (mean 90.1 [SD 9.37] vs 87.13 [SD 10.79], *P*=0.04) and motor outcomes (mean 99.62 [SD 9.45] vs 94.72 [SD 9.51], *P*=0 .001) than postnatal enrolment.

Conclusion: Psychosocial risks impact the development of children from low-income families in Singapore. Findings underscore the importance of early, integrated intervention for vulnerable families.

Ann Acad Med Singap 2024;53:132-41

Keywords: adverse childhood experiences, child development outcomes, low-income families, maternal psychosocial risks, toxic stress

CLINICAL IMPACT

What is New

- This is a Singapore study describing the psychosocial risks of mothers from low-income families and its impact on child development.
- Findings underscore the importance of upstream intervention in vulnerable families.

Clinical Implications

- Findings highlight the need for early, tailored support for vulnerable families, in particular, teenage mothers and mothers with history of sexual abuse.
- Upskilling of home visitors in infant mental health and trauma-informed care is crucial to support the needs of vulnerable families.

INTRODUCTION

Strong evidence consistently links low income to Adverse Childhood Experiences (ACEs) and children's long-term health, developmental, educational and social outcomes.[1,2]

Poverty increases parenting stress, and this is especially important in early childhood when the home environment and parent-child bond are the main contributing factors in shaping children's biological and psychosocial pathways.[3] Furthermore, research has linked early poverty to changes in brain structures critical for emotional regulation and cognitive function.[4]

Despite the growing economy, poverty remains evident in Singapore. A substantial 12% of households in Singapore do not have adequate income to meet basic consumption needs; the relative poverty rate indicates that 24% of households in Singapore do not earn enough to keep up

[1] Department of Paediatrics, KK Women's and Children's Hospital, Singapore
[2] KKH KIDS0-3, Division of Medicine, KK Women's and Children's Hospital, Singapore
[3] Duke-NUS Medical School, Singapore
[4] Department of Child Development, KK Women's and Children's Hospital, Singapore
Correspondence: Dr Kevin Wei Hao Liang, Department of Paediatric Medicine, KK Women's and Children's Hospital, 100 Bukit Timah Road, Singapore 229899.
Email: kevin.liang.w.h@singhealth.com.sg

with the rest of society.[5] A study conducted by the Singapore Children's Society showed that 62% of children from low-income families experienced at least 3 ACEs.[6]

To achieve health equity and mitigate downstream effects of poverty, it is pertinent to commence early integrated interventions from conception and early childhood, especially during the first 1000 days of life.[7–10] Interventions designed to optimise child development should aim at strengthening the capability and capacity of their caregivers.[11] These efforts can shift the odds towards more favourable developmental outcomes, especially for children at risk.[12]

The Kids Integrated Development Service (KIDS0-3) began in 2014 at KK Women's and Children's Hospital (KKH), the largest public hospital providing tertiary healthcare services for women and children in Singapore. Using an integrated health and social approach, this programme aims to mitigate adverse social determinants of health and optimise the developmental potential of children from low-income families. To date, there have been limited local data pertaining to maternal psychosocial risk factors and its association with child development outcomes in the vulnerable population.

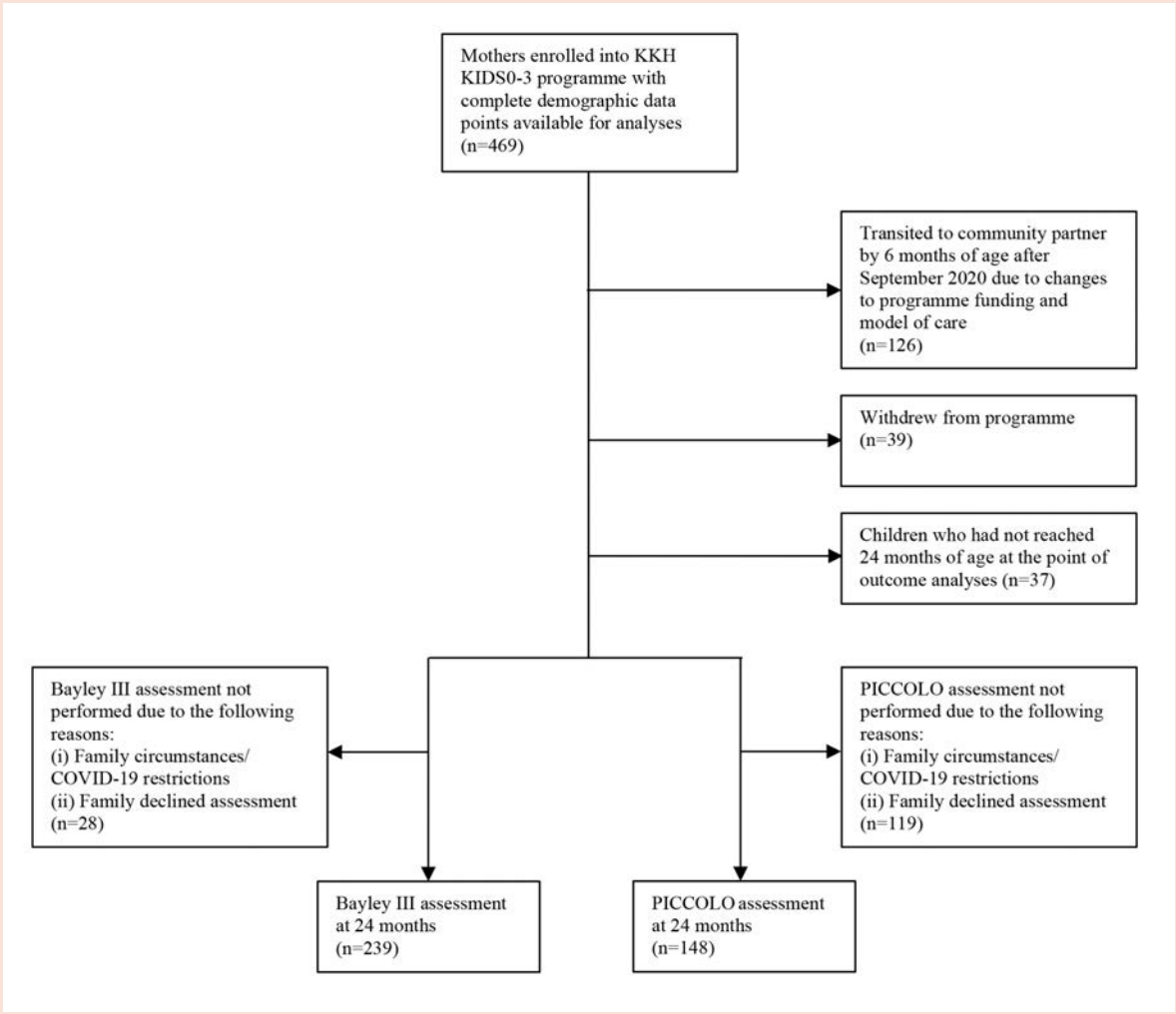
This paper aims to describe the demographics and psychosocial risk factors of mothers from low-income families enrolled into the programme at pregnancy or within 1 week postpartum, as well as to study the associations between psychosocial risk factors and child development outcomes.

METHOD

Study design

This is a descriptive retrospective study of a single-centre home visitation programme targeting low-income families in Singapore. Information was collated through a chart review of 469 mother-child dyads enrolled in the KIDS0-3 programme from June 2014 to October 2022 (Fig. 1). Three hundred and sixty mothers (77%) were enrolled prenatally, and 109 (23%) were enrolled postnatally.

Fig. 1. Number of families enrolled and assessment of demographic data and outcomes.



This programme included pregnant women whose children would be born Singapore citizens at KKH or infants that are Singapore citizens, from less than 7 days old and with per capita family income ≤ SGD650 or gross household income ≤ SGD2500. This targets to support the bottom 20th percentile by household income in the population.[13]

By virtue of the fact that specialised services would better meet their needs, babies with the following problems were excluded from the programme: malignancy; congenital cyanotic cardiac disease; severe birth defects detected antenatally; genetic conditions associated with significant developmental delay and/or associated with other organ involvement; chronic conditions with neuromuscular or musculoskeletal involvement that would impede their ability to perform developmental activities; conditions with high nursing needs (e.g. long-term nasogastric tube feeding, stoma, tracheostomy); and any other conditions with specific programmes in the hospital or community that can cater to their needs (e.g. home care, very low birth weight baby <1.5 kg, early intervention programme for infants and children).

The programme is transdisciplinary, where frontline team members, known as key workers, are the single point of contact for the family.[14] They consist of nurses who support families during the antenatal and early postnatal period up to 12 weeks post-delivery, and health visitors who provide support from 12 weeks onwards. They are supported by the team of doctors, social workers and allied health professionals through peer learning, coaching and reflective supervision.

Enrolled mothers receive regular home visits starting from pregnancy until their child reaches the age of 3. During pregnancy, nurses conduct a home visit in each trimester, and advise mothers on maternal nutrition, mental wellness, exercise, and smoking and alcohol cessation. The importance of key antenatal investigations and appointments is reinforced to improve understanding and compliance. Pre-delivery, mothers are advised on birth preparedness, home readiness, breastfeeding and parent craft. They are guided using principles of infant mental health, to promote early stimulation of their newborn through daily caregiving practices. From 12 weeks onwards, the health visitors conduct home visits every fortnight till 6 months of age and monthly visits thereafter. They focus on promoting healthy nutrition, growth and child development practices. Key workers apply the Abecedarian approach, a set of evidence-based strategies implemented through individualised adult-child interactions.[15,16] It comprises 4 interconnected elements that facilitate adult-child interaction and stimulate children's development: language priority, enriched caregiving, conversational reading and learning games. Every key worker is trained in the Abecedarian approach, and fidelity checks are conducted for quality assurance. The programme also adopts a developmental parenting approach to home visitation, which comprises evidence-based strategies to engage parents in supporting children's development.[17] In this regard, the Home Visit Rating Scales have been applied to guide observations in home visit quality and improve home visitation practices and family engagement.[18]

Through a transdisciplinary and family-centred approach, the key workers aim to build strong rapport with families and develop self-sufficient families who can navigate societal challenges. This is augmented by collaborative partnerships established with community agencies and preschools.

Study variables

Demographic data comprising age, ethnicity, education level, employment status and housing type were obtained at enrolment. Psychosocial risks were assessed using the Family and Adult Support Tool (FAST),[19] which is adapted from Child and Adolescent Needs and Strengths, or CANS — a family of planning and outcome management tools developed by the John Praed Foundation.[20] It is a multipurpose communicative tool for understanding family needs and strengths in order to prioritise intervention, and this has been standardised for use among family service centres in Singapore which provide community-based social services. It measures the domains of safety concerns, risk behaviours, individual functioning, family functioning and family strengths. Each domain is scored on a level of 0 to 3, where 3 indicates the highest risk. Families in the programme were rated as "low risk" if risk factors were scored at 0 or 1 only, "moderate risk" if at least one level-2 risk was identified, and "high risk" if at least one level-3 risk was identified. ACEs were recorded through a questionnaire adapted from the Centers for Disease Control and Prevention's ACEs study in the US.[21] Data collection for ACEs commenced mid-way during the study; hence, only 340 cases were recorded.

Outcome measures included child development and parent-child interaction at 24 months of age. Due to changes in the government's funding model from September 2020 onwards, children assessed to be of low-to-moderate risk were transited to the community partner by 6 months of age for continued follow-up through home visits. Children assessed to be of high risk continued to receive

support from the programme till 3 years of age. In addition, due to home-visiting restrictions during the COVID-19 pandemic and family circumstances, the above assessments could not be completed for every family. In particular, assessment of parent-child interaction involved a process of video recording and hence, there was a substantial number of mothers who declined. In view of these reasons, data at 24 months for development and parent-child interaction were only available for 51% and 32% of the sample, respectively (Fig. 1).

The Bayley-III tool was administered to children at 24 months of age to assess developmental outcomes.[22] Bayley-III is a standardised neurocognitive assessment tool used to assess the development of infants and young children aged between 1 and 42 months across 5 domains (cognitive, language, motor, adaptive and social-emotional) using a series of play tasks and parent report questionnaires. Scoring is dichotomous (1, 0) and the composite score for Bayley-III is calculated for the cognitive, language and motor scales.

Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO) was used to assess parent-child interaction at 24 months of age. It is a validated tool for children ages 10 to 47 months, that predicts child outcomes in cognitive development, vocabulary and behaviour using observational measures of parenting interactions.[23] It measures 4 domains of developmental parenting (affection, responsiveness, encouragement and teaching), where the observation of items is scored from 0 to 2 (0=absent, 1=barely, 2=clearly). A scoring grid is used, and the individual domains are rated as below average, average or above average.

Statistical analysis

Baseline demographics, psychosocial risk based on FAST and maternal ACEs were summarised as frequencies and percentages. Continuous variables were compared using independent-samples t-test. Independent-samples t-test was used to compare mean scores for Bayley-III among families belonging to different risk groups based on FAST (moderate/high and low); it was also performed to compare Bayley-III mean scores of children among mothers who were enrolled antenatally versus postnatally. Chi-squared test was used to compare the scores of various domains of PICCOLO among families in the moderate/high- and low-risk groups. All data analyses for this report were performed using SPSS Statistics version 28 (IBM Corp, Armonk, NY, US). A P value <0.05 was considered statistically significant.

RESULTS

Population demographics and psychosocial risk profile

In our cohort, 118 (25%) were young mothers below the age of 21 years. Two hundred and sixty-nine were Malay (57%), and the rest comprised 101 (22%) Chinese, 60 (13%) Indian, 2 (0.1%) Eurasian, and 37 (7.9%) from other ethnic groups. Two hundred and fifty-three (54%) mothers were unemployed and 267 (57%) lived in 1- or 2-room public rental flats. Three hundred and seven (65%) mothers received education of secondary school level or below. Three hundred and sixty (77%) mothers in the programme were recruited antenatally, with 50 (11%) mothers recruited in the first trimester of pregnancy (Table 1).

Based on the FAST, 272 (58%) families had significant financial needs, 112 (24%) had significant mental health issues, and 117 (25%) had criminal involvement (Table 1); 312 (67%) families were identified to have moderate-to-high risks at enrolment. One hundred and seventy mothers (50%, n=340) had experienced ≥3 ACEs (Table 2).

Child development outcomes

The Bayley-III tool was administered to 239 (51%) children at 24 months of age. Out of 239 children, 234 (98%), 196 (82%) and 232 (97%) had normal cognitive, language and motor ability respectively.

Children of teenage mothers had lower Bayley cognitive scores (teenage mothers: mean 95.16, SD 8.42; non-teenage mothers: mean 97.76, SD 8.55), and children of mothers who reported an experience of sexual abuse scored lower in Bayley cognitive (mean 93.1, SD 5.68) and language scores (mean 82.3, SD 12.87), compared with those whose mothers did not (cognitive: mean 99.7, SD 8.17; language: mean 91.3, SD 10.86).

Children from families in the moderate/high-risk group based on FAST had poorer Bayley cognitive (mean 95.88, SD 8.25) and language scores (mean 87.38, SD 10.35) at 24 months of age, compared with those in the low-risk group (cognitive: mean 98.44, SD 8.72; language: mean 90.43, SD 9.61) (Table 3). In addition, there were more families in the moderate/high-risk group scoring below average PICCOLO scores in Encouragement ($\chi 2$ [2,171]=7.12, P=0.028) and Teaching ($\chi 2$ [2,171]=9.56, P=0.008) when compared with those in the low-risk group.

Table 1. Sample demographics, FAST items, frequency and percentages (n=469).

Demographic categories		Number of families	Percentage (%)
Maternal age	Below 21 years of age	118	25
	21–25 years of age	125	27
	26–30 years of age	116	25
	Above 30 years of age	110	23
Maternal ethnicity	Chinese	101	22
	Malay	269	57
	Indian	60	13
	Eurasian	2	0.1
	Other ethnic groups	37	7.9
Maternal education level	Special needs school	6	1
	Up to primary school level	43	9
	Up to secondary school level	264	56
	Vocational school	103	22
	Diploma	33	7
	Degree and higher	20	5
Maternal employment status	Working full-time	47	10
	Working part-time/adhoc	37	8
	Homemaker	120	26
	Others	12	2
	Unemployed	253	54
Housing type	1- & 2-room public rental flat	267	57
	3-, 4-, 5-room public purchased flat	184	39
	Executive apartment or condominium	3	1
	Others	15	3
Recruitment period	First trimester	50	11
	Second trimester	140	30
	Third trimester	170	36
	Postnatal	109	23
FAST rating at point of intake	Low needs	157	33
	Moderate needs	204	44
	High needs	108	23
Moderate/high needs for the following FAST domains			
Safety concerns	Child abuse	13	3
	Intimate partner violence	17	4
	Home environment safety	20	4

Table 1. Sample demographics, FAST items, frequency and percentages (n=469). (Cont'd)

Demographic categories		Number of families	Percentage (%)
Risk behaviours	Suicide risk	10	2
	Self-harm risk	7	1
Individual functioning	Physical/medical	98	21
	Employment	51	11
	Educational	14	3
	Mental health	112	24
Family functioning	Family conflict issues	40	12
	Family communication issues	56	16
	Financial	272	58
	Social support	70	15
	Criminal involvement	117	25
Caregiving	Involvement in care	3	0.5
	Supervision/discipline	7	1
	Caregiver knowledge	9	2
	Caregiver stress	8	2
	Caregiver organisational skills	3	0.5
Family strength	Relationship	15	3
	Resiliency	4	1
	Resourcefulness	1	0.2
	Motivation to change	1	0.2

FAST: Family and Adult Support Tool

Language (mean 90.10, SD 9.37) and motor (mean 99.62, SD 9.45) development scores for children whose mothers were enrolled antenatally were higher than children whose mothers participated postnatally (language: mean 87.13, SD 10.79; motor: mean 94.72, SD 9.51) (Table 4).

DISCUSSION

This study gives insight into the psychosocial risk factors of vulnerable families, their impact on child development outcomes, and provides a better understanding of the role of a home visitation programme in mitigating adverse social determinants in Singapore.

Compared to the Singapore adult population, mothers from our cohort had a considerably higher proportion of ACEs, especially for divorce/separation, incarceration, violence, substance abuse and sexual abuse.[24]

A quarter of mothers in our cohort were less than 21 years of age. Children born to teenage mothers in our cohort had lower Bayley cognitive scores at 24 months of age. The effect of younger maternal age on cognitive outcomes in children is likely to be associated with the sociodemographic circumstances that the children are born into. Morinis et al.[25] also showed a strong association between young maternal age and poor cognitive outcomes at 5 years, but after adjustment for confounders, this effect was mostly explained by marked inequalities in sociodemographic circumstances and perinatal risks. A more tailored intervention for young adolescent mothers is important. Key areas include maternal nutrition, mental wellness and the avoidance of smoking, alcohol use and substance abuse. Key workers also emphasised contraception, family planning and role of parents. The programme has strengthened

Table 2. Maternal Adverse Childhood Experiences (ACEs; n=340).

	Number of mothers[a] (KIDS 0-3)	KIDS 0-3[a] (%)	Singapore[b] (%)
Types of ACEs			
Divorce/Separation	178	53	22
Incarceration	109	33	5
Mental illness	51	15	6
Substance abuse	85	25	5
Violence	102	30	8
Economic hardship	133	39	-
Emotional neglect	145	43	46
Physical neglect	46	14	6
Physical abuse	75	22	5
Emotional abuse	85	25	8
Sexual abuse	45	13	5
Number of ACEs			
0 ACEs	56	16	
1 ACE	51	15	
2 ACEs	63	19	
3 ACEs	39	11	
4 or more ACEs	131	39	

[a] KIDS ACEs data collection commenced mid-way of study. Hence, there were only 340 cases with ACEs recorded.
[b] Data for Singapore adult population from Subramaniam M, Abdin E, Seow E, et al. Prevalence, socio-demographic correlates and associations of adverse childhood experiences with mental illnesses: Results from the Singapore Mental Health Study. Child Abuse Negl 2020;103:104447.

Table 3. Bayley-III and FAST risk: means, standard deviations and *P* values (n=239).

	FAST risk mean (SD)		
Bayley-III domains	Low	Moderate/high	t-test statistic (*P*-value)[a]
Cognitive	98.44 (8.72)	95.88 (8.25)	2.474 (0.014)
Language	90.43 (9.61)	87.38 (10.35)	2.435 (0.016)
Motor	98.15 (9.89)	98.20 (9.55)	-0.040 (0.968)

[a] Comparing means of Bayley scores between children from families in the low versus moderate/high groups based on FAST risk.
FAST: Family and Adult Support Tool; SD: standard deviation

collaborative efforts with the hospital's teenage pregnancy clinic to enrol pregnant adolescents. The team also works closely with community partners to facilitate early infant care placement as mothers seek education or employment opportunities. Where needed, the programme works with voluntary welfare organisations that can provide baby necessities, counselling services or a temporary shelter during family crises. Importantly, the key workers journey closely with adolescent mothers to monitor their mental health, develop their parenting confidence and empower them to make responsible life decisions.

In our study, 13% of mothers reported an experience of sexual abuse. Children of these mothers scored significantly lower in Bayley cognitive scores and language scores. Mothers' early adversity has been known to affect their

Table 4. Recruitment period and Bayley-III: means, standard deviations and *P* values for variables (n=239).

	Recruitment period mean[a] (SD)		
Bayley-III domains	Antenatal	Postnatal	t-test statistic (*P*-value)[b]
Cognitive	97.75 (8.67)	96.87 (7.68)	0.733 (0.464)
Language	90.10 (9.37)	87.13 (10.79)	2.06 (0.040)
Motor	99.62 (9.45)	94.72 (9.51)	3.55 (0.000)

[a] Composite scores were compared with the standardised composite scoring provided by the developers: 69 and below: extremely low; 70–79: borderline; 80–89: low average; 90–109: average; 110–119: high average; 120–129: superior; and 130 and above: very superior.
[b] Comparing means of Bayley scores between children whose mothers were enrolled into the programme antenatally versus postnatally

children's development, largely through effects on maternal mental health and parenting confidence.[26] In particular, childhood sexual abuse results in increased risks for mothers and their children due to the intergenerational nature of this trauma and the disruption it causes in the mother-child relationship.[27,28] Consequently, disruption to maternal attachment can have a lasting impact on neurological, emotional and social development.[29] With the higher prevalence of ACEs in vulnerable families, in particular, sexual abuse, there is a pressing need to upskill the team in infant mental health and trauma-informed practice. The key workers have undergone certificate courses conducted by the University of Minnesota, Center for Early Education and Development, that provide a crucial foundation in infant mental health and intervention with infants, toddlers, and their parents. The programme has also embarked on training sessions facilitated by mental health experts to equip the team on evidence-based principles of trauma-informed practice.[30] This allows key workers to support the families more effectively, identify mental health concerns, and facilitate early referral to the Mental Wellness Service or Psychosocial Trauma Support Service where warranted. A greater emphasis is also being placed on reflective supervision, where key workers share their feelings about working with vulnerable families; this aids to decrease burnout and improve satisfaction and morale.[31]

Families in the moderate/high-risk groups based on FAST were also associated with poorer Bayley cognitive and language scores as well as poorer PICCOLO encouragement and teaching scores, compared to those who belonged to the low-risk group. Poverty increases young children's exposure to biological and psychosocial risks that affect development through changes in brain structure and function.[32] Such children are often exposed to multiple and cumulative risks, and their development is increasingly compromised with the accumulation of such risks.[33] In addition, a lack of positive parenting behaviours negatively affects child development, and this is apparent as early as 6 months of age.[34] Though the findings of lower Bayley scores among children of mothers with moderate/high risk were statistically significant, we recognise that the clinical outcomes may vary. This is in part due to the intervention from key workers and government assistance from social service agencies, which may have helped to buffer the impact of the adverse environment for children belonging to the moderate/high-risk groups. When supporting families with higher psychosocial risk, the programme focuses on promoting positive parenting behaviours. One approach is through video-feedback intervention, where key workers share with parents the strengths observed in their parent-child interaction, and potential areas of improvement in each domain. In high-risk families where mothers are unable to function as the primary caregiver, the programme works with the family and government agencies to find an alternative caregiver, who can build a safe, stable and nurturing relationship with the child, thereby buffering the impact of toxic stress.

Positive developmental outcomes were seen in children whose mothers were enrolled antenatally into the programme. The findings build on existing literature showing that investment in the vulnerable population should commence before birth because intervention during pregnancy can mould the future trajectory of health and abilities.[35] During antenatal home visits, mothers are provided with anticipatory guidance that promotes mental wellness and healthy pregnancies. Mothers are taught early developmental stimulation practices in preparation for their newborns, and this is reinforced postnatally. This has been key in helping mothers to support their child's development through daily caregiving routines and practices. Although 77% of mothers were enrolled antenatally, only 11% were enrolled in the first trimester. It is hence crucial to streamline processes and improve community networking to allow early referrals of pregnant women at risk.

CONCLUSION

It is the responsibility of every child health worker to advocate for and achieve health equity in the vulnerable population. As Singapore's attention shifts from "survive to thrive", and from "healthcare to health", it is pertinent that these families are not left behind.[36] Altering the developmental trajectory of a young child growing up in an adverse environment requires much perseverance.[12] A holistic population health approach with inter-professional collaboration within the community is crucial. We must emphasise positive parenting and strengthen the capacity of parents to fulfil their roles effectively[37]—this will give children a better start in life.

The findings highlight the psychosocial risk profile of the low-income families in Singapore and its impact on child development outcomes. Early targeted interventions such as home visitation are important to reach families who may have difficulties accessing healthcare services due to psychosocial constraints.

Further development to tailor intervention and support is pertinent, as this home visitation programme expands to benefit more families. This can be achieved with continual refining of our service delivery through data feedback, regular upskilling of team members, and strengthening of collaboration with our partners in the ecosystem.

Conflict of interest

None to declare.

Acknowledgement

The authors wish to acknowledge the following programme funders: Temasek Foundation Cares (in the pilot phase), Early Childhood Development Agency and KidSTART Singapore Limited.

The authors wish to acknowledge A/Prof Chan Yoke Hwee (Senior Consultant, Chairman Medical Board, KK Women's and Children's Hospital, Singapore) and Prof Chay Oh Moh (Emeritus Consultant, Campus Director, Education Office; Programme Director, Kids Integrated Development Service, KK Women's and Children's Hospital, Singapore) for their invaluable input and feedback in the writing of this paper.

REFERENCES

1. Centers for Disease Control and Prevention. Preventing Adverse Childhood Experiences: Leveraging the Best Available Evidence. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 2019.
2. Shonkoff JP, Garner AS, Siegel BS, et al. The lifelong effects of early childhood adversity and toxic stress. Pediatrics 2012;129:e232-46.
3. Chaudry A, Wimer C. Poverty is not just an indicator: The relationship between income, poverty, and child well-being. Academic Pediatrics 2016;16:S23-9.
4. Luby JL, Constantino JN, Barch DM. Poverty and Developing Brain. Cerebrum 2022:cer-04-22.
5. Ng IYH. Definitions and measurements of poverty 2020. SSR Snippet 2020;4:2-9.
6. Toh SM, Khong J, Liu D. What difficult life events do children from low-income families face. Research Bites 2018;6:1-2.
7. Almond D, Currie J, Duque V. Childhood circumstances and adult outcomes: Act II. Journal of Economic Literature 2018;56:1360-446.
8. Shonkoff JP, Boyce WT, McEwen BS. Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. JAMA 2009;301:2252-9.
9. Moore T, Arefadib N, Deery A, et al. The first thousand days: An evidence paper. Melbourne: Centre for Community Child Health; 2017.
10. Clark H, Coll-Seck AM, Banerjee A, et al. A future for the world's children? A WHO-UNICEF-Lancet Commission. Lancet 2020;395:605-58.
11. Chong WH, Choo H, Chay OM et al. Preventive Child Healthcare in Singapore: A Parents' Well-Being Perspective. Ann Acad Med Singap 2015;44:550-3.
12. Ho LY. Child development programme in Singapore 1988 to 2007. Ann Acad Med Singap 2007;36:898-910.
13. Department of Statistics, Singapore. General Household Survey 2015. https://www.singstat.gov.sg/publications/ghs/ghs2015content. Accessed 4 July 2023.
14. Noah's Ark Inc. The Key Worker: resources for early childhood intervention professionals. Malvern, Victoria: ECII; 2012.
15. Sparling J, Meunier K. Abecedarian: An Early Childhood Education Approach that has a Rich History and a Vibrant Present. International Journal of Early Childhood 2019; 51:207-16.
16. García JL, Heckman JJ, Leaf DE, et al. Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program. J Polit Econ 2020;128:2502-41.
17. Roggman LA, Boyce LK, Innocenti MS. Developmental parenting: A guide for early childhood practitioners. Baltimore, Maryland: Paul H. Brookes Publishing; 2008.
18. Roggman LA, Cook GA, Innocenti MS, et al. The Home Visit Rating Scales: Revised, restructured, and revalidated. Infant Ment Health J 2019;40:315-30.
19. The John Praed Foundation. Family and Adult Support Tool: The Ministry of Social and Family Development User Guide. Singapore: Ministry of Social and Family Development, John Praed Foundation; 2014.
20. Lyons JS, Weiner DA, Lyons MB. Measurement as communication in outcomes management: The child and adolescent needs and strengths (CANS). In: Maruish ME (Ed). The use of psychological testing for treatment planning and outcomes assessment. New York: Routledge; 2014.
21. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. American journal of preventive medicine 1998;14:245-58.
22. Bayley N. Bayley Scales of Infant and Toddler Development (3rd ed). San Antonio, Texas: Harcourt Assessment Inc; 2006.

23. Roggman LA, Cook GA, Innocenti MS, et al. Parenting interactions with children: Checklist of observations linked to outcomes. Baltimore, Maryland: Paul H. Brookes Publishing; 2013.
24. Subramaniam M, Abdin E, Seow E, et al. Prevalence, socio-demographic correlates and associations of adverse childhood experiences with mental illnesses: Results from the Singapore Mental Health Study. Child Abuse Negl 2020;103:104447.
25. Morinis J, Carson C, Quigley MA. Effect of teenage motherhood on cognitive outcomes in children: a population-based cohort study. Arch Dis Child 2013;98:959-64.
26. Treat AE, Sheffield-Morris A, Williamson AC, et al. Adverse childhood experiences and young children's social and emotional development: the role of maternal depression, self-efficacy, and social support. Early Child Dev Care 2020;190:2422-36.
27. Duncan KA. Healing from the trauma of childhood sexual abuse: The journey for women. Philadelphia: Penn State Press; 2004.
28. Duncan KA. The impact of child sexual abuse on parenting: A female perspective. In: Walz GR, Yep RK. (Eds), VISTAS: Compelling perspectives on counseling. Alexandria, VA:ACA; 2005.
29. Schore AN. The effects of early relational trauma on right brain development, affect regulation, and infant mental health. Infant Mental Health Journal: Official Publication of The World Association for Infant Mental Health 2001; 22:201-69.
30. Huang LN, Flatow R, Biggs T et al. SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach. U.S. Department of Health and Human Services Substance Abuse and Mental Health Services Administration Office of Policy, Planning and Innovation; 2014.
31. Shea SE, Sipotz K, McCormick A, et al. The implementation of a multi-level reflective consultation model in a statewide infant & early childcare education professional development system: Evaluation of a pilot. Infant Mental Health Journal 2022;43:266-86.
32. Walker SP, Wachs TD, Meeks Gardner J, et al. Child development: risk factors for adverse outcomes in developing countries. Lancet 2007;369:145-57.
33. Wachs TD. Necessary but not sufficient: the respective roles of single and multiple influences on individual development. American Psychological Association; 2000.
34. Glascoe FP, Leew S. Parenting behaviors, perceptions, and psychosocial risk: impacts on young children's development. Pediatrics 2010;125:313-9.
35. National Research Council (US) and Institute of Medicine (US) Committee on Integrating the Science of Early Childhood Development. From Neurons to Neighborhoods: The Science of Early Childhood Development. Shonkoff JP, Phillips DA (Eds). Washington (DC): National Academies Press (US); 2000.
36. Tan CC, Lam CSP, Matchar DB et al. Singapore's health-care system: key features, challenges, and shifts. Lancet 2021; 398:1091-104.
37. Ho LY. Raising Children in Singapore: A Paediatrician's Perspective. Ann Acad Med Singap 2009;38:158-62.

Ann Acad Med Singap 2024;53:142-51
https://doi.org/10.47102/annals-acadmedsg.2023168

ORIGINAL ARTICLE

Health practices, behaviours and quality of life of low-income preschoolers: A community-based cross-sectional comparison study in Singapore

Shang Chee Chong[1,2] *MRCPCH*, Ramkumar Aishworiya[1,2] *MRCPCH*, Woon Li Seo[1,2] *MRCPCH*, Yee Keow Chiong[1,2] *MRCPCH*, Gerald Choon-Huat Koh[3,4,5] *MBBS*, Jeremy Bingyuan Lin[1,2] *MRCPCH*, Lynn Heng[1] *MSc (Early Childhood Development and Education)*, Thahirah Habib Mohd[1] *BA (Psychology)*, Young Ern Saw[1] *BA (Psychology)*, Yiong Huak Chan[6] *PhD*, Jing Shi Chua[7] *BSS (Mathematics)*, Shefaly Shorey[7] *PhD*

ABSTRACT

Introduction: Children from low-income (LI) families often suffer from poor health, with sub-optimal health practices. This cross-sectional study examined the differences in health habits and health-related quality of life (HRQoL) of LI preschool children compared to non-low-income preschool peers (PPG).

Method: Using data from the social-health Circle of Care-Health Development Screening Programme (CoC-HDSP) in Singapore, 118 LI children and 304 PPG children aged 18 months to 6 years old and their families were recruited from 13 government-funded preschools. Health practices examined included screen time habits, sleep, nutrition, dental health and the children's HRQoL using PedsQL 4.0 Generic Core Scales.

Results: Majority of the children were aged 4–6 years in kindergarten 1 and 2. There were more Malay children in the LI than the PPG (61.9% versus [vs] 29.3%, *P*<0.001). Low-income children were more likely to have lower-educated parents (*P*<0.001). The completed vaccination rate in the LI group was lower than those in PPG (84.7% vs 98.0%, *P*<0.001). More in the LI group utilised emergency services for acute illnesses (*P*<0.05). Fewer LI children had ever visited a dentist (47.4% vs 75.4%, *P*<0.001), and more LI children consumed sweetened drinks daily (33.3% vs 8.6%, *P*<0.001). The LI group reported poorer-quality sleep (48.3% vs 27.2%, *P*<0.001), though both groups exceeded the daily recommended screen viewing duration. The LI group scored higher in the social (mean 92.4±12.2 vs 84.3±15.3, *P*<0.001) and emotional (mean 85.2±15.1 vs 76.6±17.3, *P*<0.001) domains of the PedsQL 4.0 when compared to PPG.

Conclusion: Low-income children have poorer health practices, receive less preventive paediatric care, and utilise more emergency services for acute illnesses. These findings are important for developing interventions that work towards improving the health of LI children.

Ann Acad Med Singap 2024;53:142-51

Keywords: health practices, low-income, paediatrics, public health, quality of life

CLINICAL IMPACT

What is New

- Children from low-income families have lower sleep quality, poorer dental health habits and less healthy diets than their peers.
- Low-income children received less preventive paediatric care and utilised more emergency than primary care services for acute illnesses.

Clinical Implications

- This study confirmed the need for attention on health needs of low-income children in Singapore.
- A deeper understanding about health practices and knowledge-practice gaps of parents of low-income children can facilitate future research directions and meaningful interventions.

[1] Khoo Teck Puat-National University Children's Medical Institute, National University Health System, Singapore
[2] Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[3] Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[4] Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[5] Department of Medicine, National University Hospital, National University Health System, Singapore
[6] Biostatistics Unit, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[7] Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
Correspondence: Dr Shefaly Shorey, Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Level 2, Clinical Research Centre, Block MD11, 10 Medical Drive, Singapore 117597.
Email: nurssh@nus.edu.sg

INTRODUCTION

Poverty is a serious concern that has been found to bring about various adverse psychological, social and developmental outcomes.[1] Living in poverty as a child can affect an individual's life well into adulthood[2] due to risks including poor nutrition, poorly controlled chronic ailments and unstable environments.[3,4] Overall, these children are often worse off than their counterparts with higher family income—in terms of health, academic achievements and psychological well-being.[2,5-7]

Singapore has recently stepped-up efforts to give children living in poverty equal opportunities as other children by implementing early childhood interventions, which aim to promote child health and development, facilitate parent-child bonding and improve family functioning.[8] Like in other countries, these aim to reduce health disparities. However, health services and health policies will first need a deeper understanding of priority health issues in these children, before specific recommendations can be made for such intersectoral programmes.[9,10]

The present study is part of a larger study involving a preventive health programme called the Circle of Care-Health and Development Screening Programme (CoC-HDSP). The CoC was developed by a team of social workers from Care Corner, a social service agency in Singapore, to bring social work and parenting supports to families of low-income (LI) preschoolers. Healthcare professionals from the National University Hospital Child Development Unit and Paediatrics Department enhanced CoC to introduce health screening and education through the Health and Development Screening Programme (HDSP), with case coordination among a multidisciplinary team. This study sought to do a cross-sectional review of the baseline health habits of LI children and parental perception of their quality of life, when compared to the preschool peer group (PPG) from non-LI families.

METHOD

The health habits of low-income (LI) children in Singapore under the health intervention group (CoC-HDSP) were compared to a control group of preschool children from non-LI families at baseline.

Participants

Children receiving CoC-HDSP must be: (1) aged 18 months to 6 years; (2) attending government-funded preschools in Singapore; and (3) have a household income below SGD3000 or per capita income below SGD750. Children with complex or life-threatening medical care issues requiring frequent hospitalisations were excluded from HDSP, as their baseline health characteristics and quality of life would differ from the general paediatric population. We did not exclude children who had common but chronic medical conditions that may require outpatient visits (e.g. eczema and asthma) as these are not uncommon within the paediatric population. Children who were PPG must be: (1) from households with higher household income and per capita income than the CoC-HSDP group; and (2) of similar age-range within same school classes as their LI peers.

Procedure

Data collection took place from November 2018 to November 2019. Thirteen preschools under the CoC programme were identified. In each preschool, there was an average of 1 to 3 classes across nursery and kindergarten levels, with 15 to 20 children per class. There are 1 to 2 classes per school at each level dedicated to the CoC programme. Children in CoC programme were invited to participate in the HDSP arm of the study. Data on health practices was taken at baseline for the LI children who opted to be in the CoC-HDSP, and these children were already receiving social and parenting supports under CoC. Non-LI children who were part of these classes were also invited to participate in the study, forming the PPG. Instead of additional health-screening services, they received basic vision screening and height and weight assessments under the national health screening exercise. Parents were asked to complete a health survey that included providing the health information of their children.

Outcome measures

To assess the health practices of the children, the following instruments were administered:

The health assessment questionnaire (HAQ). The HAQ was developed by clinicians and the research team, and involved questions adapted from validated sleep, health habits and parenting habits questionnaires from previous studies.[11-13] This 57-item instrument assessed various outcomes like screen time exposure, sleep, nutrition, dental health, parenting concerns, parental high-risk behaviours (e.g. smoking) and the presence of any medical concerns or known medical conditions of the child. Information regarding the child's vaccination records and other health records in the past year were also captured.

Paediatric Quality of Life Inventory Version 4.0 (PedsQL 4.0). To measure health-related quality of life (HRQoL), the PedsQL 4.0 was used. The patent-rated proxy versions, including the

23-item parent report for young children aged 5–7 and the 21-item parent report for toddlers aged 2–4, were used. The PedsQL 4.0 has 4 generic score scales: physical functioning (8 items), emotional functioning (5 items), school functioning (5 items) and social functioning (5 items). Each item was based on a 5-point Likert scale, and scale scores were calculated by dividing the sum of item scores by the number of items. Higher scores on the PedsQL indicate better HRQoL. The PedsQL 4.0 has been reported to have high internal consistency of 0.90 for the parental proxy version.[14]

Data analysis

Data were analysed using the IBM Statistical Package for the Social Sciences (SPSS) software version 26 (IBM Corp, Armonk, NY, US).[16] Linear regression and logistic regression, adjusting for sex, race and age range, were performed to assess the differences in numerical and binary outcomes, respectively, between the children in the intervention group and controls.

Ethical consideration

Ethics approval was obtained from the National Healthcare Group Domain Specific Review Board (2018/00629) before the commencement of this study. Written consent before the data was obtained and confidentiality was ensured. Voluntary participation was reinforced.

RESULTS

A total of 118 participants from the LI intervention group and 304 participants from the PPG were included in the analyses (see Table 1). Majority of the children were aged 4–6 years old in kindergarten 1 and 2 classes. There was a higher proportion of Malay children in LI group (61.9% versus [vs] 29.3%, $P<0.001$), and twice as many Chinese students in the PPG (54.9% vs 25.4%, $P<0.001$) (Table 1). No significant sex or age difference was observed.

In the LI group, there were more single-parent or divorced families, and parents who worked part-time or who were unemployed. Parents in the LI group, either fathers or mothers, tended to have lower levels of education and were more likely to be smokers ($P<0.001$).

General health

Despite more children (Table 3) from LI families than PPG having pre-existing conditions (e.g. asthma, eczema) ($P=0.005$), half of them had no follow-up for these conditions. A lower proportion of LI children had up-to-date vaccinations (84.7% vs 98%, $P<0.001$). Fewer of the LI visited primary care for acute medical conditions like fever and vomiting (19.5% vs 52.0%, $P<0.001$), but instead utilised emergency settings for such needs (22.9% vs 13.2%, $P<0.05$).

Table 1. Participant characteristics (low-income intervention group and preschool peer group).

Characteristics		Low-income intervention group (LI) (n=118)		Non-low-income preschool peer group (PPG) (n=304)		*P* value
		n	%	n	%	
Sex	Male	55	46.6	158	52.0	0.323
	Female	63	53.4	146	48.0	
Ethnicity	Chinese	30	25.4	167	54.9	<0.001
	Malay	73	61.9	89	29.3	
	Indian	7	5.9	24	7.9	
	Others	8	6.8	22	7.2	
	Unstated	0	0.0	2	0.7	
Age range (years)	<2	3	2.5	7	2.3	0.402
	2–3	16	13.6	41	13.5	
	3–4	29	24.6	69	22.7	
	4–5	36	30.5	77	25.3	
	5–6	25	21.2	95	31.3	
	>6	9	7.6	15	4.9	

Table 2. Family demographics and characteristics.

Characteristics		Low-income intervention group (LI) (n=118)		Non-low-income preschool peer group (PPG) (n=304)		*P* value
		n	%	n	%	
Marital status	Single	11	9.3	4	1.3	<0.001
	Married	90	76.3	292	96.0	
	Widowed	2	1.7	2	0.7	
	Divorced/separated	15	12.7	6	2.0	
Main caregiver	Mother	99	83.9	130	42.8	<0.001
	Father	11	9.3	18	5.9	
	Helper	0	0.0	5	1.6	
	Grandparents	2	1.7	25	8.2	
	Both parents	1	0.8	92	30.3	
	Others (e.g. aunts)	3	2.6	0	0.0	
	Unstated	2	1.7	34	11.2	
Employment status of caregiver	Full-time	59	50.0	206	67.8	<0.001
	Part-time	25	21.2	29	9.5	
	Not employed	34	28.8	62	20.4	
	Unstated	0	0.0	7	2.3	
Number of smokers	0	39	33.1	201	66.1	<0.001
	1	52	44.1	84	27.6	
	2 or more	25	21.2	18	5.9	
	Unstated	2	1.7	7	2.3	
Education level (mother)	Primary school or below	21	17.8	12	3.9	<0.001
	Completed secondary school	41	34.7	41	13.5	
	Completed ITE/NTE	21	17.8	27	8.9	
	Polytechnic diploma/A-level	19	16.1	55	18.1	
	Completed university	4	3.4	93	30.6	
	Postgraduate degree	0	0.0	53	17.4	
	Others	8	6.8	18	5.9	
	Unstated	4	3.4	5	1.6	
Education level (father)	Primary school or below	13	11.0	9	3.0	<0.001
	Completed secondary school	37	31.4	46	15.1	
	Completed ITE/NTE	29	24.6	51	16.8	
	Polytechnic diploma/A-level	9	7.6	68	22.4	
	Completed University	4	3.4	82	27.0	
	Postgraduate degree	0	0.0	37	12.2	
	Others	1	0.8	7	2.3	
	Unstated	25	21.2	4	1.3	

Responses under the "others" option for education level includes diplomas or degrees from other countries.

Table 3. Children's general health.

Characteristics		Low-income intervention group (LI) (n=118)		Non-low-income preschool peer group (PPG) (n=304)		*P* value
		n	%	n	%	
Pre-existing medical condition (common paediatric conditions, e.g. asthma, eczema)	Yes	26	22.0	33	10.9	0.005
	No	92	78.0	270	88.8	
	Parent is unsure	0	0.0	1	0.3	
Follow-up with doctor for above medical condition	Once a month or more frequent	4	15.4	1	3.0	<0.001
	Once every 3–5 months	5	19.2	6	18.2	
	Once every 6–12 months	3	11.5	14	42.4	
	Only when necessary	1	3.8	10	30.3	
	No follow-up	13	50.0	2	6.1	
Vaccinations status	Up-to-date	100	84.7	289	95.1	<0.001
	Incomplete	16	13.6	6	2.0	
	Parent is unsure	2	1.7	9	3.0	
Type of medical attendances in the past 1 year for acute illnesses (at least once in the past year)	Polyclinic	23	19.5	158	52.0	<0.001
	Children emergency	27	22.9	40	13.2	<0.05
	Hospitalisation	13	11.0	20	6.6	0.16

Only 18 participants from the LI group and 23 from the PPG indicated that their child requires follow-up for the medical condition, while remaining participants indicated that no follow-up was required.

Health awareness and practices

Dental hygiene and practices. The majority of parents from both groups recognised the importance of dental hygiene (i.e. 100% from LI and 88.8% from PPG "agreed or strongly agreed" that dental hygiene was important). Although no significant differences were found in the frequency of teeth brushing between both groups of children (61.0% LI vs 63.2% of PPG children brushed at least twice a day, *P*=0.683), parents from the PPG than the LI group were more aware of the appropriate amount of toothpaste to use (87.5% vs 40.7% , adjusted odds ratio AOR 13.6, 95% confidence interval [CI] 7.6–24.4, *P*<0.001). Significantly more children in the PPG than LI group had ever visited a dentist for check-up (75.4% vs 47.4%, AOR 3.1, 95% CI 1.9–5.2, *P*<0.001). Overall, more children in the LI group were reported to have cavities (22.0% vs 11.5%, AOR 2.1, 95% CI 1.2–3.9, *P*=0.016).

Sleep quantity. Overall, children from both groups fell short of the National Sleep Foundation's recommended guidelines of 10–13 hours of sleep daily for preschoolers at night (Table 4), though making up for the total sleep duration recommendations with daytime naps. Children in the LI group slept significantly longer than their counterparts in the PPG at night during weekdays (*P*=0.047) and also took longer naps than children in the PPG during weekdays (*P*=0.007), and marginally shorter naps during weekends (*P*=0.092).

Sleep quality. Significantly more parents in LI than PPG expressed concerns regarding their children's sleep quality (48.3% vs 27.2%, AOR 3.9, 95% CI 2.3–6.6, *P*<0.001). The top 3 sleep problems reported by LI parents were: (1) the child needs an accompanying adult or pacifier/milk bottle to fall asleep or when the child wakes up at night (33.1%); (2) teeth grinding (13.6%); and (3) snoring (12.7%). Comparatively, sleep problems reported by the PPG parents were about sleep resistance (19.8%); and sleep latency (17.0%).

Table 4. Amount of sleep (in minutes) in low-income and preschool peer groups.

Characteristics	Low-income group (n= 118)		Non-low-income preschool peer group (n= 304)		*P* value[a]
	Mean (mins)	SD	Mean (mins)	SD	
Daytime sleep (weekday)	111.4	83.8	92.0	59.6	0.007
Night-time sleep (weekday)	557.4	70.0	532.3	86.3	0.047
Daytime sleep (weekend)	88.7	76.2	95.1	71.0	0.092
Night-time sleep (weekend)	573.8	98.7	557.1	97.0	0.226

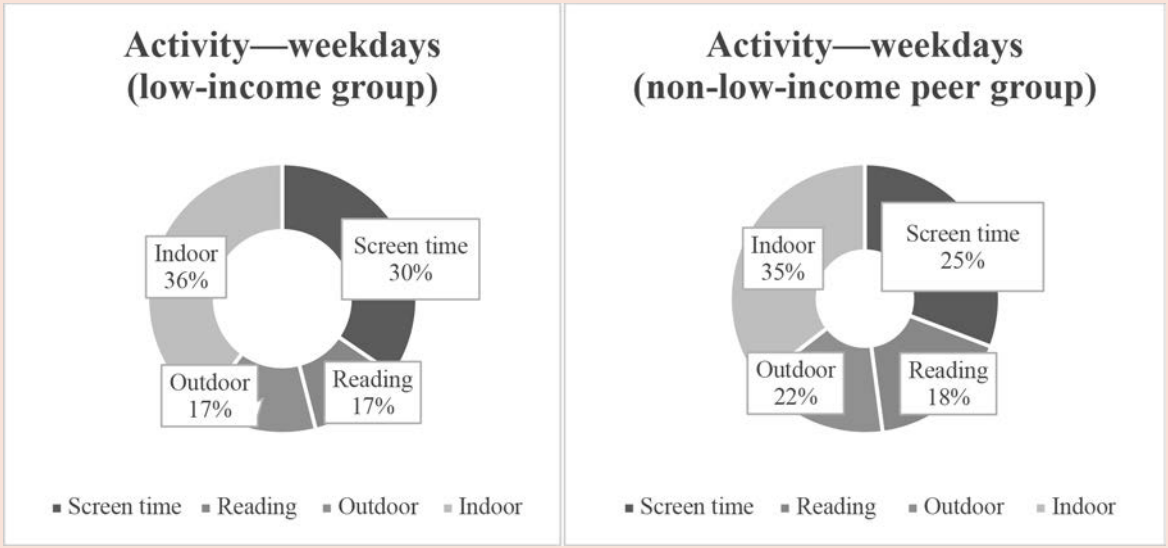
[a] Adjusting for sex, ethnicity and age range

Nutrition. Significantly more parents in the LI group than in the PPG "agreed or strongly agreed" that they knew about what makes a balanced meal for their child (89.7% vs 80.1%, AOR 2.0, 95% CI 1.0–4.0, *P*=0.050) and felt they could provide a balanced meal for their child (90.6% vs 76.9%, AOR 2.5, 95% CI 1.3–5.2, *P*=0.008). However, LI families reported a higher daily consumption of sweetened drinks (33.3% vs 8.6%, AOR 4.9, 95% CI 2.7–8.9, *P*<0.001) and confectionary (41.9% vs 25.7%, AOR 1.9, 95% CI 1.2–3.1, *P*=0.007) by the children. They also consumed more processed food (12.9% vs 4.3%, AOR 2.2, 95% CI 0.98–5.1, *P*=0.057) weekly. Parents from the LI group reported more fussy eating in their children (54.2% vs 34.9%, AOR 2.0, 95% CI 1.3–3.2, *P*=0.003).

Proportion of time spent on activities. Figure 1 presents the proportion of time the children spent on various activities. During weekdays, children from the LI group spent significantly less time on outdoor physical activities (adjusted *p*= 0.004) than those in the PPG. They also reported higher screen time (mean 144.1 minutes [mins]±SD 89.4mins) than the PPG (mean 124.1mins±SD 122.4mins). However, differences in screentime between both groups were not significant on both weekdays (*P*=0.315), and weekends (*P*=0.209). The proportion of supervised screen time also did not differ between both groups on weekdays (*P*=0.946) and weekends (*P*=0.455).

Health-related QoL. Interestingly, the LI group scored higher in the social (mean 92.4±12.2 vs 84.3±15.3, *P*<0.001) and emotional (mean

Fig. 1. Time spent on various activities.



Indoor activities included other activities like indoor play, art and craft, writing and drawing, but excluded screen time or reading, which were measured separately.

85.2±15.1 vs 76.6±17.3, *P*<0.001) domains of the PedsQL 4.0 when compared to the PPG (Table 5). Less difference was noted for physical and school function.

DISCUSSION

Social determinants are important for child health and outcomes, with health-related behaviours and home and environmental conditions as influencing factors.[15] Within the Singapore context, other studies have supported these social determinants as well—with higher socioeconomic status, healthy diets, higher physical activity, and non-smoking being related to better health.[16,17] Our study echoes findings from other studies that poorer health practices and knowledge-practice gaps within these families have the potential to negatively impact the overall health of the children.[18-20] Compared to non-LI children, more LI children have existing medical issues, with more visits to the emergency departments for acute needs. Like other studies, this indicates that such families usually have no point-of-care other than emergency visits.[21] Reasons for this may be related to parents waiting on acute conditions until they worsen, lack of access to primary care beyond office hours for parents on shift-work, or simply a lack of education on use of health services.

Childhood vaccinations, recognised globally as important for protection against vaccine-preventable diseases, covers 12 preventable diseases under the National Childhood Immunisation Schedule in Singapore.[22] UNICEF reported a high vaccine coverage of 95% to 99% of Singaporean children for the top important preventable childhood diseases.[23] Despite many of these vaccinations being fully subsidised, vaccine coverage for LI children under this schedule is lower than the national and peer-average, indicative of the disparity and need for action in this area.

Dental caries afflict around 40% of 3- to 6-year-old Singaporean children, and are more common among those from the lower social-economic group.[24] Affordability issues for fluoride toothpaste and dental health services, and difficulties accessing preventive dental services (where only selected polyclinics have paediatric services) could have contributed directly to reasons for poor dental health in our study.[25,26] Existing literature show that oral health literacy has a direct impact on perioral health,[27] and attention to oral health literacy and dental care access are important factors to explore further for improving dental health, especially for the socio-economically disadvantaged.[28]

Our study is consistent with other local studies where both income groups reflect unhealthy screen practices, which exceeded the recommended 1-hour screen-viewing duration set by the American Academy of Paediatrics for children of preschool ages.[29,30] Lower maternal education has been found to be a risk factor for higher television viewing in infants,[31] which is negatively associated with subsequent cognitive and language skills.[32] Parental knowledge of screen viewing recommendations has been suggested as important in improving children's screen practices, including balancing sedentary screen-viewing with physical activity.[33] Assisting parents with selection of higher-quality educational content may be especially beneficial for language and cognitive development of children,[34] even if these parents struggle to supervise their children while on screens.

Overall, our study supports previous studies on sleep in preschool children in Singapore, with

Table 5. PedsQL 4.0 scores based on domain age group.

Domain	Low-income group (LI) (n=118)		Non-low-income group (PPG) (n=304)		
	Mean	SD	Mean	SD	*P* value[a]
Physical function	91.0	8.92	87.2	18.0	0.102
Emotional function	85.2	15.1	76.6	17.3	<0.001
Social function	92.4	12.2	84.3	15.3	<0.001
School function	76.7	15.1	75.9	17.7	0.907
Total scale score	86.6	9.12	81.0	13.1	0.001

[a] Adjusting for sex, ethnicity and age range
Total scale score was computed based on the sum of total scores divided by the number of completed items.

both LI and PPG children still requiring naps and having shortened night-time sleep duration overall.[35] On weekdays, LI children had longer night-time sleep than their peers. Possibly, LI parents put their children to bed earlier to manage their household chores, while non-LI parents spend their night hours after work with the children on play or achievement-oriented activities. For parents of PPG children, the main sleep concerns were sleep resistance and sleep latency, which was similar to what was reported by Aishworiya and colleagues.[11] In comparison, the LI group reported issues of requiring an adult to co-sleep or a sleep-object as a top challenge—suggesting that the noisy living environments of LI families in smaller flats with poor sleep hygiene likely contributed to these practices. These children required milk bottles and pacifiers to go to bed as well, which is not exemplary of healthy sleeping habits in children aged 4 to 6. We did not capture the bedtime and wake times of the children; this would have been interesting to find out if parents in contracted shift work affect the sleep habits of the children, which can also impact wake time and school-attendance.

Consumption of sweetened beverages, confectionary and processed food by children from LI families was significantly higher than PPG children. The Singapore Longitudinal Early Development Study reported that half the children from food-insecure families were consuming sweetened beverages and sweetened and salted snacks at least thrice a week and instant noodles and fast food over twice weekly.[36] Frequent consumption of these food items is associated with increased cardiometabolic risk due to high concentration of refined carbohydrates, fats and sodium.[37] However, these processed foods, due to their longer shelf-life, form the bulk of donated food items to LI families.[38] Additionally, there were challenges of LI children being fussy-eaters, causing parents to have compounded struggles with managing feeding behaviour problems—which can lead to poorer nutrition. Our study also revealed a knowledge-practice gap among parents, where parental knowledge of preparing healthy and balanced meals for children was unmatched with practices of providing frequent sweetened and processed foods. Underlying reasons to be explored in future studies include the capacity of parents of these children to prepare meals (e.g. single parent, long and uncertain shift work hours) and affordability of fresh ingredients for meal preparation. Certainly, solutions include fresh food donations instead of processed foods, and providing parenting education on cultivating healthy eating habits.

This study provides interesting insights into the perceived quality of life of Singapore preschool children—which is overall very high on the PedsQL 4.0 in both groups. Surprisingly, LI parents rated their children higher on social and emotional functioning than PPG children, which was contrary to some existing literature.[39,40] The level of support rendered by the social workers in the programme overall might have promoted a sense of social well-being and better emotional coping, as supported by other published research on LI families.[41] This may highlight the potential beneficial impacts of such community interventions.

Strengths and limitations

The present study provided local, albeit regional, evidence of the health status and health practices of children from LI families in the community. It informs on priority health areas of focus, where meaningful interventive and health-promoting strategies can then be explored. It explored and revealed that socioeconomic gradients exist in health practices among young children and families.

There are limitations to this study. There is an inherent recall bias and interpretation variation in survey and questionnaire studies. It is also difficult to tease out the myriad of family and social factors that contribute to HRQoL parent-report measures. The data were limited to a group of children who were undergoing a community intervention, which may not fully represent LI children at a population level. As Singapore is a high-income country, the relative poverty in Singapore appears very differently from the absolute poverty present in other countries. Thus, access to LI families for research has been known to be notoriously difficult, especially when Singapore does not have an official poverty line.[42] Furthermore, there was a rather large difference in sample sizes between groups, which might have affected the representativeness of the population and the study's statistical power. This was related to the difficulties of recruitment of LI families for research, and data collection was affected by the COVID-19 pandemic. Lastly, the Personal Data Protection Act prevented sharing of sensitive social data (e.g. parental incarceration, child neglect) with the study team for analysis of health outcomes in the highest-risk children.

CONCLUSION

This study allowed an understanding of the health practices among LI children and families, which potentially allows for further research and a deeper understanding of meaningful interventions. LI

families do not perceive their children's HRQoL to be worse, and different supports may impact domains of HRQoL differently. Further improving their health literacy, designing health systems to better promote preventive care, as well as working across social and community agencies to advance health equity and outcomes through holistic partnerships will be important, for existing and future intervention programmes.

Conflicts of interest

None to declare.

Funding

This study was funded by the Singapore Population Health Improvement Center (SPHERiC) seed funding grant number SPHERiC-018.

Acknowledgment

The authors would like to thank the Lien Foundation and Care Corner Singapore Ltd for the Circle of Care programme; as well as all parents, social workers from Care Corner, preschool principals and teachers from the PAP Community Foundation, Ministry of Education (MOE) Kindergartens, and Persatuan Pemudi Islam Singapura (PPIS) preschools for participating in the study.

REFERENCES

1. Chaudry A, Wimer C. Poverty is not just an indicator: the relationship between income, poverty, and child well-being. Academic pediatrics 2016;16:S23-9.
2. Lesner RV. The long-term effect of childhood poverty. Journal of population economics 2018;31:969-1004.
3. Chong WH, Choo H, Goh EC, et al. Preventive child healthcare in Singapore: A parents' well-being perspective. Ann Acad Med Singap 2015;44:550-3
4. Evans GW, Kutcher R. Loosening the link between childhood poverty and adolescent smoking and obesity: The protective effects of social capital. Psychological Science 2011;22:3-7.
5. Bøe T, Sivertsen B, Heiervang E, et al. Socioeconomic status and child mental health: The role of parental emotional well-being and parenting practices. Journal of abnormal child psychology 2014;42:705-15.
6. Parker S, Greer S, Zuckerman B. Double jeopardy: the impact of poverty on early child development. Pediatr Clin North Am 1988;35:1227-40.
7. los Reyes-Aragon D, Jose C, Amar Amar J, et al. The care and development of children living in contexts of poverty. Journal of Child and Family Studies 2016;25:3637-43.
8. Seah YH. Improving Early Childhood Outcomes for Children and Families: Lessons From The KidSTART Journey. In: Ng IYH, Neo YW (Eds.) Working with low-income families through the life course: Challenges to social services. Singapore: National University of Singapore; 2024.
9. Black MM, Walker SP, Fernald LCH, et al. Early childhood development coming of age: science through the life course. Lancet 2017;389:77-90.
10. Xie H, Nah YH, Yang X, et al. Early childhood intervention: What we know and where we are headed. A review of local and international literature and implications for Singapore. NIE Working Series Paper 18. Singapore: National Institute of Education; 2021.
11. Aishworiya R, Chan P, Kiing J, et al. Sleep behaviour in a sample of preschool children in Singapore. Ann Acad Med Singap 2012;41:99.
12. Owens JA, Dalzell V. Use of the 'BEARS' sleep screening tool in a pediatric residents' continuity clinic: a pilot study. Sleep medicine 2005;6:63-9.
13. Wu XY, Han LH, Zhang JH, et al. The influence of physical activity, sedentary behavior on health-related quality of life among the general population of children and adolescents: A systematic review. PloS one 2017;12:e0187668.
14. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Med Care 2001;39:800-12.
15. Spencer N. The social determinants of child health. Paediatrics and Child Health 2018;28:138-43.
16. AshaRani P, Sin KY, Abdin E, et al. The Relationship of Socioeconomic Status to Alcohol, Smoking, and Health: a Population-Level Study of the Multiethnic Population in Singapore. International Journal of Mental Health and Addiction 2022;22:1-19.
17. Zhou Y-F, Song X-Y, Pan X-F, et al. Association between combined lifestyle factors and healthy ageing in Chinese adults: the Singapore Chinese Health Study. The Journals of Gerontology: Series A 2021;76:1796-805.
18. Garg A, Butz AM, Dworkin PH, et al. Screening for basic social needs at a medical home for low-income children. Clinical Pediatrics 2009;48:32-36.
19. Odgers CL, Adler NE. Challenges for low-income children in an era of increasing income inequality. Child Development Perspectives 2018;12:128-33.
20. de Buhr E, Tannen A. Parental health literacy and health knowledge, behaviours and outcomes in children: a cross-sectional survey. BMC Public Health 2020;20:1096.
21. Perry CD, Kenney GM. Preventive care for children in low-income families: how well do Medicaid and state children's health insurance programs do? Pediatrics 2007;120e1393-401.
22. Ministry of Health. Childhood Immunisation. Singapore: Ministry of Health Singapore; 2021.
23. UNICEF. Immunization coverage survey data. 2022. Accessed 23 April 2023.
24. Gao XL, Hsu CY, Loh T, et al. Dental caries prevalence and distribution among preschoolers in Singapore. Community Dent Health 2009;26:12-7.
25. Gussy MG, Waters E, Riggs E, et al. Parental knowledge, beliefs and behaviours for oral health of toddlers residing in rural Victoria. Australian dental journal 2008;53:52-60.
26. Rai NK, Tiwari T. Parental factors influencing the development of early childhood caries in developing nations: a systematic review. Frontiers in public health 2018;6:64.
27. Baskaradoss JK. Relationship between oral health literacy and oral health status. BMC Oral Health 2018;18:1-6.
28. Winter J, Glaser M, Heinzel-Gutenbrunner M, et al. Association of caries increment in preschool children with nutritional and preventive variables. Clin Oral Investig 2015;19:1913-9.
29. Bernard JY, Padmapriya N, Chen B, et al. Predictors of screen viewing time in young Singaporean children: the GUSTO cohort. International Journal of Behavioral Nutrition and Physical Activity 2017;14:1-10.

30. Hill D, Ameenuddin N, Reid Chassiakos Y, et al. Media and Young Minds. Pediatrics 2016;138:e20162591.
31. Rideout VJ, Hamel E. The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents. San Francisco: Henry J. Kaiser Family Foundation; 2006.
32. Aishworiya R, Magiati I, Phua D, et al. Are There Bidirectional Influences Between Screen Time Exposure and Social Behavioral Traits in Young Children? Journal of Developmental & Behavioral Pediatrics 2022;10:1097.
33. Ebbeck M, Yim HYB, Chan Y, et al. Singaporean parents' views of their young children's access and use of technological devices. Early Childhood Education Journal 2016;44:127-34.
34. Wright JC, Huston AC, Murphy KC, et al. The relations of early television viewing to school readiness and vocabulary of children from low-income families: The early window project. Child Development 2001;72:1347-66.
35. Mindell JA, Sadeh A, Kwon R, et al. Cross-cultural differences in the sleep of preschool children. Sleep Medicine 2013; 14:1283-89.
36. Chen X. Nutrition, Family Stress and Preschool Children's Behavior Problems. In: Singapore Longitudinal EArly Development Study (SG LEADS). Singapore: National University of Singapore; 2021.
37. Shin HJ, Cho E, Lee H-J, et al. Instant noodle intake and dietary patterns are associated with distinct cardiometabolic risk factors in Korea. The Journal of Nutrition 2014;144: 1247-55.
38. Goh CT, Yap C. Pulling together in creative ways, to better feed food insecure Singaporeans. In: Channel News Asia. 2020. Accessed 23 April 2023.
39. Fagan J, Wildfeuer R. Low-Income Parental Risk and Engagement in Early Childhood and Child Social-Emotional Functioning in Middle Childhood. Journal of Child and Family Studies 2022;31:70-85.
40. Ren L, Garcia AS, Esteraich JM, et al. Parent–child relationships and preschoolers' social-emotional functioning among low-income families: the moderating role of parental nativity. Infants & Young Children 2019;32:123-38.
41. Chan KL, Lo CKM, Ho FK, et al. Modifiable Factors for the Trajectory of Health-Related Quality of Life among Youth Growing Up in Poverty: A Prospective Cohort Study. Int J Environ Res Public Health 2 2021;18:9221.
42. Donaldson JA, Loh J, Mudaliar S, et al. Measuring poverty in Singapore: frameworks for consideration. Social Space 2013;58-66.

Ann Acad Med Singap 2024;53:152-69
https://doi.org/10.47102/annals-acadmedsg.2023237

REVIEW ARTICLE

The impact of media reporting of suicides on subsequent suicides in Asia: A systematic review

PV Asharani[1] *PhD*, Yen Sin Koh[1] *MPH*, Rachel Hsiao Shen Tan[1] *BSocSci*, Yoke Boon Tan[1] *BSocSci*, Savita Gunasekaran[1] *BSc Hons*, Benedict Lim[1] *BA*, Lorainne Tudor Car[2,3*] *PhD*, Mythily Subramaniam[1,2,4*] *PhD*

ABSTRACT

Introduction: This systematic review is aimed at (1) evaluating the association between media portrayals of suicides and subsequent copycat suicides or attempts among the general public in Asia, (2) understanding the factors associated with copycat suicides and (3) determining the positive impacts of the media reporting of suicides (e.g. increased help-seeking, coping).

Method: A systematic review and narrative synthesis of English and Chinese articles from 8 electronic databases (i.e. PsycINFO, MEDLINE, Embase, CINAHL, Web of Science, Ariti, China National Knowledge Infrastructure and OpenGrey) from January 2000 to May 2023 was conducted. Observational studies were included, and the data were analysed through narrative synthesis. The protocol was registered with PROSPERO (CRD42021281535).

Results: Among the 32 studies included (n=29 for evidence synthesis) in the review, there is good-quality evidence to show that copycat suicides and suicide attempts increase after media reports of a suicide, regardless of country, celebrity status, study design, type of media, mode of suicide or follow-up period. Females, younger age groups and those sharing similar characteristics as the deceased in publicised suicides (age, gender) were more susceptible to negative impact. Reporting of the mode of death of the deceased increased suicides by the same method among the public.

Conclusion: Media portrayals of suicide appear to have a negative impact on copycat suicides at the population level in Asia. Thus, in addition to tighter media control, healthcare systems, professional medical bodies and community outreach services should work collaboratively to promote early help-seeking in those with psychological distress.

Ann Acad Med Singap 2024;53:152-69

Keywords: celebrity, copycat, media reporting, suicide, Werther effect

CLINICAL IMPACT

What is New

- This systematic review examined the effect of media portrayal of suicides on subsequent copycat suicides among public and found substantial evidence for the Werther effect.
- Those who share similar characteristics as the celebrity, younger age groups and females were more susceptible to negative impact.

Clinical Implication

- Policymakers should work together with media professionals and health systems to create awareness and identify the groups at increased risk, to promote help-seeking and reduce copycat effects.

INTRODUCTION

Globally, around 703,000 people die by suicide annually, 77% of which occur in low- to middle-income countries.[1] In 2019, suicide was the fourth leading cause of death among younger adults (15–29 years).[1] Approximately 60% of the global suicide rates were attributed to Asian countries in 2012,[2] with the data from 2019 showing a higher average rate of suicide in Southeast Asia of 10.2 per 100,000 compared to the global average of 9.0 per 100,000.[3] The characteristics of suicide victims in Asia are different from the rest of the globe: the male–female ratio of suicides is 1:1 (compared to 3:5 in other countries); these deaths occur in rural areas, include specific methods of suicide such as poisoning, and the populations have a lower prevalence of psychiatric conditions.[4]

[1] Research Division, Institute of Mental Health, Singapore
[2] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
[3] Department of Primary Care and Public Health, Imperial College London, London, United Kingdom
[4] Saw Swee Hock School of Public Health, National University of Singapore, Singapore
* Joint last authors
Correspondence: Dr PV Asharani, Research Division, Institute of Mental Health, 10 Buangkok View, Singapore 539747.
Email: Asharani_PEZHUMMOOTTIL_VASUDEVAN_N@imh.com.sg

Religion, cultural beliefs and tradition have a strong influence on the suicide rates in many Asian countries, where suicides are often glorified as acts of valour, devotion or sacrifice.[5] For example, India has the highest rates of suicide among Asian countries (82% in 2012), and there is a strong influence of tradition in people's lives. While the holy scriptures of Hinduism clearly discourage and condemn suicide for selfish reasons, they also condone suicides offered as sacrifices and to uphold duty. The mythology of the sacrifice of sages for the well-being of the world, of royals to escape apostasy (*Jauhar*), and self-immolation of widows in their husband's pyre (*Sati*)[5] have a great influence on the populace. There were tales about mass suicides in the Kingdom of Ayodhya upon the passing of Lord Ram in the Ramayana, an epic that has deep-rooted religious sentiments in people's lives.[6] These strong religious beliefs have resulted in the continued practice of such rituals even in the modern era.[7] This strong devotion is not limited to deities but often extends to the deification of celebrities and politicians who are worshipped with strong passion and have temples built under their names.[8] Worshipping of celebrities and idols is omnipresent and often leads to a tendency to idolise celebrities and follow their every move, appearance and other attributes.[9]

Celebrity obsession tends to spark copycat suicides following the deaths of their idols.[10] A study conducted in South Korea marked a clear spike in copycat suicide by 16.4% within a day of the death of a celebrity.[11] Fu and Chan showed that the increase in copycat suicides is only specific to those who shared similar characteristics (e.g. age, gender, suicide method) as the celebrity.[12] The Werther effect, that is, a rise in copycat suicides following a celebrity suicide, has also been observed in other Asian countries, such as India and Hong Kong.[13,14]

The media guidelines for reporting suicides restrict prominent placement of stories, repeated reporting of the same suicide, usage of sensationalised language, indication of suicide method or location, showing of visuals related to suicide, and links to the celebrity's social media page. Nonetheless, media outlets often disregard the guidelines and glamourise the suicide with the inclusion of sensitive information, which brings about a series of copycat suicides in a short period after a suicide has been reported.[15] A meta-analysis of 31 studies from across the globe showed that the risk of suicide increased by 13% following celebrity suicide reporting.[16] Moreover, a 30% increase in suicides using the same method was noted when the method of suicide was included in the media.[16] Another retrospective time series conducted in 2009 across Hong Kong, Taiwan and China showed a higher risk of copycat suicides by 43%, 29% and 25%, respectively, in the first, second and third weeks following the reports of celebrity suicide.[17] While there is quality evidence from systematic reviews for the Werther effect across the world, no systematic reviews have looked at the trend in Asian countries. Given that the suicide rates, sociodemographic characteristics and cultural influences in Asian populations are different from those in Western countries, a synthesis of relevant evidence through a systematic review is warranted. This information on copycat suicides in Asia is critical to understanding if existing guidelines for suicide reporting and adherence to the media guidelines are adequate to tackle the ongoing situation.

The primary objective was to examine the association between media reporting of suicide and subsequent incidences of copycat suicides and attempts among the public in Asia. The secondary objectives were to understand the factors associated with copycat suicides and whether there were positive impacts of the media reporting of suicides (e.g. higher help-seeking, coping).

We followed the list of countries that fall under the United Nations classification of Asia,[18] which encompasses the geographical area in the eastern and northern hemispheres, extending from the Mediterranean Sea to the Western Pacific Ocean. This includes countries that fall under Central, Eastern, South, Southeastern and Western Asia.

METHOD

Protocol and registration

The systematic review followed the reporting guidelines for the meta-analyses of observational studies in epidemiology (MOOSE) guidelines, and the review was performed as per Cochrane standards. The protocol was registered with PROSPERO (CRD42021281535). A mixed-methods review was initially planned, which was later modified. Qualitative studies and case reports were excluded due to the high heterogeneity and irrelevant outcomes (e.g. attitudes of professionals or the public towards media reporting of suicides) that did not align with the primary outcomes (i.e. copycat suicides and suicide attempts).

Search sources and search

Eight databases (i.e. MEDLINE, Embase, PsycINFO, CINAHL, Web of Science, Airiti, China National

Knowledge Infrastructure and OpenGrey) were searched. The keywords used for each database are included in the Supplementary Appendix S1. We also conducted bibliographic searches from the articles downloaded and subsequent follow-up searches for related articles from Google Scholar using strings of keywords (from the article) or using the related articles retrieval function in Google Scholar.

Screening and eligibility criteria

Studies were included if they were (1) they were in English or Chinese. published between January 2000 and May 2023; (2) they were observational studies; (3) they have primary data; (4) the study setting was in Asia; (5) the exposure was media reporting; and (6) the outcome was suicide rate or attempts. Studies that had secondary data, those reporting unrelated outcomes, data from clinical cohorts published in languages other than English or Chinese, and those with specific study designs (i.e. conference abstracts, qualitative studies and case series or reports) were excluded. Articles were screened against these criteria. Overlapping publications (the same incident in similar settings reported in multiple articles) were screened based on their quality and number of cases. The most robust and comprehensive ones were included. We excluded articles whose eligibility could not be determined based on the information provided in the article and whose authors were uncontactable.

Data collection

Data extraction was done in parallel and independently. The screening of title, abstract and full text was conducted by reviewer pairs (AR and SG; RT and YBT; YSK and BL), and disagreements were resolved by further discussion or by a third reviewer. Disagreements that were not solved by the third reviewer were directed to MS for resolution. We developed a data extraction sheet that was piloted using 5 studies and refined accordingly. The celebrity recognition (i.e. local or international) was determined either by following the article or using Wikipedia[16] when the information was not provided in the article.

Risk of bias (RoB) and publication bias assessment

RoB assessments were performed to determine the sources of bias in various stages of the study, including analysis, reporting and interpretation of results, which could affect the quality of the evidence presented in the systematic review. The RoB of the included studies in this review was assessed using the methods reported by Niederkrotenthaler et al.[16] The RoB was assessed using a modified Robins I tool that assessed biases arising from confounding, intervention classification, preparatory phase, missing data, measurement of outcome, and selection of reported results. Studies were classified according to 4 levels of RoB (low, moderate, serious and critical). The overall bias was determined by the highest classification among the 6 domains measured e.g. when 1 of 6 domains is coded moderate and the remaining domains are coded low, the study will be coded moderate RoB. RoB was assessed for the outcome of interest, and studies with a critical RoB were excluded from the narrative synthesis. Publication bias was assessed using contour-enhanced funnel plots and Egger's regression test.

Data analysis

In the protocol, we indicated that data will be pooled through random effect meta-analysis, and heterogeneity will be measured using the I^2 statistic and the I^2 test. If a high heterogeneity (>75%) was to be identified, we proposed a narrative synthesis approach. The I^2 value for the current dataset was 99.47%, which was not attributable to common confounding variables, such as region (specific countries), type of media (traditional versus [vs] new media), study design (interrupted time series [ITS] vs pre-post design), and publication year, using meta-regression. Therefore, the initial plans for a meta-analysis were switched to a narrative synthesis. Nonetheless, the summarised data have been presented quantitatively through forest plots (rate ratios; Stata MP Version 18 using meta analysis [StataCorp, College Station, TX, US]) to understand the direction of evidence and extent of heterogeneity in the literature that could provide insights for future research.

We performed subgroup analysis based on celebrity status, entertainers vs non-entertainers, and geographical areas (South Asia vs East Asia). We excluded 3 studies with critical RoB from the analysis.

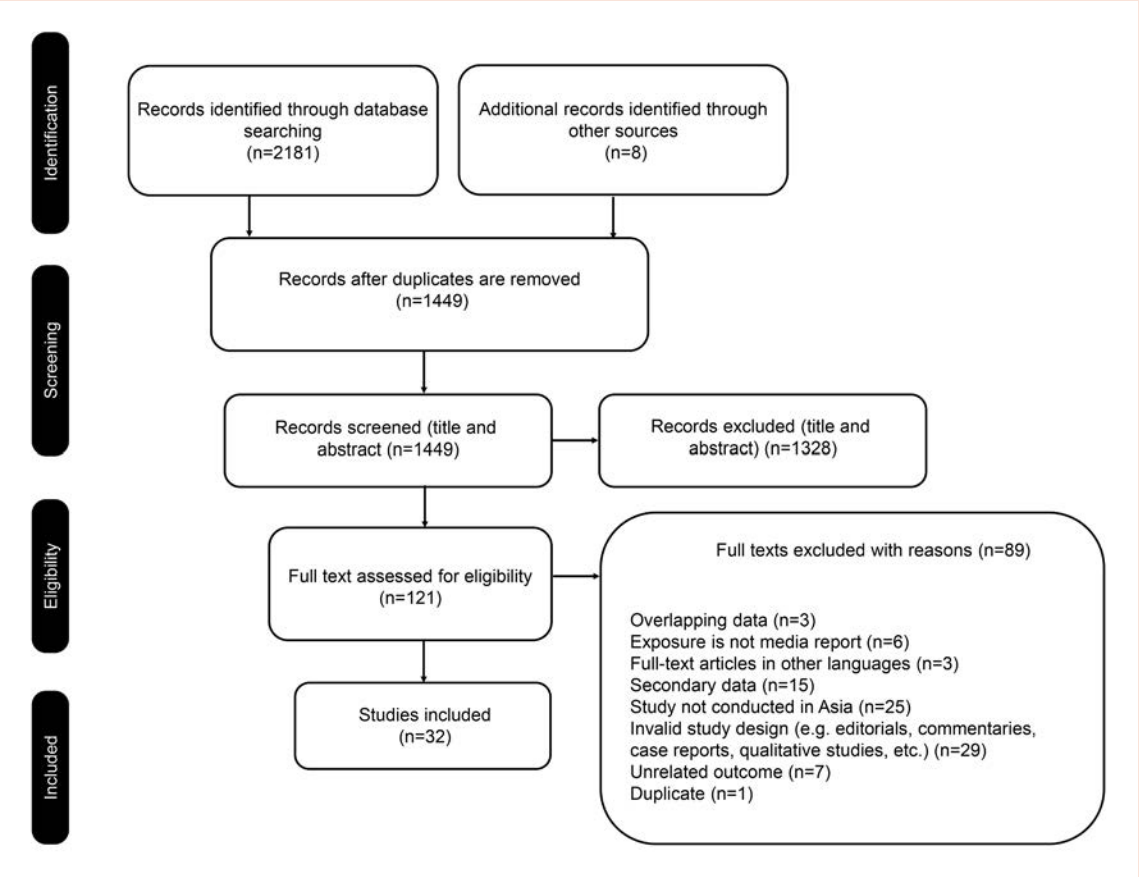
RESULTS

The search retrieved a total of 2181 articles, and 1449 were screened to include 32 eligible articles in this review (Fig. 1).

Study characteristics

The study characteristics and quality assessments of the included studies are in Table 1. All the studies included a community sample, and 2 included specific groups—youths and adolescents. The studies examined a general population with diverse characteristics (all age groups and gender). Eight studies captured suicide attempts as an outcome, of which 2 included combined data for completed suicides and attempts. There were 12

Fig. 1. Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) diagram showing details of the selection and screening process.



studies conducted in South Korea; 9 in Taiwan (of which 2 included data from China, Hong Kong and South Korea); 4 in Hong Kong (2 of which included data from other countries); 5 in Japan; 1 in Israel and 1 in India. A total of 26 studies were ITS, while 4 studies had a pre-post comparison (multiple arms) and 2 had a cross-sectional design. Studies examined the reporting of suicides in mass media (n=13; print, digital and broadcast); newspapers alone (n=15); newspapers along with television (n=2); newspapers with electronic media (n=1); and television alone (n=1). Specific modes of suicide, such as charcoal burning (n=5), hydrogen sulphide (n=2), and other mixed modes of suicide (n=14), were reported; the remaining articles (n=11) did not specify the role of the mode of suicide. The included studies looked at celebrity (n=22) and non-celebrity (n=10) suicides. The specific characteristics are summarised in Supplementary Table S1. Among the 32 studies, 3 with critical RoB[13,19,20] were excluded from the narrative synthesis.

Overall association between media reporting and risk of copycat suicides and suicide attempts

Meta-analysis

The rate ratio based on the meta-analysis for the association between the media reporting of suicide and actual suicide, ranged from 0.961 to 1.430 among 18 studies. Only 11 out of 18 studies that presented the quantitative data showed statistical significance (Fig. 2A). For the association between media reporting of suicide and suicide attempts, the rate ratio ranged from 1.153 to 1.550 among 2 studies, both of which showed statistical significance (Fig. 2B). The heterogeneity was high (99.47%), and meta-regression failed to identify the source of variation. Therefore, the data from the narrative synthesis were used to summarise the summary of evidence.

Narrative synthesis

Among the 29 studies included in the narrative synthesis, only 2 studies[21,22] showed no association

Table 1. Study characteristics.

Author and year	Country	Design	Study period	Study length	Follow-up time	Data source	Participants	Main outcome	Type of media	Method of suicide	Celebrity	Celebrity type	Celebrity recognition	Number of events
Bakst et al. (2019)	Israel	Pre-post design (multiple arm)	1 Jan 2008–31 Dec 2012	1827 days	28 days	Central Bureau of Statistics	n=2119	Daily suicides	Mass media	Suffocation, firearms, jumping or falling from high-rise buildings, poisoning and burning	No	NA	NA	48 (n)
Chang et al. (2015)	Taiwan	Interrupted time series	27 Jul 2008–31 Dec 2011	1253 days	1253 days	National Death Registry of Taiwan	n=14,920	Weekly suicide rates	Newspaper	Charcoal burning	No	NA	NA	NA
Chen et al. (2010)	Taiwan	Interrupted time series	14 Nov 2008–12 Dec 2008	29 days	29 days	Suicide attempt surveillance database	n=63 (male: 19, female: 44)	Suicide attempts	Mass media	Charcoal burning, drug overdose	Yes	Singer	Local	1 (c)
Chen et al. (2011)	Taiwan	Interrupted time series	Jan 2002–Dec 2005	1461 days	1461 days	Official death records of Taiwan	NR	Daily suicide attempts and rates	Newspaper	NR	No	NA	NA	NR
Chen et al. (2012)	Taiwan	Interrupted time series	2006–2008	1096 days	48 days	Mortality data from Department of Health of the Executive Yuan of Taiwan	n=929; 463 (after reporting) 466 (before reporting)	Biweekly suicide rates	Newspaper	Charcoal burning	Yes	Singer	Local	1 (c) 78 (l)
Chen et al. (2013)	Taiwan	Interrupted time series	1998–2002	1826 days	1826 days	Official death records of Taiwan	NR	Daily suicide rates	Newspaper	Charcoal burning	No	NA	NA	334 (l)
Cheng et al. (2011)	China, Hong Kong, Taiwan	Interrupted time series	1 Jan 2010–1 Dec 2010	365 days	342 days	Media reports from Baidu search engine	n=13	Daily suicide attempts and suicide rates	Mass media	Jumping	No	NA	NA	13 (n), 1279 (l)

Table 1. Study characteristics. (Cont'd)

Author and year	Country	Design	Study period	Study length	Follow-up time	Data source	Participants	Main outcome	Type of media	Method of suicide	Celebrity	Celebrity type	Celebrity recognition	Number of events
Cheng et al. (2007)	Taiwan	Interrupted time series	1 Jan 2003–27 Aug 2005	970 days	118 days	List of repeated suicide attempters from the community, mental health centre	n=270 (n=139 agreed)	Weekly suicide attempts	Mass media	NA	Yes	Actor	Local	1 (c)
Cheng et al. (2007)	Taiwan	Interrupted time series	1 Jan 2003–31 Dec 2005	1096 days	244 days	Department of Health, Taiwan	n=10,945	Weekly counts of suicide	TV, newspaper	Hanging	Yes	Actor	Local	1 (c)
Cheng et al. (2017)	Hong Kong	Interrupted time series	1 Sep 1998–31 Dec 2005	2679 days	2618 days	Hong Kong Coroner's Court and Census & Statistics Department	n=6958	Daily suicide rates	Newspaper	Charcoal and non-charcoal	No	NA	NA	1 (c) 1 (n) 6308 (l)
Cheng et al. (2018)	Hong Kong	Interrupted time series	1 Jan 2003–1 Dec 2016	5114 days	1218 days	Coroner's Court, police investigation files	NR	Daily counts of suicides	Newspaper	Non-specified methods	No	NA	NA	101 (n) 1010 (l)
Choi et al. (2016)	South Korea	Interrupted time series	1 Jan 1997–31 Dec 2009	4748 days	2341 days	Statistics Korea	n=129,571	Monthly suicide rates	Mass media	Jumping, hanging, gas poison, overdose	Yes	Mixed (chief executive officer, actors, singers, entertainers, models, former president)	15 local, 7 international	22 (c)
Fu et al. (2009)	Hong Kong, Taiwan, South Korea	Interrupted time series	1 Jan 2001–1 Jan 2006	1827 days	1006 days	Hong Kong Census and Statistics Department and the Coroner's Court, the Department of Health of Taiwan, and the National Statistical Office of Korea	NR	Weekly suicide rates	Newspaper	Jumping, hanging	Yes	Mixed (pop singer, actors)	Local	3 (c)

Table 1. Study characteristics. (Cont'd)

Author and year	Country	Design	Study period	Study length	Follow-up time	Data source	Participants	Main outcome	Type of media	Method of suicide	Celebrity	Celebrity type	Celebrity recognition	Number of events
Ha and Yang (2021)	South Korea	Interrupted time series	2005–2018	5113 days	10, 20 and 40 days	Mortality microdata from Korean National Statistics Office	n=189,985	Daily suicide rates	Newspaper	Mixed (hanging, fire/ explosion/ gas, drowning, deadly objects/ jumping, unspecified, etc.)	Yes	Mixed (singer, actor, politician and entrepreneur)	NR	13 (c), 99,467 (I) ((I) for celebrity: 47,100; (I) for control: 52,367)
Hagihara et al. (2007)	Japan	Interrupted time series	1 Jan 1987–31 Mar 2005	6665 days	6664 days	Vital and Health Statistics summary	n=455,282	Monthly suicide rates	Newspaper	Non-specified methods	No	NA	NA	Not specified. Monthly mean: 480.33 (I)
Hagihara et al. (2012)	Japan	Interrupted time series	Feb 2003–Dec 2009	2526 days (83 months)	672 days (post-suicide reporting) 598 days (stop of sale)	Vital and Health Statistics summary	NR	Monthly suicide rates	Mass media	Hydrogen sulphide	No	NA	NA	1 (I)
Hagihara et al. (2014)	Japan	Interrupted time series	27 Mar 2008–21 May 2008	56 days	56 days	Fire and Disaster Management Agency, Ministry of Internal Affairs and Communications in Japan	n=145	Daily suicide attempts	Newspaper	Hydrogen sulphide	No	NA	NA	539 (n), 713 (I)
Jang et al. (2016)	South Korea	Interrupted time series	1 Jan 2005–Dec 2008	1461 days	1409 days	Statistics Korea	NR	Monthly suicide rates	Mass media	Hanging, carbon monoxide poisoning	Yes	Mixed (actor, singer, entertainer, model)	NR	7 (c)
Jang (2021)	South Korea	Multiple arm pre-post comparison	1 Dec 2004–31 Jan 2018	4810 days	30 days	Korea National Statistical Office	n=176,336	Daily suicide rates	Mass media	Mixed; hanging, gas, jumping	Yes	Mixed (entertainers, businessmen, politicians, broadcasters, athletes and writers)	NR	24 (c) 3519 (I)

Table 1. Study characteristics. (Cont'd)

Author and year	Country	Design	Study period	Study length	Follow-up time	Data source	Participants	Main outcome	Type of media	Method of suicide	Celebrity	Celebrity type	Celebrity recognition	Number of events
Jeong et al. (2012)	Korea	Pre-post design (multiple arm)	Jan 2005–Dec 2008	1461 days	28 days	National Emergency Department Information System database	n=27,605	Weekly suicide attempts	TV	Hanging, suffocation	Yes	Entertainer	NR	5 (c)
Ju Ji et al. (2014)	South Korea	Interrupted time series	Jan 2003–Dec 2005	1096 days	313 days (28 days)	National Statistical Office of Korea	n=34,237 (male = 23,309, female = 10,928)	Weekly suicide rates	Mass media	Hanging, non-hanging	Yes	Entertainer	Local	1 (c), 1101 (l)
Kim JH et al. (2013)	South Korea	Interrupted time series	May 2007–Oct 2010	1280 days	540 days (30 days)	National Statistical Office of Korea	n=2651 (male = 1595, female = 1056)	Weekly suicide rates	Newspaper	Hanging, jumping, others	Yes	Mixed (politician, entertainer)	Local	2 (c), 1265 (l)
Kim WJ et al. (2013)	South Korea	Pre-post design (multiple arm)	1 Apr 2008–31 Mar 2009	365 days	180 days	Hospital medical records	n=299 (male = 64, female = 235)	Suicide attempts	Mass media	Hanging, self-poisoning, wrist cutting, drowning, gunshot, carbon monoxide poisoning, starving and jumping	Yes	Entertainer	Local	1 (c)
Lee SY (2021)	South Korea	Interrupted time series	1 Jan 2007–31 Dec 2015	3287 days	NR	Statistics Korea	n=128,169	(1) Daily suicide rate (2) Web search frequency	Mass media	All	Yes	Mixed (entertainers, politicians, famous business professionals, sports players, TV anchors, and TV announcers)	Mixed	69 (c), 10,091 (l)

Table 1. Study characteristics. (Cont'd)

Author and year	Country	Design	Study period	Study length	Follow-up time	Data source	Participants	Main outcome	Type of media	Method of suicide	Celebrity	Celebrity type	Celebrity recognition	Number of events
Lirng-jing (2008)	Taiwan	Cross sectional	2 weeks	15 days	54 days	Self-report	n=2576 students	Suicide attempts	Newspaper and electronic media (unspecified)	Unspecified	Yes	Artist (otherwise unspecified)	Local	1
Menon et al. (2021)	India	Cross-sectional	14 Jun 2020–15 Jul 2020	32 days	32 days	Regional and English online news portals	n=1160 (male: 757; female: 379; not mentioned: 24)	Number of suicide cases as captured by news articles	Newspapers (online)	Hanging, poisoning, shooting, drowning, burning, jumping from height, jumping in front of train/vehicle, others	Yes	Entertainer (actor)	Local	1 (c)
Park et al. (2016)	South Korea	Interrupted time series	1 Jan 1996–31 Dec 2010	5479 days	30 and 60 days	National Statistical Office of Korea	n=150,736	Monthly suicide rate	Newspaper	NR	Yes	Mixed (business executive, entertainers)	NR	9 (c)
Suh et al. (2015)	South Korea	Cross-sectional	1 Jan 1991–31 Dec 2010	7305 days	2707 days	Korean Statistical Information Service database	NR	Daily suicide rates	Newspapers, TV	NR	Yes	Mixed (politicians, entertainers, conglomerate)	Local	15 (c)
Ueda et al. (2014)	Japan	Interrupted time series	1 Jan 1989–31 Dec 2010	8035 days	8035 days	Vital Statistics of Japan database (Ministry of Health, Labour and Welfare)	n=596,707	Daily suicide counts	Newspapers	NR	Yes	Mixed (politicians, entertainers, athletes, professors, executives)	Local	109 (c)

Table 1. Study characteristics. (Cont'd)

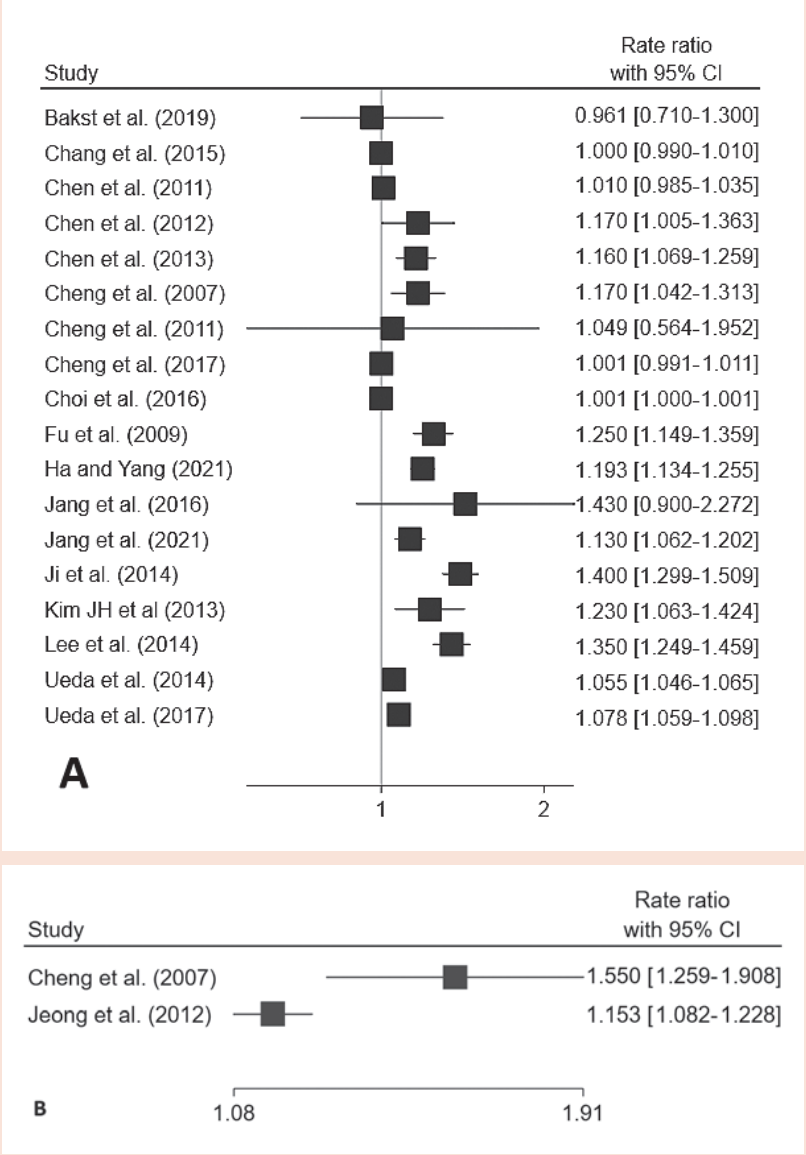
Author and year	Country	Design	Study period	Study length	Follow-up time	Data source	Participants	Main outcome	Type of media	Method of suicide	Celebrity	Celebrity type	Celebrity recognition	Number of events
Ueda et al. (2017)	Japan	Interrupted time series	1 Jul 2010–31 Dec 2014	1645 days	1638 days	Vital Statistics Report database (Ministry of Health, Labour and Welfare)	NR	Daily suicide counts and number of Twitter posts	Mass media	NR	Yes	Mixed (entertainers, business owners, public officials, journalists, athlete, scientist)	Local	26 (c) 985,735 Twitter (l)
Yang et al. (2012)	Taiwan	Interrupted time series	1 Jan 2003–31 Dec 2010	2922 days	2922 days	Collaboration Centre of Health Information Application, Department of Health database	n=31,364 (male = 21,484; female = 9880)	Monthly suicide counts	Mass media	Hanging, jumping from heights, poisoning, drug intoxication, charcoal burning, others (unspecified)	Yes	Entertainers	NR	2 (c) 1 (n) 16,795 (l)
Yi et al. (2019)	South Korea	Interrupted time series	1 Jan 1993–31 Dec 2013	7670 days	1916 days	Korean National Statistics Office database	NR	Suicide rates (expected and actual)	Newspaper	NR	Yes	Mixed (actors, politicians, businessmen, TV anchors)	NR	10 (c, suicide) 6 (c, control)

NA: not applicable; NR: not reported; (c): number of celebrities; (n): number of non-celebrities; (l): number of media reports

between media reporting and suicide rates, while the majority found that media reporting of suicides could trigger subsequent copycat suicides among the public. A marked increase in suicide attempts following the media report of suicides was noted in all 6 studies that reported suicide attempts separately. The 2 studies that showed no association involved mainly non-celebrity suicides that were publicised by the media. Cheng et al.[22] looked at the point clusters of copycat suicides among a local network rather than mass clusters (general population) or among those who shared similar risk factors. Bakst et al., on the other hand, focused on a diverse sample of "publicised" suicides involving dentists, police officers, social activists and celebrities, making it a diverse mixed group.[21] There were no other differences between these 2 studies and the rest of the articles that could explain this discrepancy. Overall, the negative influence of media reporting (higher rates of suicides or attempted suicides) was evident regardless of country, celebrity status, study design, type of media, mode of suicide, length of follow-up period, or RoB rating. Many of the suicides and attempted suicides happened within the 4 weeks of media reporting, with higher rates observed in earlier weeks.[17,23-34]

Univariable meta-regressions were run to determine the heterogeneity observed among different studies. None of the variables (i.e. study,

Fig. 2. Forest plot showing the association between media reporting of suicide and actual suicide (A) and suicide attempts (B).



CI: confidence interval

design, publication year, confounders adjusted, and study length [per 10,000 days]) were significant enough to explain the source of heterogeneity.

Factors affecting copycat suicides

Sociodemographic characteristics

The impact of sociodemographic characteristics (age and gender) of the participants on copycat suicides was explored in 15 studies. Twelve studies reported gender differences in copycat suicides.[10,23,24,28,30,35-37] A higher risk of copycat suicides among females was reported in 8 studies[10,23,24,30,31,35,36,38] and among males in four studies.[26-27,39-40] Three studies showed that the risk of copycat suicides or suicide attempts is similar in both genders.[31,39,41] Many studies found that younger age groups are at a higher risk of copycat suicides following media reporting.[10,24,27,28,30,37,40]

Idol's characteristics

The impact of celebrity suicides on the rates of copycat suicides was reported in 22 studies, and all of them evidenced a Werther effect. Celebrity death by suicide as opposed to other causes of death, such as accidents, was associated with higher subsequent suicide rates in 2 studies.[28,37] A comparison of the consequences of media reporting of entertainers' suicides as opposed to non-entertainers' suicides was reported in 5 studies showing mixed results.[31,33,34,37,38] The role of an idol's characteristics (age, gender) in copycat suicides was noted by 7 studies showing a higher Werther effect in those who share similar characteristics as the idol.[10,17,30,31,33,35,37] Thirteen studies showed that reporting the method of suicide of the idol could lead to higher rates of copycat suicides following the same methods.[10,17,23-28,30,31,42-44]

Vulnerability of the participants

The association of vulnerability to the Werther effect was assessed by 3 articles and showed that those who had a previous history of suicide ideation, attempts, stress, negative attitudes towards life (life despair), depression, and financial problems had a higher tendency for copycat suicides or attempts.[27,32,39]

Media-specific factors

The association between media-specific factors and subsequent suicides was analysed in 3 studies. Competition among media could trigger an increase in the reporting intensity of suicides, which could lead to a higher rate of copycat suicides, as reported by Chen et al.[45] Furthermore, the number of reports could affect copycat suicides. For example, more than 10,000 social media posts featuring young celebrities were associated with higher copycat deaths.[46] High-intensity descriptive reporting resulted in higher copycat suicides, while preventive reporting reduced the suicides.[47] There was only one article among the included studies that described a positive outcome (coping) of the media reporting.[47] The study showed that preventive reporting (i.e. sharing stories of how people overcame suicide ideations, introducing community resources for suicide prevention and coping) had a protective effect on copycat suicides.

Additional information, including the findings from the individual studies, is included in Supplementary Table S2. A summary of the potential factors is shown in Fig. 3.

RoB and strength of evidence

Among the included studies, there were 3 with critical, 10 with serious and 19 with moderate RoB. Those with critical RoB were excluded from the narrative synthesis. Therefore, the synthesis includes 29 articles. The RoB of the included studies is shown in Fig. 5. A detailed list of individual domains of RoB is included in Supplementary Table S3. Based on the results and quality of the evidence, there was indeed a copycat effect following media reporting of suicide.

Publication bias

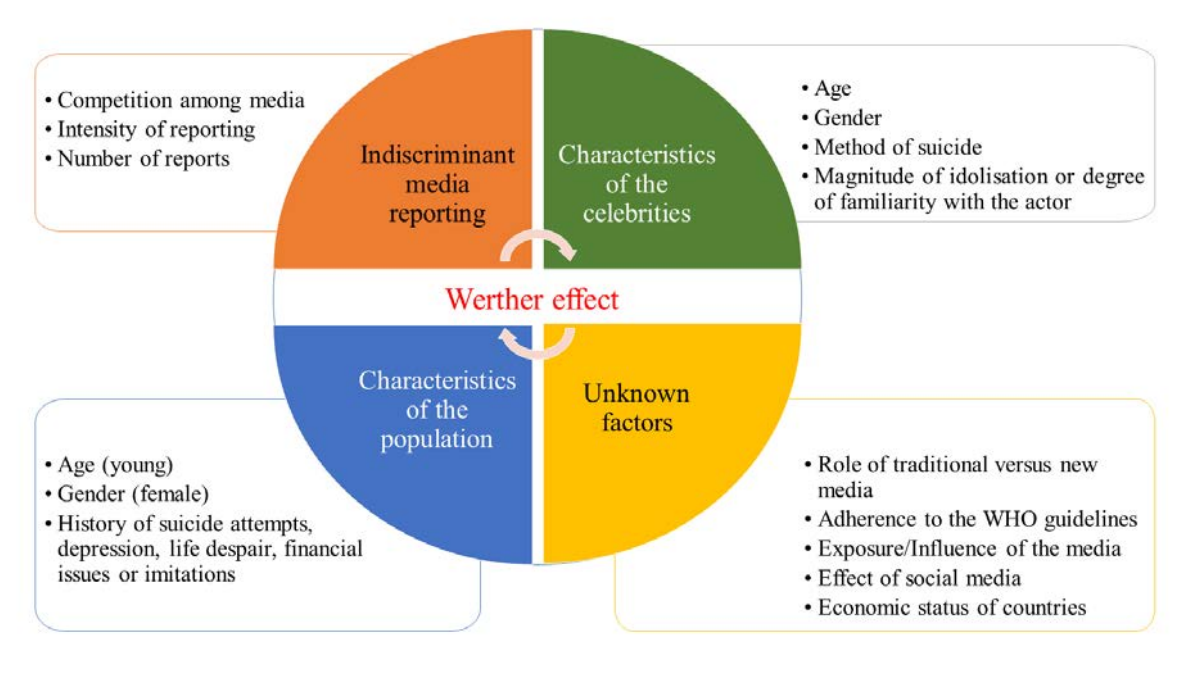
We performed publication bias assessments on 18 studies, as 8 studies did not have relevant outcome measures. The asymmetric funnel plots (Fig. 5) indicated publication bias. The bias was confirmed by Egger's regression test (P=0.03).

DISCUSSION

Our systematic review found substantial evidence that media's portrayal of suicide was associated with higher rates of copycat suicides. The evidence showed a positive trend but was not sufficient for suicide attempts due to the small number of studies examining this outcome. It was also noted that the population (age and gender) and characteristics of the deceased (age, gender, suicide method) were associated with the copycat deaths.

Previous systematic reviews that looked at studies across the world shared similar findings in their meta-analyses.[16,48] The previous meta-analysis showed a 13% higher suicide risk following the media report of suicide while an 18% risk was noted when the media report involved a fictional portrayal of suicide.[16] We have noted similar findings in the current review, with the majority of the studies showing a higher rate of copycat suicides and attempts following publicised suicides. This shows

Fig. 3. Factors associated with higher rates of copycat suicides.



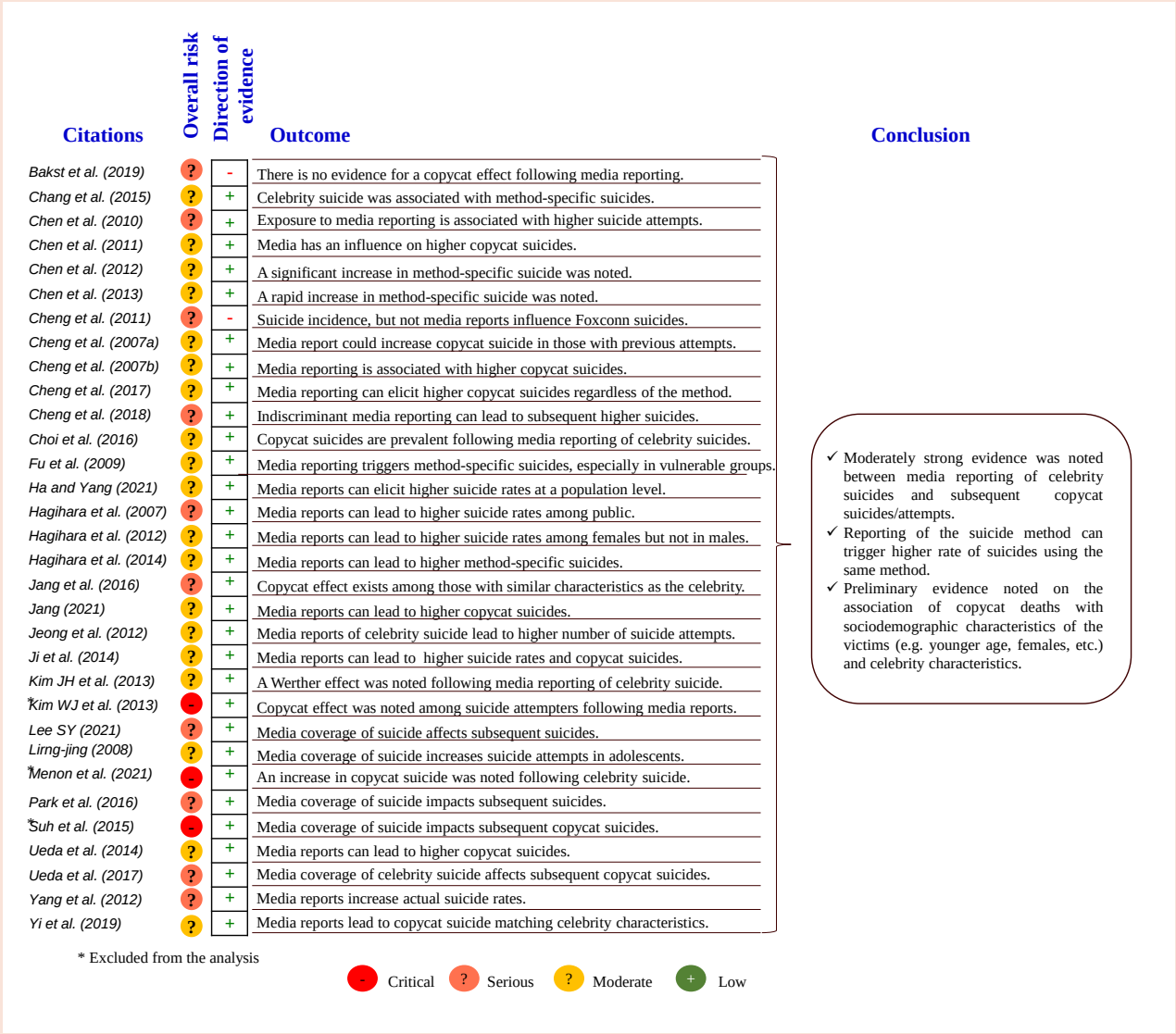
that the Werther effect exists in Asian countries, which is a finding of great concern given the higher suicide statistics in Asian countries.

Almost all the included studies supported the association between media reporting of suicide and subsequent copycat suicides. Copycat suicides commonly occur in individuals with strong worship or idolisation of another person (a role model or celebrity).[35,49] People often follow the lifestyle, looks, clothing and mannerisms of a specific person (e.g. celebrity, writer or politician) out of devotion,[50,51] which leads to a fictional attachment to them even with no direct contact. This attachment, especially among those with underlying vulnerability (depression or anxiety) or a history of previous suicide attempts or self-harm, trigger copycat suicides imitating the same method when their idol dies.[35,49] Thus, identifying with the person or celebrity in the media reports triggers a Werther effect, especially in susceptible people, based on the degree of vulnerability and idolisation they have.

We identified sociodemographic characteristics of the population, characteristics of the idol, vulnerability of individuals, and media-specific factors to be associated with the Werther effect. Younger age groups and predominantly females were more susceptible to the Werther effect, as noted in the current review, which was corroborated by others.[13,52] While 8 studies reported female gender to be associated with copycat suicides, 4 showed an association with male gender. While the RoB between these studies remained comparable, the studies that found an effect of male gender were from Taiwan. Thus, the reason for this discrepancy could be related to a specific population. Research has also shown that the magnitude of the copycat effect (2.31-fold) and the mortality of copycat suicides (22.7-fold) are higher in younger females (20–29 years).[37] This could be explained based on psychological vulnerability, where younger females tend to have more susceptibility to mental health conditions,[53] especially mood disorders and obsessive-compulsive disorder.[54] We had also noted a higher copycat effect in the population that shared similar characteristics as the idol. While none of the previous systematic reviews looked at these characteristics, this similarity was noted in many other studies.[37]

Only a single study in the current review linked the copycat effects with non-adherence to the World Health Organization's reporting guidelines.[43] Research showed that media coverage of the famous actor Robin Williams' suicide resulted in a 10% increase in suicides in the US. Furthermore, analysis of the media reports showed that 46% reported the method of suicide, 27% of articles romanticised suicide, and only 11% included information on help-seeking.[55] A report from India

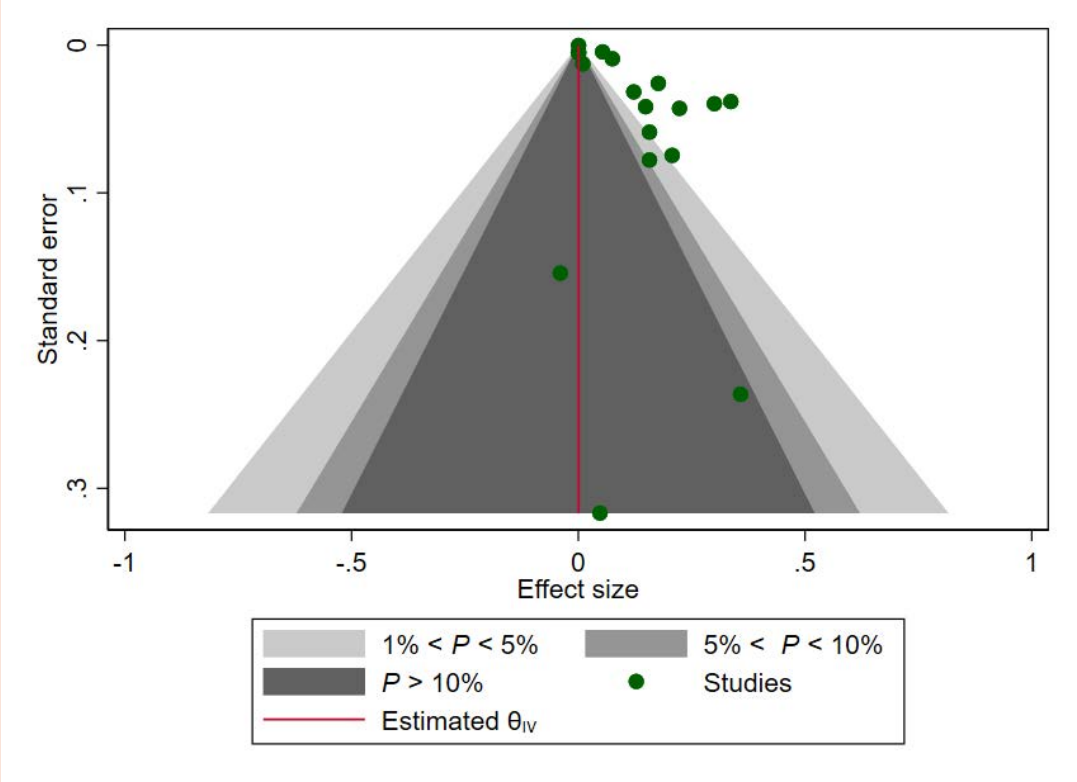
Fig. 4. RoB and the evidence.



showed that all the news articles indicated the name of the victim, and 10% included the method and location of suicide in the title, while 80% reported the method and location in the main contents with no guidance for help-seeking or coping.[56] Given that media reporting has a strong impact on the Werther effect, reports should emphasise more on help-seeking and coping strategies, rather than publicising or romanticising the suicides. The majority of Asian countries have no national-level guidelines for suicide prevention.[57] Thus, along with the attempts to enforce adherence to the reporting guidelines, there is an urgent need to identify the population at risk and develop specific public health suicide prevention interventions for them in the wake of a celebrity suicide.

The role of specific factors that could have a significant influence on the extent of the Werther effect, such as the recognition of celebrities (local vs international), gender and age of the celebrities (male vs female; younger vs older), magnitude of fame (popular actresses vs TV anchors), type of celebrities (entertainers vs non-entertainers), and the media coverage that different types of celebrity suicides (e.g. entertainers) receive, are currently unknown or unclear. To our knowledge, only a limited number of studies have addressed these factors. Similarly, the impact of media reporting through electronic media vs print media and social media was not taken into account by the included articles. In particular, since social media is used by many people in the world, the platform represents a new face of global networking and thus affects people's social behaviour. Early reports showed that specific emotional responses to social media reports of celebrity suicides were associated with a

Fig. 5. Funnel plot for suicide rates indicating publication bias.



subsequent rise in national-level suicides.[58] On the other hand, Shoib et al. noted that social media could also have a positive impact by providing the necessary help-seeking and support online, along with potential predictions on the suicide behaviours of the users, aiding in prevention.[59] Given the higher penetration and accessibility of social media, its influence on copycat suicides should be further explored. The analysis in the current review showed evidence for publication bias, which suggests that the evidence could be weaker than reported. However, in the absence of unpublished data to compare and contrast, the extent of publication bias and its impact cannot be determined.

The current review includes data mainly from East Asia, and there is a dearth of literature from many Asian countries (e.g. India, China, Singapore and Thailand). Current data represent high-income countries, which do not give a clear understan-ding of the Asian situation. Hence, it is not possible to compare how the Werther effect varies between different Asian regions (e.g. Southeast Asia vs East Asia). There are reports on the extreme worship and huge fan base of celebrities in India.[8] However, minimal research was conducted in this area. Menon et al.[13] showed that there was an increase in news articles about copycat suicides among the public following the death of a popular actor. More studies representing the country need to be done in view of the rising prevalence of celebrity suicides in India[60] which would provide further clarity to the country's position in the Asian situation.

The strengths of the review include its robust search strategy through multiple databases, including Chinese literature databases, and its adherence to Cochrane standards. The comprehen-sive synthesis methods and quality assessments further add to the strengths of the review. The high-level heterogeneity within the studies, publication bias, and the inability to establish a causation effect between the media reports and subsequent deaths are the limitations of the review. Also, some studies had overlapping data, and it was difficult to exclude the overlapping data in instances where the identity of the celebrity was masked. We have tried to contact the authors for clarifications and have excluded studies where data overlap could be confirmed. The homogenous design could have led to the loss of important data from case series, reports and qualitative studies.

Given the high heterogeneity of data in the existing literature, future studies should employ robust study designs, valid outcome measures, and quantitative reporting of results (rate ratios, odds ratios or standardised mortality ratios) to facilitate future meta-analysis and thus advance the reach of

the data. The studies should also employ vigorous data collection methods that address potential confounders, consider cultural and contextual variations, and examine the effects of various media and celebrity-related factors (i.e. impact of celebrity gender, entertainers vs non-entertainers, suicide vs other modes of death, electronic media vs traditional media vs social media, other celebrity characteristics and sensationalised vs non-sensationalised reporting) on suicide outcomes. Longitudinal studies should be conducted in diverse populations (males, females, adolescents, adults, older adults and children) to track individuals' exposure to suicide-related media content from various specific types of media over time and assess its impacts on their mental health behaviours. These details are indispensable for preventive efforts to design new guidelines or interventions to curtail this growing challenge.

Even with years of lobbying to enforce compliant reporting, non-compliant reporting by the media persists with lucrative business outcomes in mind. This persistent problem underscores the necessity for collaboration between researchers, media professionals, professional medical bodies and mental health experts to create awareness among media professionals about the public health impact of non-adherent reporting and implement guidelines for responsible reporting on suicide. Proactive involvement of healthcare systems and community services to raise population-wide awareness about copycat suicides and to promote early help-seeking behaviour in the context of suicide prevention is a necessity. Mental healthcare organisations serve as critical anchors for early identification, assessment, intervention and follow-up care for vulnerable subjects, offering them a comprehensive suicide prevention platform. Long-term support and crisis intervention, together with collaborative efforts among various professionals, might ultimately contribute to improved public health outcomes and lower suicide rates.

CONCLUSION

The current systematic review showed higher suicide rates among the public following media reporting of suicides in Asian countries. While the meta-analysis showed high heterogeneity and little evidence, the narrative synthesis of 29 studies showed modest evidence with the reporting of suicide methods shown to have a negative impact on subsequent suicides. Factors identified as potentially associated with a higher risk of copycat suicides were younger age and female gender. This also calls for stringent adherence to the reporting guidelines and the development of national-level, stringent guidelines for suicide by media professionals, which include not romanticising the suicide or giving unwanted attention to the details of the event as a social responsibility. Since many Asian countries do not have national-level guidelines for suicide prevention, media professionals and policymakers should take up the responsibility to reduce the cluster effect due to this avoidable cause. Future studies should explore additional factors specific to Asia in greater detail. In addition to reporting guidelines for the media, a collaborative effort between healthcare systems, media professionals and other community partners to raise awareness and administer early interventions targeting those at risk may also be helpful.

Acknowledgement

The authors would like to thank Zhang YunJue for assisting in the Chinese literature database search.

REFERENCES

1. World Health Organization. Suicide, Geneva: 28 August 2023. https://www.who.int/news-room/fact-sheets/detail/suicide. Accessed 30 June 2023.
2. World Health Organization. Preventing suicide: A global imperative, 17 August 2014. https://www.who.int/publications/i/item/9789241564779. Accessed 30 June 2023.
3. World Health Organization. One in 100 deaths is by suicide: WHO guidance to help the world reach the target of reducing suicide rate by 1/3 by 2030, 17 June 2021. https://www.who.int/news/item/17-06-2021-one-in-100-deaths-is-by-suicide. Accessed 30 June 2023.
4. Vijayakumar L, Daly C, Arafat Y, et al. Suicide Prevention in the Southeast Asia Region. Crisis: The Journal of Crisis Intervention and Suicide Prevention 2020;41:S21-29.
5. Vijayakumar L. Altruistic suicide in India. Arch Suicide Res 2004;8:73-80.
6. Balodhi JP. Indian mythological views on suicide. NIMHANS Journal 1992;10:101-05.
7. Sangari K, Vaid S. Sati in modern India: A report. Econ Polit Wkly, 1981;16:1284-8.
8. Ram S. Temples in India Dedicated to Politicians and Celebrities. Updated 25 April 2018. https://www.nativeplanet.com/travel-guide/temples-in-india-dedicated-politicians-celebrities-002216.html. Accessed 7 March 2023.
9. Levy MS. Celebrity and Entertainment Obsession Understanding Our Addiction. Maryland: Rowman & Littlefield Publishers; 2015.
10. Jang SA, Sung JM, Park JY, et al. Copycat suicide induced by entertainment celebrity suicides in South Korea. Psychiatry Investig 2016;13:74-81.
11. Sinyor M, Williams M, Zaheer R, et al. The association between Twitter content and suicide. Aust N Z J Psychiatry 2021;55:268-76.
12. Fu KW, Chan CH. A Study of the Impact of Thirteen Celebrity Suicides on Subsequent Suicide Rates in South Korea from 2005 to 2009. PLOS ONE 2013;8:e53870.
13. Menon V, Kar SK, Marthoenis M, et al. Is there any link between celebrity suicide and further suicidal behaviour in India? Int J Soc Psychiatry 2021;67:453-60.
14. Yip PS, Fu KW, Yang KC, et al. The effects of a celebrity suicide on suicide rates in Hong Kong. J Affect Disord 2006;93:245-52.

15. McTernan N, Spillane A, Cully G, et al. Media reporting of suicide and adherence to media guidelines. Int J Soc Psychiatry 2018;64:536-44.
16. Niederkrotenthaler T, Braun M, Pirkis J, et al. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. BMJ (Clinical research ed) 2020;368:m575.
17. Fu KW, Yip PSF. Estimating the Risk for Suicide Following the Suicide Deaths of 3 Asian Entertainment Celebrities: A Meta-Analytic Approach. J Clin Psychiatry 2009;70:869-78.
18. United Nations. Methodology: Standard country or area codes for statistical use (M49). https://unstats.un.org/unsd/methodology/m49/. Accessed 8 March 2023.
19. Kim WJ, Song YJ, Namkoong K, et al. Does a copycat effect exist in the emergency department? Int J Psychiatry Med 2013;45:59-72.
20. Suh S, Chang Y, Kim N. Quantitative exponential modelling of copycat suicides: Association with mass media effect in South Korea. Epidemiol Psychiatr Sci 2015;24:150-7.
21. Bakst SS, Berchenko Y, Braun T, et al. The Effects of Publicized Suicide Deaths on Subsequent Suicide Counts in Israel. Arch Suicide Res 2019;23:440-54.
22. Cheng Q, Chen F, Yip PS. The Foxconn suicides and their media prominence: is the Werther Effect applicable in China? BMC Public Health 2011;11:841.
23. Chang SS, Kwok SS, Cheng Q, et al. The association of trends in charcoal-burning suicide with Google search and newspaper reporting in Taiwan: a time series analysis. Soc Psychiatry Psychiatr Epidemiol 2015;50:1451-61.
24. Chen YY, Liao SF, Teng PR, et al. The impact of media reporting of the suicide of a singer on suicide rates in Taiwan. Soc Psychiatry Psychiatr Epidemiol 2012;47:215-21.
25. Chen Y-Y, Chen F, Gunnell D, et al. The Impact of Media Reporting on the Emergence of Charcoal Burning Suicide in Taiwan. PLOS ONE 2013;8:e55000.
26. Cheng AT, Hawton K, Chen TH, et al. The influence of media coverage of a celebrity suicide on subsequent suicide attempts. J Clin Psychiatry 2007;68:862-6.
27. Cheng ATA, Hawton K, Lee CTC, et al. The influence of media reporting of the suicide of a celebrity on suicide rates: A population-based study. Int J Epidemiol 2007;36:1229-34.
28. Ha J, Yang HS. The Werther effect of celebrity suicides: Evidence from South Korea. PLOS ONE 2021;16.
29. Jeong J, Shin SD, Kim H, et al. The effects of celebrity suicide on copycat suicide attempt: a multi-center observational study. Soc Psychiatry Psychiatr Epidemiol 2012;47:957-65.
30. Ju Ji N, Young Lee W, Seok Noh M, et al. The impact of indiscriminate media coverage of a celebrity suicide on a society with a high suicide rate: Epidemiological findings on copycat suicides from South Korea. J Affect Disord 2014;156:56-61.
31. Kim JH, Park EC, Nam JM, et al. The werther effect of two celebrity suicides: An entertainer and a politician. PLOS ONE 2013;8:e84876.
32. Lee SY. Do effects of copycat suicides vary with the reasons for celebrity suicides reported by the media? Soc Sci J 2021;58:106-18.
33. Park J, Choi N, Kim SJ, et al. The Impact of Celebrity Suicide on Subsequent Suicide Rates in the General Population of Korea from 1990 to 2010. J Korean Med Sci 2016;31:598-603.
34. Ueda M, Mori K, Matsubayashi T. The effects of media reports of suicides by well-known figures between 1989 and 2010 in Japan. Int J Epidemiol 2014;43:623-9.
35. Choi YJ, Oh H. Does Media Coverage of a Celebrity Suicide Trigger Copycat Suicides? Evidence from Korean Cases. J Media Econ 2016;29:92-105.
36. Hagihara A, Abe T. Effects of media reports and the subsequent voluntary withdrawal from sale of suicide-related products on the suicide rate in Japan. Eur Arch Psychiatry Clin Neurosci 2012;262:245-51.
37. Yi H, Hwang J, Bae H-J, et al. Age and sex subgroups vulnerable to copycat suicide: evaluation of nationwide data in South Korea. Sci Rep 2019;9:17253.
38. Jang J, Myung W, Kim S, et al. Effect of suicide prevention law and media guidelines on copycat suicide of general population following celebrity suicides in South Korea, 2005-2017. Aust N Z J Psychiatry 2022;56:542-50.
39. Lirng-jing P. Perceived Impacts of Media Suicide Reports upon Mental Health among Adolescents: Taking the Suicide of One Actor as an Example. Bull Educ Psychol 2008;39:355-76.
40. Chen YY, Tsai PC, Chen PH, et al. Effect of media reporting of the suicide of a singer in Taiwan: the case of Ivy Li. Soc Psychiatry Psychiatr Epidemiol 2010;45:363-9.
41. Hagihara A, Tarumi K, Abe T. Media suicide-reports, Internet use and the occurrence of suicides between 1987 and 2005 in Japan. BMC Public Health 2007;7:321.
42. Cheng Q, Chen F, Yip PSF. Media effects on suicide methods: A case study on Hong Kong 1998-2005. PLOS ONE 2017;12:e0175580.
43. Hagihara A, Abe T, Omagari M, et al. The impact of newspaper reporting of hydrogen sulfide suicide on imitative suicide attempts in Japan. Soc Psychiatry Psychiatr Epidemiol 2014;49:221-9.
44. Yang AC, Tsai SJ, Yang CH, et al. Suicide and media reporting: a longitudinal and spatial analysis. Soc Psychiatry Psychiatr Epidemiol 2013;48:427-35.
45. Chen YY, Chen F Fau - Yip PSF, Yip PS. The impact of media reporting of suicide on actual suicides in Taiwan, 2002-05. J Epidemiol Community Health 2011;65:934-40.
46. Ueda M, Mori K, Matsubayashi T, et al. Tweeting celebrity suicides: Users' reaction to prominent suicide deaths on Twitter and subsequent increases in actual suicides. Soc Sci Med 2017;189:158-66.
47. Cheng Q, Chen F, Lee EST, et al. The role of media in preventing student suicides: A Hong Kong experience. J Affect Disor 2018;227:643-8.
48. Niederkrotenthaler T, Kirchner S, Till B, et al. Systematic review and meta-analyses of suicidal outcomes following fictional portrayals of suicide and suicide attempt in entertainment media. EClinicalMedicine 2021;36:100922.
49. Cheng AT, Hawton K, Chen TH, et al. The influence of media reporting of a celebrity suicide on suicidal behavior in patients with a history of depressive disorder. J Affect Disord 2007;103:69-75.
50. Chan K, Prendergast GP. Social comparison, imitation of celebrity models and materialism among Chinese youth. Int J Advert 2008;27:799-826.
51. Elliott, A. (2011). 'I Want to Look Like That!': Cosmetic Surgery and Celebrity Culture. Cult Sociol 2011;5:463-77.
52. Huang CY, Huang YT, Lin YH, et al. Factors associated with psychological impact of celebrity suicide media coverage: An online survey study. J Affect Disord 2021;295:839-45.
53. Subramaniam M, Abdin E, Vaingankar JA, et al. Tracking the mental health of a nation: prevalence and correlates of mental disorders in the second Singapore mental health study. Epidemiol Psychiatr Sci 2019;29:e29.

54. Picco L, Subramaniam M, Abdin E, et al. Gender differences in major depressive disorder: findings from the Singapore Mental Health Study. Singapore Med J 2017;58:649-55.
55. Carmichael V, Whitley R. Media coverage of Robin Williams' suicide in the United States: A contributor to contagion? PLOS ONE 2019;14:e0216543.
56. Menon V, Arafat SY, Akter H, et al. Cross-country comparison of media reporting of celebrity suicide in the immediate week: A pilot study. Asian J Psychiatr 2020;54:102302.
57. Ortiz-Ospina E. The rise of social media, 18 September 2019. https://ourworldindata.org/rise-of-social-media. Accessed 21 April 2023.
58. Fahey RA, Matsubayashi T, Ueda M. Tracking the Werther Effect on social media: Emotional responses to prominent suicide deaths on twitter and subsequent increases in suicide. Soc Sci Med 2018;219:19-29.
59. Shoib S, Chandradasa M, Nahidi M, et al. Facebook and Suicidal Behaviour: User Experiences of Suicide Notes, Live-Streaming, Grieving and Preventive Strategies-A Scoping Review. Int J Environ Health Res Public Health 2022;19:13001.
60. Kar SK, Arafat SY, Ransing R, et al. Repeated celebrity suicide in India during COVID-19 crisis: An urgent call for attention. Asian J Psychiatr 2020;53:102382.

Ann Acad Med Singap 2024;53:170-86
https://doi.org/10.47102/annals-acadmedsg.2023391

REVIEW ARTICLE

Singapore tuberculosis (TB) clinical management guidelines 2024: A modified Delphi adaptation of international guidelines for drug-susceptible TB infection and pulmonary disease

The Clinical Tuberculosis Guidelines Development Team comprises the Clinical Tuberculosis Guidelines Panel and the Public Health Translational Team.

The Panel comprises, in alphabetical order:

Ang Michelle Lay Teng[1] *PhD*, Chan Si Min[2,3] *MBBS*, Cheng Lionel Tim-Ee[4] *MBBS*, Cheong Hau Yiang[5] *MBBS*, Chew Ka Lip[6] *MBBS*, Chlebicki Piotr Maciej[7] *MBBS*, Hsu Li Yang[8,9] *MPH*, Kaw Gregory Jon Leng[10] *MBBS*, Kee Adrian Chin Leong[9,11] *MBBS*, Ng Mark Chung Wai[12] *MBBS*, Ong Rick Twee Hee[8] *PhD*, Ong Catherine Wei Min[9,11] *PhD*, Quah Jessica Lishan[13] *MBBS*, Selvamani D/O Balasubramaniam[14] *MBBS*, Sng Li Hwei[15] *MBBS*, Tan Jamie Bee Xian[15] *MBBS*, Tan Cher Heng[10] *MBBS*, Tay Jun Yang[16] *MBBS*, Teo Lynette Li San[17,18] *MBBS*, Thoon Koh Cheng[19] *MBBS*, Yan Gabriel Zherong[6] *MBBS*

The Public Health Translational Team comprises, in alphabetical order:

Chen Jacinta I-Pei[8] *MA*, Hud bin Mohammed Helmi[8], Khoo Benjamin Bing Jie[8], Lee Dawn Yi Xin[8], Ng Bob Xian Yi[8] *BBA*, Park Jia Ying[8] *BSS*, Tan Belinda Ying Ting[8] *BSc*, Yang Qian[8] *MSc*

ABSTRACT

Introduction: Tuberculosis (TB) remains endemic in Singapore. Singapore's clinical practice guidelines for the management of tuberculosis were first published in 2016. Since then, there have been major new advances in the clinical management of TB, ranging from diagnostics to new drugs and treatment regimens. The National TB Programme convened a multidisciplinary panel to update guidelines for the clinical management of drug-susceptible TB infection and disease in Singapore, contextualising current evidence for local practice.

Method: Following the ADAPTE framework, the panel systematically reviewed, scored and synthesised English-language national and international TB clinical guidelines published from 2016, adapting recommendations for a prioritised list of clinical decisions. For questions related to more recent advances, an additional primary literature review was conducted via a targeted search approach. A 2-round modified Delphi process was implemented to achieve consensus for each recommendation, with a final round of edits after consultation with external stakeholders.

Results: Recommendations for 25 clinical questions spanning screening, diagnosis, selection of drug regimen, monitoring and follow-up of TB infection and disease were formulated. The availability of results from recent clinical trials led to the inclusion of shorter treatment regimens for TB infection and disease, as well as consensus positions on the role of newer technologies, such as computer-aided detection-artificial intelligence products for radiological screening of TB disease, next-generation sequencing for drug-susceptibility testing, and video observation of treatment.

Conclusion: The panel updated recommendations on the management of drug-susceptible TB infection and disease in Singapore.

Ann Acad Med Singap 2024;53:170-86

Keywords: clinical practice guidelines, directly observed treatment, drug-susceptible tuberculosis, infectious diseases, pulmonary, registry, video-observed treatment

[1] National Public Health Laboratory, National Centre for Infectious Diseases, Singapore
[2] Division of Paediatric Infectious Diseases, Khoo Teck Puat-National University Children's Medical Institute, National University Hospital, Singapore
[3] Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[4] Department of Diagnostic Radiology, Singapore General Hospital, Singapore
[5] Department of Infectious Diseases, Changi General Hospital, Singapore
[6] Department of Laboratory Medicine, National University Hospital, Singapore
[7] Department of Infectious Diseases, Singapore General Hospital, Singapore
[8] Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[9] Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[10] Diagnostic Radiology, Tan Tock Seng Hospital, Singapore
[11] Department of Medicine, National University Hospital, Singapore
[12] SingHealth Polyclinics, Singapore Health Services, Singapore
[13] Department of Respiratory & Critical Care Medicine, Changi General Hospital, Singapore
[14] Bukit Panjang Polyclinic, National University Polyclinics, Singapore
[15] Department of Microbiology, Singapore General Hospital, Singapore
[16] Department of Infectious Diseases, National Centre for Infectious Diseases, Singapore
[17] Department of Diagnostic Imaging, National University Hospital, Singapore
[18] Diagnostic Radiology, Yong Loo Lin School of Medicine, Singapore
[19] Division of Medicine, KK Women's & Children's Hospital, Singapore
Correspondence: Dr Li Yang Hsu, Saw Swee Hock School of Public Health, National University of Singapore, 12 Science Drive 2, #10-01, Singapore 117549.
Email: ephhly@nus.edu.sg

CLINICAL IMPACT

What is New

- This update to the 2016 Ministry of Health Clinical Practice Guidelines on the prevention, diagnosis and management of tuberculosis (TB) in Singapore incorporates new shortened regimens and recommendations on major new advances.
- Major new advances include the use of computer-aided detection-artificial intelligence products for radiological screening for TB disease, next-generation sequencing for drug susceptibility testing of *Mycobacterium tuberculosis*, and video-observed treatment of TB disease in lieu of direct observation.

Clinical implications

- These guidelines will largely be implemented by the National TB Programme team in Singapore.
- The primary aim is to standardise and improve the clinical management of TB infection and disease in Singapore, contributing to the global goal of TB elimination.

INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by the *Mycobacterium tuberculosis* complex. For decades, it was the leading cause of death worldwide from a single infectious disease before being displaced by COVID-19 during the pandemic years.[1]

TB is endemic in Singapore, with over 2000 cases of TB disease (formerly active TB) diagnosed each year. The prevalence of TB infection (formerly latent TB) in the Singapore resident population was recently estimated to be 12.7%, ranging from 2.4% in young adults (18–29 years) to 23.2% in the elderly (70–79 years).[2] Critical to the elimination of TB are early diagnosis and treatment of TB disease as well as preventive treatment of those with TB infection who are at risk of progression to disease.[1]

Singapore's clinical practice guidelines for the management of TB was first published in 2016.[3] Over the past 7 years, there have been major new advances in the clinical management of TB, including but not limited to the use of computer-aided detection-artificial intelligence (CAD-AI) products for radiological screening for TB disease,[4,5] next-generation sequencing (NGS) for drug susceptibility testing of *M. tuberculosis*,[6,7] new drug treatment regimens incorporating rifapentine for TB disease[8-11] and infection,[8,12-18] and video-observed treatment (VOT) of TB disease in lieu of directly observed treatment (DOT).[19,20]

Scope

The National TB Programme commissioned this update with the aim of providing healthcare professionals in Singapore with evidence-based and contextualised best practices for screening, diagnosis and treatment of both adults and children with drug-susceptible TB infection and pulmonary disease. These guidelines are not intended to replace the clinical judgment of the healthcare practitioner.

Recommendations for the clinical management of extrapulmonary TB disease as well as drug-resistant TB infection and disease are beyond the scope of these guidelines. All persons with rifampicin- or multidrug-resistant TB should be referred to the TB Control Unit (TBCU) at Tan Tock Seng Hospital for further management.

METHOD

Guidelines development team

The Clinical Tuberculosis Guidelines Panel (Panel) was formed in August 2022, and was supported by the Public Health Translational Team (PHTT) from the Saw Swee Hock School of Public Health.

Selection of questions

The Panel agreed on the clinical questions for recommendations, based on prioritisation of the key clinical decisions in managing drug-susceptible TB infection and disease.

Development strategy

Adaptation of guidelines in accordance with the ADAPTE framework was performed for the majority of this update.[21] A copy of the RIGHT-Ad@pt checklist[22] is provided in the Supplementary Materials Appendix S9.

The guidelines adaptation process is described in Supplementary Materials Appendix S1. All major national and international TB management guidelines published in the English language between 1 January 2016 and 5 March 2023 are listed in Table 1. These guidelines were assessed with a modified Appraisal of Guidelines for Research & Evaluation II (AGREE II) instrument,[23] with assessment results provided in Supplementary Materials Appendix S1. PHTT conducted an additional primary literature review using a targeted search approach for questions regarding recent advancements like CAD-AI, NGS and VOT.

PHTT summarised guidelines and primary literature into reports for the Panel. Through a

Table 1. List of international guidelines reviewed and used in the preparation of the 2024 TB guidelines.

Guidelines	Year published	Developing organisations	Region	Funder	Abbreviation[Ref]
European Union Standards for Tuberculosis Care - 2017 update	2018	European Respiratory Society and European Centre for Disease Prevention and Control (ECDC)	European Union	European Respiratory Society	ECDC-S[24]
Scientific Advice: Programmatic management of latent tuberculosis infection in the European Union	2018	European Respiratory Society and European Centre for Disease Prevention and Control	European Union	European Respiratory Society	ECDC-L[13]
Tuberculosis: NICE guideline	2019	National Institute for Health and Care Excellence (UK)	United Kingdom (UK)	UK government	NICE[25]
Guidelines for Tuberculosis Control in New Zealand, 2019	2019	Ministry of Health New Zealand	New Zealand	New Zealand government	NZ[14]
Clinical Practice Guidelines: Diagnosis of Tuberculosis in Adults and Children	2017	American Thoracic Society, Infectious Diseases Society of America, Centers for Disease Control and Prevention (CDC)	United States (US)	US CDC	CDC-2017[26]
Clinical Practice Guidelines: Treatment of Drug-Susceptible Tuberculosis	2016	American Thoracic Society, Infectious Diseases Society of America, Centers for Disease Control and Prevention (CDC)	US	US CDC	CDC-2016[27]
Screening for Latent Tuberculosis Infection in Adults: US Preventive Services Task Force Recommendation Statement	2016	US Preventive Services Task Force	US	US Department of Health	USPSTF[28]
Tuberculosis Screening, Testing, and Treatment of U.S. Health Care Personnel: Recommendations from the National Tuberculosis Controllers Association and CDC, 2019	2019	National Tuberculosis Controllers Association, Centers for Disease Control and Prevention (CDC)	US	US CDC	CDC-2019[29]
Guidelines for the Treatment of Latent Tuberculosis Infection: Recommendations from the National Tuberculosis Controllers Association and CDC, 2020	2020	National Tuberculosis Controllers Association, CDC	US	US CDC	CDC-2020[12]
WHO Consolidated Guidelines on Tuberculosis. Modules 1-6	2020-2022	World Health Organization	Global	United States Agency for International Development and the Russian Federation	WHO[8]
Clinical Practice Guidelines: Management of Tuberculosis (Fourth Edition)	2021	Malaysian Health Technology Assessment Section, MOH Malaysia	Malaysia	MOH Malaysia	Malaysia[18]
Canadian Tuberculosis Standards	2022	Canadian Thoracic Society (CTS) of the Canadian Lung Association and Public Health Agency of Canada	Canada	CTS and Public Health Agency of Canada	Canada[17]

Table 1. List of international guidelines reviewed and used in the preparation of the updated TB guidelines. (Cont'd)

Guidelines	Year published	Developing organisations	Region	Funder	Abbreviation[Ref]
National Position Statement for the Management of Latent Tuberculosis Infection	2017	National Tuberculosis Advisory Committee	Australia	Australia	Australia-NTAC[15]
Tuberculosis: CDNA National Guidelines for Public Health Units	2022	Communicable Diseases Network Australia	Australia	Australia	Australia-CDNA[30]
Guidelines on targeted tuberculin testing and treatment of latent tuberculosis infection	2020	Department of Health of the Government of the Hong Kong SAR	Hong Kong SAR	Department of Health, Hong Kong SAR	Hong Kong[16]

[Ref] Superscript numbers: refer to REFERENCES

2-round modified Delphi process, consensus—meaning agreement or no objections from any Panel member—was reached on recommendations for each question. These recommendations are marked (*Adapted*), (*Adopted*) or (*De novo*) accordingly.

Draft guidelines were circulated to Singapore stakeholder bodies for review, with relevant changes incorporated into the final version. These stakeholders include:

- Academy of Medicine, Singapore:
  - Chapter of Family Medicine Physicians
  - College of Paediatrics & Child Health
  - College of Physicians:
    - Chapter of Infectious Disease Physicians
    - Chapter of Respiratory Physicians
- Agency for Care Effectiveness, Ministry of Health (methodology review only)

SCREENING FOR TB (QUESTIONS 1 AND 2)

1. What are the indications for screening for TB infection?

Recommendation

People living with human immunodeficiency virus (HIV), i.e. PLHIV, close contacts of persons with infectious TB disease and persons with clinical risk factors for TB should be screened for TB infection. (*Adapted*)

Remarks

Specific indications for screening for each at-risk group are given in Table 2. The definition of close contacts as well as the protocol for identifying and screening them had previously been established by the National TB Programme (Contact Tracing Manual, internal document). Recommendations for programmatic-level screening of other population groups (i.e. migrants from high TB burden countries, prisoners and healthcare workers, etc.) will not be further discussed here.

Multiple clinical and epidemiological factors increase the risk of TB disease progression and/or worse outcomes. However, the majority opinion of the Panel was that systematic TB screening for individuals with diabetes mellitus, tobacco/alcohol or other substance abuse, and the elderly in long-term care facilities might not be cost-effective in Singapore. Similarly, due to the heterogeneous nature of cancer, one single recommendation was deemed impractical.

Several guidelines highlighted prolonged moderate/high-dose corticosteroid use as a risk for TB reactivation.[14,15,17,25,28] However, the dose threshold and duration varied considerably. The Panel's majority consensus in the absence of clear evidence was a corticosteroid dose of ≥15 mg prednisolone equivalent for a duration of >8 weeks. Although other non-anti-tumour necrosis factor (TNF) biologics variably increase the risk of TB reactivation, these risks are lower and/or less well quantified compared with anti-TNF therapy.

2. What are the indications for screening for TB disease?

Recommendation

PLHIV, close contacts of persons with infectious TB disease and persons with clinical risk factors for TB should be screened for TB disease. (*Adapted*)

Remarks

Details regarding the specific indications for screening for each population group are given above in Section 1 and Table 2.

Table 2. Specific recommendations for screening for TB infection or disease.

Population	Screening for TB infection	Screening for TB disease	Comments	Guidelines referenced[a]
PLHIV	Universal screening upon diagnosis of HIV	1. Universal screening upon diagnosis of HIV 2. Workup from TB infection diagnosis[b]	HIV and TB co-infection rates are <3% of all HIV cases as of 2014[31]	WHO, NICE, CDC-2017, USPSTF, Canada, ECDC-L, Australia-NTAC, New Zealand, Malaysia, Hong Kong
Close contacts	Yes, to follow National TB Programme protocol	Yes, to follow National TB Programme protocol	Nil	WHO, NICE, CDC-2017, Canada, ECDC-L, Australia-NTAC, New Zealand, Malaysia, Hong Kong
Other clinical risk factors: • Anti-TNF therapy • Steroids[c] • Organ transplant • Dialysis • Silicosis	Universal screening before initiation of treatment or on diagnosis of condition (silicosis)	1. Universal screening before initiation of treatment or on diagnosis of condition (silicosis) 2. Workup from TB infection diagnosis[b]	Nil	WHO, NICE, CDC-2017, USPSTF, Canada, ECDC-L, Australia-NTAC, New Zealand, Hong Kong

HIV: human immunodeficiency virus; PLHIV: people living with HIV; TB: tuberculosis; TNF: tumour necrosis factor
[a] Refer to Abbreviation of names of guidelines in Table 1.
[b] If TB infection is diagnosed at any point after the initial TB screening event.
[c] Steroid dose ≥15 mg of prednisolone equivalent with duration >8 weeks.
Superscript number: refer to REFERENCES

DIAGNOSIS OF TB INFECTION (QUESTIONS 3 AND 4)

At present, there is no gold-standard test for the diagnosis of TB infection. Available tests in Singapore include the tuberculin skin tests (TST) and interferon-gamma release assays (IGRAs) that measure the adaptive immune response to *M. tuberculosis*. Both tests are less sensitive in immunocompromised persons and poorly predict likelihood of progression from TB infection to disease.[32,33]

3. What are the laboratory tests that should be performed to diagnose TB infection in adults?

Recommendations

- In adults, IGRAs are preferred to the TST for diagnosis of TB infection. (*Adopted*)
- The TST can be an alternative diagnostic test, with a cut-off induration of 10 mm in general and 5 mm in at-risk immunocompromised individuals. (*Adopted*)
- Parallel or sequential IGRA/TST testing is not recommended for immunocompromised individuals. (*De novo*)

Remarks

Despite higher test costs, the Panel recommended IGRAs over the TST in view of their ease of testing and convenience for individuals.

In significantly immunocompromised adults (i.e. PLHIV with CD4 counts <200 cells/mm^3; individuals on anti-TNF therapy and/or high-dose steroids; recipients of organ or stem cell transplantation, etc.), the T-SPOT.TB test (Oxford Immunotec, UK) may be more sensitive than the QuantiFERON-TB Gold Plus (Qiagen, US) QFT-Plus or the older QuantiFERON-TB Gold In-Tube (Qiagen, US), albeit with wide confidence intervals in meta-analyses.[32,33] In view of the weak evidence, the Panel chose not to recommend sequential or parallel IGRA/TST testing in immunocompromised populations despite such recommendations in several guidelines.[12-15,17,25,26]

4. What are the laboratory tests that should be performed to diagnose TB infection in children?

Recommendations

- In children aged ≥5 years or children >2 years with prior Bacillus Calmette-Guerin (BCG) vaccination, commercially available IGRAs are preferred to the TST for diagnosis of TB infection. (*Adapted*)
- In children aged 2–5 years without prior BCG vaccination, either TST or IGRA may be used for diagnosis of TB infection. (*Adopted*)
- In children aged 6 months to <2 years, the TST is recommended for diagnosis of TB infection. (*Adopted*)

- A TST cut-off induration of 10 mm is recommended (5 mm in at-risk children who are immunocompromised). (*Adopted*)

Remarks

Several guidelines recommended TST only or preferentially over IGRA in children <5 years old.[12-15,17,25,26] Phlebotomy is challenging in young children, and their lower functional immune response increases the likelihood of indeterminate IGRA results.[12-15,17,25,26,34] WHO guidelines supported both tests but stated that a positive result should not be a prerequisite for TB infection treatment for close contacts and PLHIVs.[8] Both tests are not recommended in infants <6 months old due to very low test sensitivity.

TREATMENT OF TB INFECTION (QUESTIONS 5 TO 11)

Prevention of TB disease by treatment of TB infection is a critical component of the WHO End TB Strategy.[8] The efficacy of current treatment regimens ranges from 60–90%, with a protective effect possibly lasting beyond a decade.[35] The decision to start TB infection treatment should be made jointly between the individual or his/her legal guardian, and the physician, taking into consideration the benefits and risks of treatment as well as the individual's preferences and values. In all special populations (i.e. PLHIV, pregnant women, children, etc.), treatment should be initiated by or in close consultation with the relevant clinical specialists specific to these populations.

5. What are the baseline tests to be performed before and after starting TB infection treatment?

Recommendations

- The possibility of TB disease must be excluded via a symptom screen and chest X-ray prior to initiating treatment for TB infection. (*Adopted*)
- Baseline aspartate aminotransferase (AST) and alanine aminotransferase (ALT) testing can be performed prior to starting TB infection treatment, with subsequent AST and ALT testing on follow-up visits in individuals at high risk for hepatotoxicity from treatment, or full liver function testing in those who have clinical features suggestive of liver dysfunction. (*Adapted*)

Remarks

It is important to rule out TB disease before initiating treatment for TB infection via assessment for symptoms suggestive of disease and a chest X-ray,[8,15-18,26] with sputum testing where indicated.[16-18,26]

Several guidelines recommended baseline liver function testing for individuals at risk of hepatotoxicity,[8,15,16,25] including those over the age of 35 years,[8,15] those with a history of liver disease, harmful use of alcohol/other drugs and HIV infection,[8,15,16,25] as well as pregnant and postpartum (<3 months) women.[8,16] Others recommended baseline testing for all individuals.[14,17,18] Risk-benefit assessments of treatment should be performed in individuals with abnormal liver function results.[8]

6. What are the preferred and alternative regimens for the treatment of TB infection in adults?

Recommendations

- The 2 preferred regimens for the treatment of TB infection in adults are rifampicin daily for 4 months (4R), or isoniazid daily for either 6 or 9 months (6H/9H). (*Adopted*)
- Alternative treatment regimens include isoniazid and rifampicin daily for 3 months (3HR), or isoniazid and rifapentine weekly for 3 months (3HP) should rifapentine become available in Singapore. (*Adapted*)
- For PLHIV, isoniazid and rifapentine daily for 1 month (1HP) may be another alternative regimen. (*Adopted*)
- Pyridoxine supplementation is recommended for all adults on an isoniazid-containing regimen. (*Adapted*)

Remarks

The preferred and alternative treatment regimens are listed in Table 3. There is no clear benefit of administering 6H/9H as a thrice-weekly regimen except when prescribed as part of DOT or video-observed treatment (VOT).[12,14,17]

WHO, ECDC-L and Malaysia guidelines (Table 1) favoured shorter regimens and those requiring less frequent administration by providers and patients, aiming to increase adherence and expedite completion.[8,13,18] This is particularly important for certain patient populations (i.e. the homeless, prisoners or those due to start anti-TNF therapy).

For PLHIV, 6H/9H is preferred to avoid potential drug interactions between rifamycin-containing regimens and antiviral therapy.[8,12,13,18,25] The data for 1HP comes from a single randomised clinical trial in PLHIV living in a high TB prevalence region,[37] and it is recommended as an alternative regimen for adult PLHIV in recent guidelines (2020 to 2022).[8,12,18]

Table 3. Recommended treatment regimens for TB infection.

Population	Preferred regimen (dose)	Alternative regimen	Comments
Adults	• 4R daily (10 mg/kg BW) • 6H/9H daily (5 mg/kg BW) with pyridoxine daily (10–25 mg)	• 3HP weekly (isoniazid 15 mg/kg BW up to 900 mg; rifapentine up to 900 mg) with pyridoxine daily • 3HR with pyridoxine daily	• 3HP weekly is not recommended in pregnant/breastfeeding women
PLHIV	• 6H/9H daily (5 mg/kg BW) with pyridoxine daily (10–25 mg) • 4R daily (10 mg/kg BW)	• 3HP weekly with pyridoxine daily • 3HR with pyridoxine daily • 1HP daily (isoniazid 300 mg/day; rifapentine 600 mg/day)	• To be aware of potential drug interactions with antiretroviral therapy
Children	• 4R daily (<10 years: 15 mg/kg BW; ≥10 years: 10 mg/kg BW) • 6H/9H daily (<10 years: 10 mg/kg BW; ≥10 years: 5 mg/kg BW)	• 3HR daily • 3HP weekly • (age ≥2 years only; weight-adjusted dosing[39])	• Isoniazid/rifapentine weekly dosing (2 –14 years): ○ 10 to <16 kg = 300 mg/300 mg ○ 16 to <24 kg = 500 mg/450 mg ○ 24 to <31 kg = 600 mg/600 mg ○ ≥31 kg = 700 mg/750 mg

1HP: isoniazid and rifapentine daily for 1 month; 3HP: isoniazid and rifapentine weekly for 3 months; 3HR: isoniazid and rifampicin daily for 3 months; 4R: rifampicin daily for 4 months; 6H/9H: isoniazid daily for either 6 or 9 months; BW: body weight; PLHIV: people living with HIV
Superscript number: refer to REFERENCES

Rifapentine is not recommended in pregnant women at present due to limited data on its pharmacokinetics and safety during pregnancy.[8,12,18] While Canadian guidelines recommended avoiding isoniazid-containing regimens in pregnancy due to potential risk of maternal hepatoxicity,[17] other guidelines had a different stance, citing insufficient evidence to warrant a separate recommendation.[8]

Pyridoxine supplementation can reduce the risk of isoniazid-induced peripheral neuropathy. It is variously recommended for all persons,[14,18] or only at-risk individuals[17,26,29] on an isoniazid-containing regimen. To simplify implementation, the Panel recommended pyridoxine supplementation for all adults on an isoniazid-containing regimen.

7. What are the preferred and alternative regimens for the treatment of TB infection in children?

Recommendations

- The 2 preferred regimens for the treatment of TB infection in children are 4R or 6H/9H. (*Adopted*)
- Alternative treatment regimens include isoniazid and rifampicin daily for 3HR or weekly 3HP in children age ≥2 years, should rifapentine become available in Singapore. (*Adapted*)
- Pyridoxine supplementation is recommended for children at risk of peripheral neuropathy while on an isoniazid-containing regimen. (*Adopted*)

Remarks

3HP is currently recommended to children ≥2 years old due to the limited data on the efficacy and pharmacology of rifapentine in younger children.

8. How should treatment be monitored?

Recommendation

- TB infection treatment should be self-administered except in the case of the weekly 3HP regimen or in a situation where adherence to treatment is poor. In the latter cases, DOT or VOT will be preferred. (*Adapted*)

Remarks

Most guidelines endorsed self-administered treatment for TB infection, reserving observed treatment only for the 3HP regimen,[12,17,18] or significant non-adherence.[12-14] Enforcing observed treatment for TB infection can pose a significant barrier for treatment.[8,36]

9. How should treatment interruptions be managed?

Recommendation

- For interruptions due to drug toxicity, the risk-benefit of continued treatment should be re-evaluated before resuming treatment. (*Adopted*)
- For interruptions due to hepatotoxicity, treatment drugs can be re-introduced once liver function of the affected individual has normalised. (*Adopted*)

Remarks

The treating physician should carefully weigh the risk and costs of potential additional hepatic injury against the benefit of TB preventive treatment with the affected individual before continuing or re-initiating treatment. Further details on treatment interruptions are in Supplementary Materials Appendix S2.

10. How often should patients on TB infection treatment be followed up and for how long?

Recommendations

- Follow-up for patients on TB infection treatment can be at 4–6 week intervals, with longer intervals for adherent patients who are at low risk for drug adverse events. (*Adapted*)
- Follow-ups can be discontinued upon treatment completion. (*Adopted*)

Remarks

Most guidelines recommended monthly follow-up for treatment monitoring.[8,12,14,16,17] Canadian guidelines supported extending the interval between visits for adherent patients who are at low risk of drug toxicity.

11. How should post-treatment re-exposure to infectious TB be managed?

Recommendation

- Re-treatment of individuals re-exposed to infectious TB can be considered. (*Adopted*)

Remarks

Currently tests cannot determine if re-infection has occurred in an individual who had previously completed treatment for TB infection or disease. Canada guidelines conditionally recommended against re-treatment of re-exposed individuals unless they are at high risk of progression to TB disease.[17] However, WHO stated that prior treatment should not be a contraindication to TB preventive treatment for close contacts of infectious individuals.[8] As a precautionary measure, the Panel elected to adopt WHO's recommendation.

DIAGNOSIS OF TB DISEASE (QUESTIONS 12 TO 17)

The emergence of new technologies has prompted a consideration for their incorporation into TB diagnosis algorithms and guidelines.

12. What tests should be performed to diagnose TB disease?

Recommendations

- The workup for TB disease should include clinical history taking and examination, chest X-ray, and collection of samples for acid-fast bacilli (AFB) smear microscopy, nucleic acid amplification test (NAAT) and mycobacterial culture. (*Adapted*)
- In children but not adults, the IGRA test or TST in the absence of prior BCG vaccination may be a supplementary test for diagnosing TB disease. (*Adapted*)

Remarks

WHO does not recommend AFB smear microscopy or mycobacterial culture in general for diagnosis of TB disease,[8] partially due to resource constraints in low-middle income and high-burden TB countries. The Panel debated the value of AFB smear microscopy given the widespread availability of NAAT in Singapore. The consensus was to remain aligned with other major guidelines in recommending AFB smear microscopy at present,[14,17,18,24-26,30] since these results still guide contact tracing operations.

Most guidelines support using either TST or IGRA as a supplementary diagnostic test in children, due to the challenges of obtaining clinical samples from young children as well as the paucibacillary nature of childhood pulmonary TB.[8,14,17,18,24-26,30] However, these tests are not recommended in adults due to their lower specificity and relative ease of obtaining clinical samples.

13. What are the types of clinical samples that should be collected and sent to the laboratory for testing?

Recommendations

- Sputum is the preferred clinical sample for the diagnosis of TB disease in adults and older children. (*Adopted*)
- In adults and children who are unable to properly expectorate, induced sputum is preferred, with early morning gastric aspirate considered an equivalent option in very young children and adults for whom sputum induction is unfeasible. (*Adopted*)
- Bronchoscopy with aspiration and lavage is an alternative option for sample collection in both adults and children. (*Adopted*)

- Nasopharyngeal aspiration is an alternative option for sample collection in children. (*Adopted*)

Remarks

Obtaining high-quality sputum increases the yield of microbiological testing. NZ and Malaysia guidelines (Table 1) recommended induced sputum over expectorated sputum for the diagnosis of pulmonary TB disease,[14,18] but the former is a secondary or equivalent option in other guidelines.[8,17,24-26,30]

Nasopharyngeal aspirate—while relatively easy to perform—has considerably lower sensitivity for diagnosing TB disease (NAAT sensitivity of approximately 46% versus 73% in gastric aspirate) in children.[38] It is recommended in 3 guidelines as an alternative when sputum or gastric aspirate collection is unfeasible.[8,14,18]

Testing stool via NAAT is newly recommended by WHO for the diagnosis of pulmonary TB in children, with reasonable diagnostic sensitivity.[8] However, in view of the familiarity of obtaining induced sputum or gastric fluid samples at present coupled with additional laboratory processing steps for stool,[17,38] the Panel elected not to recommend stool as an alternative sample type for TB diagnosis.

14. Excluding bronchoscopic sampling, what should be the minimum number of samples sent and when should the samples be collected?

Recommendations

- At least 2 samples should be sent for laboratory testing for TB disease, of which at least 1 sample should be an early morning sample. (*Adapted*)
- Same-day collection of samples is preferred except in the case of gastric aspirate, for which alternate day (i.e. day 1 and day 2) collection of early morning samples is ideal. (*Adopted*)

Remarks

The recommendations from international guidelines on this question are listed in Supplementary Materials Appendix S4. Most guidelines recommended a minimum of 3 samples,[17,24-26,30] with most also preferring early morning collection.[14,18,24-26] Considering the marginal additional diagnostic yield for a third sample (2–3% on average),[24] the Panel elected to set the minimum at 2 samples, with at least 1 collected in the early morning. Fasted early morning samples are recommended for gastric aspirate collection as gastric emptying following food or water intake will significantly reduce diagnostic yield.[38]

15. Under what circumstances should computed tomography (CT) scan of the thorax be considered for the diagnosis of pulmonary TB?

Recommendation

- CT thorax may be considered as an additional imaging modality for the purposes of excluding another underlying diagnosis or for the evaluation of extrapulmonary TB disease with pulmonary involvement (i.e. pleural or pericardial disease). (*Adopted*)

Remarks

CT thorax was positioned across several guidelines as a non-routine and additional imaging modality to be considered for use for further diagnostic investigations.[8,14,17,25]

16. What is the role of next-generation sequencing (NGS) as a clinical laboratory tool?

Recommendations

- Whole genome sequencing (WGS) can be deployed in clinical laboratories in Singapore for genotypic prediction of TB drug resistance. It has demonstrable high sensitivity and specificity especially for first-line drugs, and the output can also be used for national TB surveillance. (*Adapted*)

Remarks

NGS applications for TB currently include prediction of drug resistance via both WGS[6-8] and targeted NGS (tNGS),[8] and molecular epidemiological investigations of transmission via WGS. WGS is ideally performed on cultured *M. tuberculosis* isolates due to the need for high-quality deoxyribonucleic acid (DNA), with techniques to enable direct sequencing from clinical samples still very much a work-in-progress. In Singapore, WGS is primarily used for public health surveillance. Routine WGS for drug susceptibility testing in clinical laboratories could eliminate the requirement for phenotypic testing in genotypic pan-sensitive *M. tuberculosis* isolates.

tNGS assays determine drug resistance via targeted panels that interrogate specific relevant regions of the genome, and can be used on DNA directly obtained from a clinical sample with a more rapid turnaround time. Commercially available tNGS assays have been developed, but none are available in Singapore at the time of writing. While tNGS shows good concordance with WGS in predicting drug resistance, it is currently limited to known resistance genotypic markers. WHO is currently developing a guideline on the use of

tNGS for detecting drug resistance,[39] and the Panel has no specific recommendation regarding tNGS at present.

17. What is the role of machine learning in the screening, diagnosis and follow-up of pulmonary TB?

Recommendations

- The panel recognises that a number of automated CAD-AI products have been certified by the Health Sciences Authority, Singapore and more are expected in the future. Some products have been deployed in other countries for chest X-ray screening for TB disease in various settings. While multiple studies have demonstrated comparable diagnostic performance (relative to a human radiologist) for the detection of chest X-ray changes associated with pulmonary TB, such findings will have to be validated within the local practice setting in Singapore. At present, deployment of CAD-AI products will require a human-in-the-loop, where a certified radiologist is still needed to provide a final read. In addition, due consideration must be given to the fact that current products have narrow capabilities, with potential to miss other non-specified abnormalities on a chest X-ray. (*De novo*)
- The panel also recommended vigilance in this rapidly evolving field and timely updates of major developments. (*De novo*)

Remarks

WHO guidelines recommended that CAD-AI products may be used in place of human readers for interpreting digital chest X-rays in screening and triaging individuals ≥15 years old for TB disease where TB screening is recommended.[8] Canada guidelines note that CAD-AI products may be valuable in closing diagnostic gaps in resource-limited and remote settings.[17] An online and regularly updated data repository of commercially available CAD-AI products for TB screening is maintained jointly by the Stop TB Partnership and FIND (https://www.ai4hlth.org/). Further details on CAD-AI products for TB disease screening are available in the Supplementary Materials Appendix S5.

TREATMENT OF TB DISEASE (QUESTIONS 18 TO 25)

Early diagnosis and completion of TB disease treatment are cornerstones of global and local strategy for TB elimination.[8] However, adherence to treatment can be challenging because of its lengthy duration, pill burden and side effects.[8,14,17,18,24,25,27,30] Incomplete adherence increases the risk for drug resistance and relapse. In Singapore, treatment completion, relapse and case fatality rates are 65%, 6.1% and 0.4%, respectively, in 2020 (data by the National Tuberculosis Programme).

Although TB disease should be tackled at a programmatic level, a patient-centred approach focusing on increasing treatment literacy and enabling adherence is necessary. Comprehensive treatment support and tailored adherence monitoring interventions should be developed in collaboration with persons with TB disease.[8] A multidisciplinary team involving physicians, nurse practitioners, public health personnel, social workers and interpreters where necessary is the standard of care in Singapore. In all special populations (i.e. PLHIV, pregnant women, children, etc.), treatment of TB disease should be initiated by or in close consultation with the relevant clinical specialists. Table 4 outlines the aspects and components of a TB disease treatment plan.

18. What are the preferred and alternative regimens for the treatment of drug-susceptible pulmonary TB disease in adults?

Recommendations

- The preferred regimen for the treatment of drug-susceptible pulmonary TB disease in adults is 2-month intensive phase of isoniazid (H), rifampicin (R), pyrazinamide (Z) and ethambutol (E), followed by 4-month continuation phase of H and R (2HRZE/4HR). (*Adopted*)
- In persons aged >70 years with low risk of drug resistance, pyrazinamide should be avoided; the preferred regimen is a 2-month intensive phase of H, R and E followed by 7-month continuation phase of H and R (2HRE/7HR). (*Adapted*)
- An alternative regimen, should rifapentine become available in Singapore, is an 8-week intensive phase of daily H, rifapentine (P), moxifloxacin (M) and Z, followed by 9-week continuation phase of daily H, P and M (2HPMZ/2HPM)—except in pregnant/breastfeeding women and PLHIV. (*Adopted*)
- For PLHIV with TB disease, care should be taken to avoid heightened risk of drug adverse effects due to interactions with antiretroviral therapy. (*Adopted*)
- Pyridoxine supplementation is recommended for all adults on an isoniazid-containing regimen. (*Adopted*)

Table 4. Aspects and components of a TB disease treatment plan.

Aspects	Components	Guidelines referenced[a]
Principles	Patient-centered care with comprehensive and individualised treatment support	11,17,20,27,30
	Balancing patient rights/choices with public health and safety	11,20,27,28,30
Enablers, incentives and adherence interventions	Patient health education and counselling; improving treatment literacy	11,20,21,27,28,30
	Psychological, material (including food, financial and transport subsidies) and social support for patients	11,20,21,27,28,30
	Healthcare worker education and training	11,28
	Integration with patients' primary/specialty care where possible	11,20,27,28,30
	Observed treatment (DOT or VOT or family supervision for children)	11,17,20,21,27,28,30

DOT: directly observed treatment; VOT: video-observed treatment
[a] Numbers: refer to REFERENCES

Table 5. Recommended treatment regimens for TB disease.

Population	Regimen and dose	Comments
Adults, general	• 2HRZE/4HR with pyridoxine ○ Isoniazid 5 mg/kg BW daily OR 10 mg/kg BW thrice-weekly ○ Rifampicin 10 mg/kg BW daily or thrice-weekly ○ Pyrazinamide 20–25 mg/kg BW daily or 30–40 mg/kg BW thrice-weekly ○ Ethambutol 15–20 mg/kg BW daily or 25–40 mg/kg BW thrice-weekly ○ Pyridoxine 10–25 mg daily or thrice-weekly	• Daily dosing is preferred, particularly during the intensive phase[11] • Thrice-weekly dosing is more convenient for observed treatment (DOT or VOT)
	• 2HPMZ/2HPM with pyridoxine ○ Isoniazid, pyrazinamide and pyridoxine doses as above ○ Rifapentine 1200 mg daily ○ Moxifloxacin 400 mg daily	• Daily dosing only
Children	• 2HRZ(E)/4HR OR 2HRZ(E)/2HR ○ Isoniazid 10–15 mg/kg BW daily (maximum 300 mg) or 20 mg/kg BW thrice-weekly (maximum 900 mg) ○ Rifampicin 15–20 mg/kg BW daily (maximum 600 mg) or 20 mg/kg BW thrice-weekly (maximum 600 mg) ○ Pyrazinamide 30–40 mg/kg daily (maximum 2 g) ○ Ethambutol 15–25 mg/kg daily (maximum 1 g)	• Daily dosing is preferred[11]

BW: body weight; DOT: directly observed treatment; VOT: video-observed treatment;
E: ethambutol; H: isoniazid; M: moxifloxacin; P: rifapentine; R: rifampicin; Z: pyrazinamide;
2HPMZ/2HPM: 8-week intensive phase of daily H, P, M and Z, followed by 9-week continuation phase of daily H, P and M;
2HRZ(E)/2HR: 2-month intensive phase of H, R and Z with or without E followed by 2-month continuation phase of H and R;
2HRZE/4HR: 2-month intensive phase of H, R, Z and E, followed by 4-month continuation phase of H and R;
2HRZ(E)/4HR: 2-month intensive phase of H, R and Z with or without E followed by 4-month continuation phase of H and R
Superscript numbers: refer to REFERENCES

Remarks

Treatment may be initiated based on strong clinical and radiological suspicion of TB disease, although clinical samples for microbiological diagnosis should be collected first if possible. Table 5 lists the preferred and alternative treatment regimens for TB disease where drug resistance is not expected. The first-line regimen of choice in all guidelines remains 2HRZE/4HR,[8,14,17,18,24,25,27,30] which has a global treatment success rate exceeding 85% and low risk of adverse events.[8]

Adverse events resulting in treatment interruption or discontinuation are more common among the elderly, particularly for pyrazinamide, which accounted for >15% of such events.[40] A number of age thresholds (>65 to >80 years) were proposed in guidelines where use of pyrazinamide was cautioned,[8,17,27] with the Panel setting it at >70 years based on existing practice in Singapore. Treatment duration should be extended to 9 months (2HRE/7HR) if pyrazinamide is excluded.

Following the completion of the Tuberculosis Trials Consortium (TBTC) Study 31 trial, where a shortened 2HPMZ/2HPM regimen was shown to be non-inferior to 2HRZE/4HR,[9] WHO and CDC included 2HPMZ/2HPM as an alternative regimen in updated guidelines.[8,10] The Panel therefore also recommended this as an alternative treatment regimen, pending rifapentine availability in Singapore. The higher cost of the drug regimen will likely be offset by the shorter duration of treatment for both the individual and the national programme. However, insufficient data precluded the Panel from recommending this regimen for pregnant/breastfeeding women, PLHIV and children <12 years of age.

Detailed recommendations for PLHIV with TB disease can be found in the Supplementary Materials Appendix S6.

19. What are the preferred and alternative regimens for the treatment of drug-susceptible pulmonary TB disease in children?

Recommendations

- The preferred regimen for the treatment of drug-susceptible pulmonary TB disease in children is 2HRZ(E)/4HR. (*Adopted*)
- In children with non-severe disease, a shortened 2HRZ(E)/2HR regimen may be considered after consultation with an experienced paediatric specialist. (*Adapted*)
- An alternative regimen of 2HPMZ/2HPM can be considered in children >12 years of age should rifapentine become available in Singapore. (*Adapted*)
- Pyridoxine supplementation is recommended for children at risk of peripheral neuropathy while on an isoniazid-containing regimen. (*Adopted*)

Remarks

Where the population prevalence of drug-resistant TB is low, ethambutol may be omitted from the intensive phase of treatment for HIV-negative children.[8]

The shorter treatment for minimal TB in children (SHINE) trial demonstrated non-inferiority of a 4-month regimen 2HRZ(E)/2HR to the standard 6-month regimen in children <16 years of age with non-severe (defined as respiratory TB confined to 1 lobe with no cavities, no signs of miliary TB, no complex pleural effusion, and no clinically significant airway obstruction or peripheral lymph-node tuberculosis) smear-negative TB.[41] The Panel supported consideration of this shortened regimen for non-severe TB in children in Singapore under the guidance of an experienced paediatric TB specialist, in line with WHO guidelines.[8]

20. What are the other considerations for prolonging or shortening treatment for TB disease?

Recommendations

- In persons with TB disease who present with baseline cavitation and/or extensive disease on chest X-ray AND positive mycobacterial cultures after 2 months of treatment, the Panel strongly recommended that the treatment duration be extended to 9 months (2HRZE/7HR). (*Adopted*)
- Other considerations for extending treatment to 9 months (2HRZE/7HR) in persons with microbiologically-proven TB disease include the following: (*Adopted*)
    - Positive mycobacterial cultures after 2 months of treatment
    - Baseline cavitation and/or extensive disease on chest X-ray
    - Slow radiological improvement
    - PLHIV with CD4 count <200 cells/mm^3 and not on anti-retroviral therapy
    - Poorly controlled diabetes mellitus throughout treatment course
    - In persons with culture-negative but probable TB disease, treatment duration may be shortened to 4 months (2HRZE/2HR) (*Adopted*)

Remarks

CDC-2016 and Canada guidelines recommended extending treatment to 9 months in individuals with baseline cavitation and/or extensive disease on chest X-ray who remain culture-positive after 2 months of treatment.[17,27] These patients are at the highest risk of relapse post-treatment (>10%) with the 2HRZE/4HR regimen.[17,27,42] Other factors associated with a higher risk of relapse are as listed in the immediate Recommendations above.[17,27,42] Although the Panel recommended considering treatment extension in the presence of these factors, current evidence does not suggest this results in lower relapse rates, and WHO guidelines do not recommend extending TB disease treatment beyond 6 months in such cases.[8]

In persons with culture-negative TB, considerable evidence including from Singapore suggests that a shortened treatment duration is effective.[27,43]

21. How should the treatment of TB disease be monitored?

Recommendations

- In adults, observation of treatment is recommended for those with infectious TB disease, in particular those who are at risk of adverse outcomes and/or those who are non-adherent to treatment. (*Adapted*)
- All children with TB disease should ideally undergo some form of observed treatment, including home supervision by an adult family member or caregiver. (*Adapted*)
- VOT is equivalent to DOT in terms of ensuring treatment adherence in those who are able to undergo VOT. (*Adapted*)

Remarks

DOT is a strategy to enhance TB disease treatment adherence and completion, and a central tenet in TB control in Singapore.[3] While all international TB guidelines support DOT under varied situations, several guidelines also emphasise the importance of balancing individual rights with public health objectives.[14,25,27,30] In Singapore, nurse-led clinic and home DOT are practised alongside family supervision of TB treatment in young children. Despite concerns raised in guidelines including by WHO,[8,27] Singapore's experience with family supervision in young children has demonstrated high treatment success rates.

In recent years, VOT has emerged as a cost-effective and patient autonomy-enabling alternative to DOT,[19,20] endorsed by most guidelines.[8,14,17,18,25] However, it should be properly organised,[8] with patient education[18] and in-person support within a proper monitoring and evaluation framework.[17] In general, the least restrictive and effective interventions should be applied, with patients meaningfully involved in treatment-related decisions.[30]

A detailed discussion of DOT and VOT along with other enablers for treatment adherence can be found in the Supplementary Materials Appendix S7.

Table 6. Management of treatment interruptions (adapted from WHO, NZ and CDC-2016 guidelines).[8,14,27]

Timepoint of interruption	Details of interruption	Recommendation
Intensive phase	Lapse was <14 days in duration	Continue treatment to complete planned total number of doses[a]
	Lapse was ≥14 days in duration	Restart treatment
Continuation phase (6-month 2HRZE/4HR regimen)	Received ≥80% of doses and was sputum AFB microscopy negative initially	To determine if further treatment is necessary
	Received ≥80% of doses within 16 weeks and was sputum AFB microscopy positive initially	Continue treatment until all doses are completed
	Received <80% of doses and accumulated interruptions were <3 months in duration	Continue treatment until all doses are completed if all continuation phase doses can be completed within 6 months; but to restart treatment again from the intensive phase if treatment cannot be completed within this timeframe
	Received <80% of doses and accumulated interruptions were ≥3 months in duration	To restart treatment all over again from the intensive phase
Continuation phase (4-month 2HRZE/2HR regimen)	Completed ≥80% of doses within 8 weeks	To determine if further treatment is necessary
	Completed <80% of doses and accumulated interruptions were <1 month	Continue treatment until all doses are complete
	Completed <80% of doses and accumulated interruptions were ≥1 month	Restart treatment all over again from intensive phase

AFB: acid-fast bacilli; E: ethambutol; H: isoniazid; R: rifampicin; Z: pyrazinamide
2HRZE/2HR: 2-month intensive phase of H, R, Z and E followed by 2-month continuation phase of H and R;
2HRZE/4HR: 2-month intensive phase of H, R, Z and E, followed by 4-month continuation phase of H and R
[a] If all intensive phase doses can be completed within 3 months.
Superscript numbers: refer to REFERENCES

22. How should treatment interruptions be managed?

Recommendations

- The decision to continue or restart TB treatment is based on the duration of the interruption, whether it occurred during the intensive (i.e. 2HRZE) or continuation (i.e. 4HR) phase and the bacteriological status prior to interruption (Table 6). (*Adapted*)
- For interruptions due to hepatotoxicity, it is important to identify the culprit drug(s), with sequential re-introduction once liver enzymes return to <2 times the upper limit of normal. (*Adopted*)
- For interruptions due to other drug adverse events, continuation of TB treatment with symptom alleviation should generally be the norm for mild adverse events, but serious adverse events should result in the discontinuation of the offending drug(s). (*Adopted*)
- In severe or highly infectious persons with TB disease, initiation of an alternative treatment regimen is recommended while waiting for a serious drug adverse event to resolve. (*Adopted*)

Remarks

Repeated and/or prolonged TB treatment interruptions lead to worse outcomes.[44] WHO, CDC-2016 and NZ guidelines offer similar algorithms for clinical decisions regarding resuming or restarting TB treatment after interruptions (Table 6). However, most recommendations are based on expert opinion and experience as there is no strong body of evidence available.[8,27] A more comprehensive discussion of interruptions and how these might be addressed is provided in the Supplementary Materials Appendix S3.

23. How often should patients on TB disease treatment be followed up and for how long?

Recommendations

- Follow-up for patients on TB disease treatment can be at 2–4 week intervals during the intensive phase, and at longer 4–6 week intervals during the continuation phase. (*Adapted*)
- More frequent follow-ups are recommended for patients at high risk of adverse events or who have developed adverse events to treatment. (*Adapted*)
- Follow-ups can be discontinued upon treatment completion, except in patients at higher risk of poor outcomes, for whom follow-up until 1–2 years post-treatment can be considered. (*Adopted*)

Remarks

The recommended follow-up schedule aligns with both local practice and international guidelines.[8,14,17,18,24,27] Post-treatment, Canada guidelines conditionally recommended that patients with a high risk for TB recurrence (extensive/disseminated disease, cavitation on chest X-ray with smear/culture positive disease, immunosuppression, a history of treatment interruptions or non-adherence, and/or an atypical treatment regimen) be followed up for 1–2 years.[17] WHO guidelines recommended post-treatment monitoring for potential relapse for shorter (<6 months) treatment regimens, including children and adolescents on 2HRZE/2HR.[8]

24. What investigations should be performed prior to initiating and during treatment of TB disease?

Recommendations

- Before initiating TB treatment, if not already done elsewhere, work-up for TB disease should be performed as recommended above (Questions 13 to 18).
- Other baseline laboratory investigations in adults include—if not already done elsewhere—blood testing for HIV, AST/ALT, full blood count, diabetes screening, and renal function. (*Adapted*)
- In children, similar baseline laboratory investigations can be considered except for diabetes screening, which is unnecessary. (*Adapted*)
- Chest X-rays should be performed after completion of the intensive phase of treatment, and upon completion of TB treatment. (*Adopted*)
- Clinical samples should be collected for AFB smear microscopy and mycobacterial culture after completion of the intensive phase of treatment, and upon completion of TB treatment. (*Adopted*)
- In adults, visual assessments should be performed at initial and subsequent clinic visits if on an ethambutol-containing regimen. (*Adopted*)
- The patient should be weighed at all clinic visits. (*Adapted*)

Remarks

Clinical assessments coupled with weight measurements, chest X-ray and clinical sample (generally sputum) testing are important for evaluating treatment response and the risk of relapse/need for prolongation of treatment. The

final clinical sample testing is recommended to confirm bacteriological cure.[45]

Blood tests other than for liver function assessment need not be repeated if normal and if no drug adverse events manifest. AST and ALT should be repeated monthly or at each clinic visit in persons with a higher risk of hepatotoxicity.

Ocular toxicity from ethambutol is extremely rare in children, hence baseline and/or follow-up visual assessments are at the discretion of the treating specialist.

25. Should there be post-treatment evaluation for post-TB lung disease?

Recommendations

- Clinical assessment for post-TB lung disease, including a chest X-ray, can be performed at the end of treatment. (*Adapted*)

Remarks

Post-TB lung disease refers to an overlapping spectrum of diverse chronic respiratory conditions experienced after TB disease treatment.[8,17,46] They are under-recognised but contribute significantly to excess morbidity and mortality after post-TB treatment.[46] The Canada guidelines are the only ones that formally suggest assessing for post-TB lung disease. The guidelines advise that, within 6 months of finishing treatment, everyone should undergo lung function tests.[17] WHO guidelines recommended a chest X-ray at the end of treatment to manage post-treatment TB pulmonary sequelae.[8] Consensus-based standards for assessing, managing and rehabilitating post-TB lung disease have been developed by international experts.[47]

IMPLEMENTATION CONSIDERATIONS

TB control in Singapore is overseen by the National TB Programme at the National Centre for Infectious Diseases, while the majority of persons with pulmonary TB disease are managed at the TB Control Unit, Tan Tock Seng Hospital. These guidelines will largely be implemented by the National TB Programme team, including their dissemination and education of TB care providers, with sourcing for rifapentine through the Global Drug Facility.

SCHEDULED REVIEW AND UPDATE

The National TB Programme will convene a multidisciplinary team in 2026 to update these guidelines via a systematic review of available guidelines and evidence.

Acknowledgements

This work was funded by the Ministry of Health Singapore. Special thanks to the Evidence to Practice Office of the Agency for Care Effectiveness (ACE), for the insightful and thorough review on the methodology of these guidelines. We are also grateful to our colleagues at the Academy of Medicine, Singapore who spent time reviewing and providing comments on an earlier draft. Further appreciation is given to Dr Shera Tan, former Clinical Director of the TB Control Unit, Tan Tock Seng Hospital, for her comments on selected areas of these guidelines. We also thank Dr Dhiya Ramizah of NCID for her support. Last but not least, we thank Ms Juliet Seah, Executive Assistant at the National TB Programme, for organising the meetings for the Panel.

Author OCWM has received speaking fees from Qiagen, conference sponsorship from Merck Sharpe & Dohme, and a grant from Institut Mérieux. Other Panel members have no potential conflict of interest to declare.

Supplementary Materials

Appendix S1. Methodology
Appendix S2. Managing treatment interruptions in TB infection
Appendix S3. Managing treatment interruptions in TB disease
Appendix S4. Recommendations for clinical sample collection
Appendix S5. CAD-AI products for TB disease screening
Appendix S6. Recommendations for PLHIV with TB disease
Appendix S7. Enablers for treatment adherence
Appendix S8. References for Supplementary Materials
Appendix S9. RIGHT-Ad@pt checklist
Appendix S10. Glossary
Appendix S11. Members of the Clinical TB Guidelines Development Team

REFERENCES

1. World Health Organization. Global Tuberculosis Report 2022. https://www.who.int/publications/i/item/9789240061729. Accessed 9 August 2023.
2. Yap P, Tan KHX, Lim WY, et al. Prevalence of and risk factors associated with latent tuberculosis in Singapore: a cross-sectional survey. Int J Infect Dis 2018;72:55-62.
3. Ministry of Health Singapore. Prevention, Diagnosis and Management of Tuberculosis. MOH Clinical Practice Guidelines 1/2016. https://www.moh.gov.sg/docs/librariesprovider4/guidelines/moh-tb-cpg-full-version-for-website.pdf. Accessed 18 February 2024.
4. World Health Organization on behalf of the Special Programme for Research and Training in Tropical Diseases (TDR).

Determining the local calibration of computer-assisted detection (CAD) thresholds and other parameters: A toolkit to support the effective use of CAD for TB screening. 2021. https://apps.who.int/iris/bitstream/handle/10665/345925/9789240028616-eng.pdf. Accessed 12 August 2023.

5. The Stop TB Partnership. AI-powered computer-aided detection (CAD) software. https://www.stoptb.org/digital-health-technology-hub/ai-powered-computer-aided-detection-cad-software. Accessed 12 August 2023.
6. CRyPTIC Consortium and the 100,000 Genomes Project, Allix-Béguec C, Arandjelovic I, et al. Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. N Engl J Med 2018;379:1403-15.
7. World Health Organization. Catalogue of mutations in *Mycobacterium Tuberculosis* complex and their Association with drug resistance: supplementary document. 2021. https://apps.who.int/iris/handle/10665/341906. Accessed 11 August 2023.
8. World Health Organization. WHO Consolidated Guidelines on Tuberculosis. Modules 1-6. 2020-2022. https://tbksp.org/en/guidance-books-solr. Accessed 12 August 2023.
9. Dorman SE, Nahid P, Kurbatova EV, et al. Four-month rifapentine regimens with or without moxifloxacin for tuberculosis. N Engl J Med 2021;384:1705-18.
10. Carr W, Kurbatova E, Starks A, et al. Interim guidance: 4-month rifapentine-moxifloxacin regimen for the treatment of drug-susceptible pulmonary tuberculosis – United States, 2022. MMWR Morb Mortal Wkly Rep 2022;71:285-9.
11. Pettit AC, Phillips PPJ, Kurbatova E, et al. Rifapentine with and without moxifloxacin for pulmonary tuberculosis in people with human immunodeficiency virus (S31/A5349). Clin Infect Dis 2023;76:e580-9.
12. Sterling TR, Njie G, Zenner D, et al. Guidelines for the treatment of latent tuberculosis infection: Recommendations from the National Tuberculosis Controllers Association and CDC, 2020. MMWR Recomm Rep 2020;69:1-11.
13. European Centre for Disease Prevention and Control. Scientific Advice: Programmatic management of latent tuberculosis infection in the European Union. European Centre for Disease Prevention and Control. 2018. https://www.ecdc.europa.eu/en/publications-data/programmatic-management-latent-tuberculosis-infection-european-union. Accessed 10 August 2023.
14. Ministry of Health, New Zealand. Guidelines for tuberculosis control in New Zealand, 2019. 2019. https://www.health.govt.nz/system/files/documents/publications/guidelines-tuberculosis-control-new-zealand-2019-august2019-final.pdf. Accessed 10 August 2023.
15. Stock D, National Tuberculosis Advisory Committee. National position statement for the management of latent tuberculosis infection. Commun Dis Intell Q Rep 2017;41:E204-8.
16. Tuberculosis & Chest Service, Department of Health of the Government of Hong Kong SAR. Guidelines on targeted tuberculin testing and treatment of latent tuberculosis infection. 2020. https://www.info.gov.hk/tb_chest/doc/2_LTBI_guide_TBCS_update%2019%20Dec%202020_Clean.pdf. Accessed 10 August 2023.
17. Canadian Thoracic Society. Canadian Tuberculosis Standards 8th edition. Canadian J Respiratory, Crit Care, and Sleep Medicine. 2022;6 Suppl 1. https://www.tandfonline.com/toc/ucts20/6/sup1. Accessed 12 August 2023.
18. Malaysian Health Technology Assessment Section (MaHTAS). Medical Development Division, Ministry of Health Malaysia. Clinical Practice Guidelines: Management of Tuberculosis (Fourth edition). 2021. https://www.moh.gov.my/moh/resources/Penerbitan/CPG/Respiratory/CPG-_Management_of_Tuberculosis_(4th_Edition).pdf. Accessed 10 August 2023.
19. Story A, Aldridge RW, Smith CM, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. Lancet 2019;393:1216-24.
20. Truong CB, Tanni KA, Qian J. Video-observed therapy versus directly observed therapy in patients with tuberculosis. Am J Prev Med 2022;62:450-8.
21. ADAPTE. The ADAPTE process: resource toolkit for guideline adaptation. Version 2.0. 2010. https://g-i-n.net/wp-content/uploads/2021/03/ADAPTE-Resource-toolkit-March-2010.pdf. Accessed 11 August 2023.
22. Song Y, Alonso-Coello P, Ballesteros M, et al. A reporting tool for adapted guidelines in health care: the RIGHT-Ad@pt checklist. Ann Intern Med 2022;175:710-9.
23. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 2010;182:E839-42.
24. European Centre for Disease Prevention and Control and European Respiratory Society. European Union Standards for Tuberculosis Care - 2017 update. https://www.ecdc.europa.eu/en/publications-data/european-union-standards-tuberculosis-care-2017-update. Accessed 18 February 2024.
25. National Institute for Health and Care Excellence. Tuberculosis: NICE guideline (NG33). 2019. https://www.nice.org.uk/guidance/ng33/resources/tuberculosis-pdf-1837390683589. Accessed 12 August 2023.
26. Lewinsohn DM, Leonard MK, LoBue, PA, et al. Official American Thoracic Society/Infectious Diseases Society of America/Centers for Disease Control and Prevention Clinical Practice Guidelines: diagnosis of tuberculosis in adults and children. Clin Infect Dis 2017;64:111-5.
27. Nahid P, Dorman SE, Alipanah N, et al. Official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America Clinical Practice Guidelines: Treatment of Drug-Susceptible Tuberculosis. Clin Infect Dis 2016;63:e147-95.
28. US Preventive Services Task Force. Screening for latent tuberculosis infection in adults: US Preventive Services Task Force recommendation statement. JAMA 2016;316:962-9.
29. Sosa LE, Njie GJ, Lobato MN, et al. Tuberculosis screening, testing, and treatment of U.S. health care personnel: recommendations from the National Tuberculosis Controllers Association and CDC, 2019. MMWR Morb Mortal Wkly Rep 2019;68:439-43.
30. Australian Government Department of Health, Communicable Diseases Network Australia. Tuberculosis: CDNA national guidelines for public health units. Version 3.0, 3 March 2022. https://www.health.gov.au/sites/default/files/documents/2022/06/tuberculosis-cdna-national-guidelines-for-public-health-units.pdf. Accessed 12 August 2023.
31. Nandar K, Ang LW, Tey J, et al. Epidemiology of tuberculosis and HIV coinfections in Singapore, 2000–2014. HIV Med 2018;19:59-64.
32. World Health Organization. Use of alternative interferon-gamma release assays for the diagnosis of TB infection: WHO policy statement. 2022. https://www.who.int/publications-detail-redirect/9789240042346. Accessed 13 August 2023.
33. Oh CE, Ortiz-Brizuela E, Bastos ML, et al. Comparing the Diagnostic Performance of QuantiFERON-TB Gold Plus to Other Tests of Latent Tuberculosis Infection: A Systematic Review and Meta-analysis. Clin Infect Dis 2021;73:e1116-25.
34. Zhou G, Luo Q, Luo S, et al. Indeterminate results of interferon gamma release assays in the screening of latent tuberculosis

infection: a systematic review and meta-analysis. Front Immunol 2023;14:1170579.

35. Getahun H, Matteelli A, Abubakar I, et al. Management of latent *Mycobacterium tuberculosis* infection: WHO guidelines for low tuberculosis burden countries. Eur Respir J 2015; 46:1563-76.
36. World Health Organization. WHO operational handbook on tuberculosis: module 1: prevention: tuberculosis preventive treatment. Chapter 5.1: recommended dosages of TPT medication. 2020. https://tbksp.org/en/node/1271. Accessed 13 August 2023.
37. Swindells S, Ramchandani R, Gupta A, et al. One Month of Rifapentine plus Isoniazid to Prevent HIV-Related Tuberculosis. N Engl J Med 2019;380:1001-11.
38. World Health Organization. WHO operational handbook on tuberculosis: module 5: management of tuberculosis in children and adolescents. 2022. https://tbksp.org/en/node/1730. Accessed 13 August 2023.
39. World Health Organization. Public notice: guideline development group meeting on targeted next-generation sequencing for detection of TB drug resistance. 2023. https://www.who.int/publications/m/item/public-notice--guideline-development-group-meeting-on-targeted-next-generation-sequencing-for-detection-of-tb-drug-resistance. Accessed 14 August 2023.
40. Kwon BS, Kim Y, Lee SH, et al. The high incidence of severe adverse events due to pyrazinamide in elderly patients with tuberculosis. PLoS One 2020;15:e0236109.
41. Turkova A, Wills GH, Wobudeya E, et al. Shorter treatment for nonsevere tuberculosis in African and Indian children. N Engl J Med 2022;386:911-22.
42. Jo KW, Yoo JW, Hong Y, et al. Risk factors for 1-year relapse of pulmonary tuberculosis treated with a 6-month daily regimen. Respir Med 2014;108:654-9.
43. Teo SK, Tan KK, Khoo TK. Four-month chemotherapy in the treatment of smear-negative pulmonary tuberculosis: results at 30 to 60 months. Ann Acad Med Singap 2002;31:175-81.
44. Burman WJ, Cohn DL, Rietmeijer CA, et al. Noncompliance with directly observed therapy for tuberculosis: epidemiology and effect on the outcome of treatment. Chest 1997; 111:1168-73.
45. Stadler JAM. Updated WHO definitions for tuberculosis outcomes: simplified, unified and future-proofed. Afr J Thorac Crit Care Med 2022;28:10.7196/AJTCCM.2022.v28i2.224.
46. Allwood BW, Byrne A, Meghji J, et al. Post-tuberculosis lung disease: clinical review of an under-recognised global challenge. Respiration 2021;100:751-63.
47. Migliori GB, Marx FM, Ambrosino N, et al. Clinical standards for the assessment, management and rehabilitation of post-TB lung disease. Int J Tuberc Lung Dis 2021;25:797-813.

Ann Acad Med Singap 2024;53:187-207
https://doi.org/10.47102/annals-acadmedsg.2023113

REVIEW ARTICLE

Clinical performance of automated machine learning: A systematic review

Arun James Thirunavukarasu[*1,2] *MB BChir*, Kabilan Elangovan[1] *BEng*, Laura Gutierrez[1] *MD*, Refaat Hassan[2] *MB BChir*, Yong Li[1,3] *MD*, Ting Fang Tan[1] *MBBS*, Haoran Cheng[1,3,4] *MPH*, Zhen Ling Teo[5] *FRCOphth*, Gilbert Lim[1] *PhD*, Daniel Shu Wei Ting[*1,3,5] *PhD*

ABSTRACT

Introduction: Automated machine learning (autoML) removes technical and technological barriers to building artificial intelligence models. We aimed to summarise the clinical applications of autoML, assess the capabilities of utilised platforms, evaluate the quality of the evidence trialling autoML, and gauge the performance of autoML platforms relative to conventionally developed models, as well as each other.

Method: This review adhered to a prospectively registered protocol (PROSPERO identifier CRD42022344427). The Cochrane Library, Embase, MEDLINE and Scopus were searched from inception to 11 July 2022. Two researchers screened abstracts and full texts, extracted data and conducted quality assessment. Disagreement was resolved through discussion and if required, arbitration by a third researcher.

Results: There were 26 distinct autoML platforms featured in 82 studies. Brain and lung disease were the most common fields of study of 22 specialties. AutoML exhibited variable performance: area under the receiver operator characteristic curve (AUCROC) 0.35–1.00, F1-score 0.16–0.99, area under the precision-recall curve (AUPRC) 0.51–1.00. AutoML exhibited the highest AUCROC in 75.6% trials; the highest F1-score in 42.3% trials; and the highest AUPRC in 83.3% trials. In autoML platform comparisons, AutoPrognosis and Amazon Rekognition performed strongest with unstructured and structured data, respectively. Quality of reporting was poor, with a median DECIDE-AI score of 14 of 27.

Conclusion: A myriad of autoML platforms have been applied in a variety of clinical contexts. The performance of autoML compares well to bespoke computational and clinical benchmarks. Further work is required to improve the quality of validation studies. AutoML may facilitate a transition to data-centric development, and integration with large language models may enable AI to build itself to fulfil user-defined goals.

Ann Acad Med Singap 2024;53:187-207

Keywords: AI, artificial intelligence, automated machine learning, autoML, machine learning, deep learning

CLINICAL IMPACT

What is New

- This systematic review identified 26 distinct autoML platforms that have been trialled and/or applied in a clinical context.
- AutoML exhibited variable performance; in head-to-head comparisons, AutoPrognosis and Amazon Rekognition performed the strongest with unstructured and structured data, respectively.

Clinical Implications

- The performance of autoML compares well to bespoke computational and clinical benchmarks across clinical tasks ranging from diagnosis to prognostication.
- Exemplar use cases include identifying pathology on common imaging modalities (e.g. chest X-ray) and predicting hospitalisation and mortality based on tabulated demographics and blood test results.
- AutoML may facilitate a transition from model-centric to data-centric development, and integration with large language models may enable automated development of AI applications to fulfil user-defined goals.

[1] Artificial Intelligence and Digital Innovation Research Group, Singapore Eye Research Institute, Singapore
[2] University of Cambridge School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom
[3] Duke-NUS Medical School, National University of Singapore, Singapore
[4] Rollins School of Public Health, Emory University, Atlanta, Georgia, USA
[5] Singapore National Eye Centre, Singapore
Correspondence: A/Prof Daniel Ting, Singapore Eye Research Institute, The Academia, 20 College Road, Level 6 Discovery Tower, Singapore 169856.
Email: daniel.ting@duke-nus.edu.sg
Dr Arun Thirunavukarasu, University of Cambridge School of Clinical Medicine, University of Cambridge, Addenbrooke's Hospital, Hills Rd, Cambridge CB2 0SP, United Kingdom.
Email: ajt205@cantab.ac.uk

INTRODUCTION

In medicine, machine learning (ML) has been applied in a wide variety of contexts ranging from administration to clinical decision support, driven by greater availability of healthcare data and technological development.[1-5] Automated ML (autoML) enables individuals without extensive computational expertise to access and utilise powerful forms of artificial intelligence (AI) to develop their own models.[6] AutoML thereby enables developers to focus on curating high-quality data rather than optimising models manually, to facilitate a transition from model-driven to data-driven workflows.[7] AutoML has been posited as a means of improving the reproducibility of ML research, and even generating superior model performance relative to conventional ML techniques.[8]

AutoML technologies aim to automate some or all of the ML engineering process, which otherwise requires advanced data or computer science skills.[6] The first stage is data preparation, involving data integration, transformation and cleaning. Next is feature selection, where aspects of the data to be utilised in designing the ML model are decided; this may involve data imputation, categorical encoding and feature splitting.[9] Model selection, training and optimisation are then executed, with model performance evaluation being critical for identification of an optimal solution. AutoML systems use various methods and optimisation techniques to achieve state-of-the-art performance in some or all of the engineering process, such as Bayesian optimisation, random search, grid search, evolutionary based neural architecture selection, and meta-learning.[8,10] The optimised model may then be outputted for further work, such as clinical deployment, explainability analysis or external validation.

AutoML exhibits 4 major strengths, which may support its application in clinical practice and research. First, studies have reported comparable performance of autoML to conventionally developed models.[11] This raises the possibility of clinical deployment of autoML models and use in pilot studies preceding further model development. Second, autoML may improve the reproducibility of ML research by reducing the influence of human technicians who currently engage with an idiosyncratic process of tuning until a satisfactory result is achieved; this supports a transition toward more reproducible data-centric development.[7] Third, the reduction in computational experience and hardware conferred by autoML adoption should act as a major democratising force, providing a much larger number of clinicians with access to AI technology.[10] Last, the time spent on developing models is significantly reduced with autoML, as manual tuning is abolished—this improves efficiency and facilitates an acceleration of exploratory research to establish potential applications of AI.[10,12]

With the myriad of available autoML tools, democratisation of AI beyond those with clinical and computational expertise is feasible, and potential applications are diverse.[10,11] However, rigorous validation is necessary to justify deployment. Here, a systematic review was conducted to examine the performance of autoML in clinical settings. We aimed to evaluate the quality of result reporting; describe the specialties and clinical tasks in which autoML has been applied; and compare the performance of autoML platforms with conventionally developed models, as well as each other.

METHOD

The reporting of this study adheres to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidance, and the systematic review protocol was prospectively registered on international prospective register of systematic reviews (PROSPERO) (identifier CRD42022344427).[13,14] The protocol was amended to use a second quality assessment tool, the Prediction model Risk Of Bias ASsessment Tool (PROBAST) in addition to the Developmental and Exploratory Clinical Investigations of DEcision support systems driven by Artificial Intelligence (DECIDE-AI) reporting guidelines, as described below.

Data sources and searches

The Cochrane Library, Embase (via Ovid), MEDLINE (via Ovid) and Scopus were searched from inception until 11 July 2022, with no initial restrictions on publication status or type. Our search strategy isolated autoML in clinical contexts with the use of Boolean operators, as detailed in Supplementary Appendix S1. Before screening, duplicates were removed using Zotero version 6.0.14 (Corporation for Digital Scholarship, Vienna, VA, US) and Rayyan.[13,15]

Study selection

Abstract screening was conducted in Rayyan by 2 independent researchers, with a separate third arbitrator possessing autoML expertise who resolved cases of disagreement.[15] Full-text screening was similarly conducted by 2 researchers with a separate arbitrator, in Microsoft Excel for Mac version 16.57 (Microsoft Corporation, Redmond, WA, US). The explicit, hierarchical criteria for inclusion during abstract

and full-text screening are listed below in descending order of hierarchy, with full details provided in Supplementary Appendix S2:

(1) Written in the English language
(2) Peer-reviewed primary research article
(3) Not a retracted article
(4) Utilises autoML
(5) AutoML is applied in a clinical context

Data extraction and quality assessment

For articles satisfying the inclusion criteria, data extraction was conducted by 2 researchers: a clinical researcher extracting data first, with subsequent verification by a second computational researcher. Quality assessment was conducted by a single researcher, using implicit criteria based on the DECIDE-AI framework.[16] Risk of bias and concerns regarding applicability were assessed similarly by 2 researchers using the PROBAST framework and guidance questions.[17]

Other data collected include citation details; the trialled autoML platform/s; processing location (cloud or local); code intensity of the autoML platform; technical features of the autoML platform; clinical task autoML was applied towards; medical or surgical specialty defined anatomically where possible; sources of data used to train and test models; training and validation dataset size; dataset format (i.e. structured or unstructured); evaluation metrics used to gauge performance; and benchmark figures if presented, such as with comparisons to expert clinician or conventional ML performance. Specifically, figures for area under the receiver operator characteristic curve (AUCROC), F1-score, and area under the precision-recall curve (AUPRC) were collected. If F1-score was not provided but precision (positive predictive value) and recall (sensitivity) were, F1-score was calculated as the harmonic mean of the 2 metrics. If metrics were not stated in text form but were clearly plotted in graphical form, figures were manually interpolated using WebPlotDigitizer v4.6.0 (Ankit Rohatgi, Pacifica, CA, US). Metrics were excluded if the source or modality of the tested model was unclear.[18] Where researchers disagreed, resolution was achieved through discussion or arbitration by a third researcher.

Data synthesis and analysis

A narrative synthesis was conducted because meta-analysis was precluded by heterogeneity of datasets, platforms and use cases. Quantitative comparisons of autoML models were based on performance metrics (F1-score, AUCROC and AUPRC) to judge the clinical utility of applied autoML.[19] AutoML platforms were compared using the same metrics where platforms were applied to an identical task with the same training and testing data. A statistically significant difference in metrics was defined as featuring non-overlapping 95% confidence intervals. To establish the congruence between studies' conclusions and their presented data, the discussion and conclusion sections of each study were appraised by a single researcher to identify if autoML was compared with conventional techniques, and if so whether the comparison favoured autoML, conventional techniques or neither. AutoML platforms were compared in terms of their requirements and capabilities, with researchers contacted to clarify any questions regarding code intensity, processing location or data structure. Figures were produced with R version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria),[20-22] and Affinity Designer version 1.10.4 (Pantone, Carlstadt, NJ, US).

RESULTS

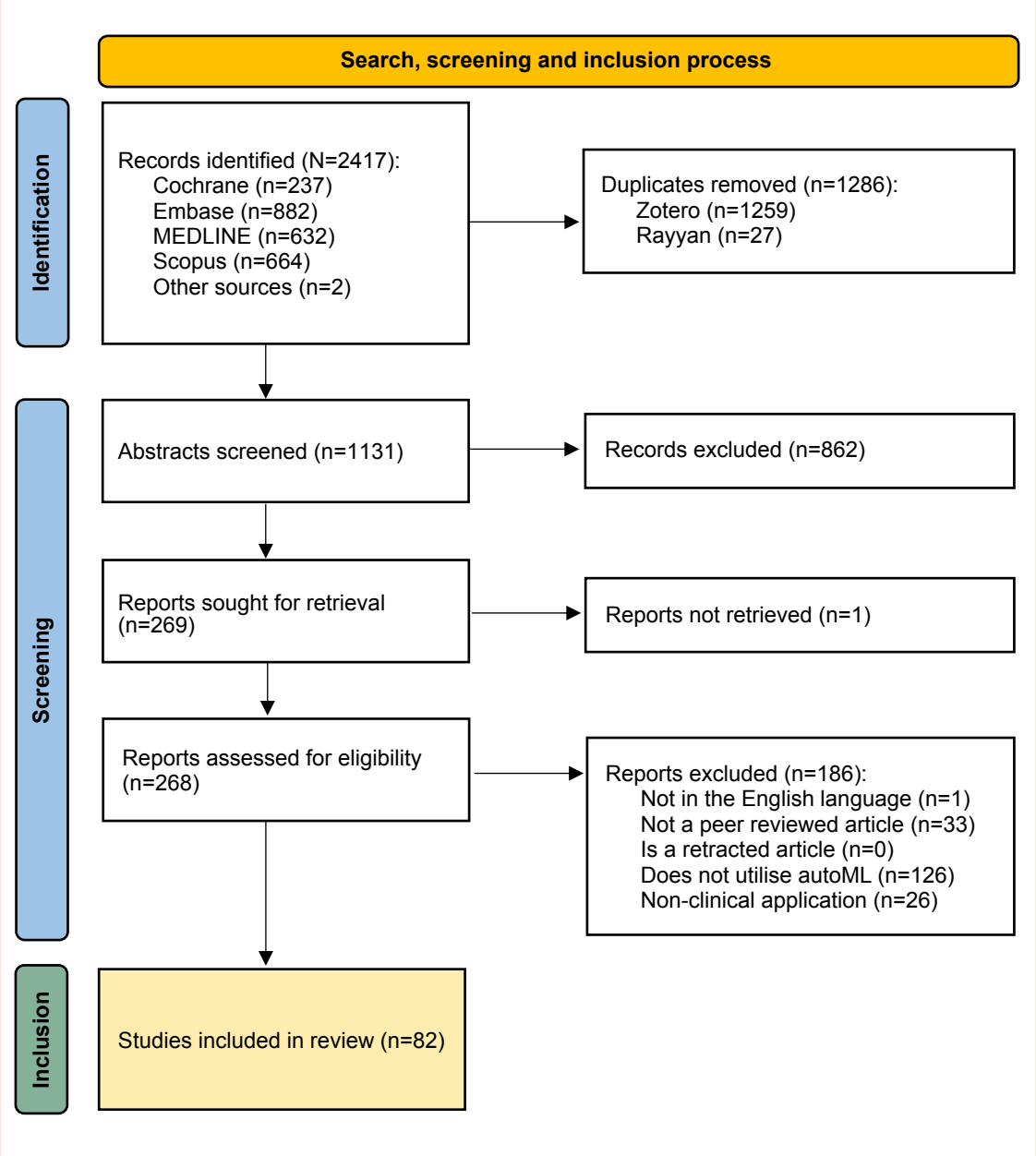
Record inclusion

Of 2417 records initially identified, 82 were included in the final analysis (Fig. 1 and Supplementary Appendix S3). In rare cases, research reports referred to autoML or similar terms in the broader context of "ML that automates", despite not utilising autoML technology: these articles were excluded for not meeting inclusion criterion 4.[23,24] Other borderline cases considered to be outside the scope of this review based on inclusion criterion 5 involved uses of autoML in clinical contexts, but without contributing to patient diagnosis, management or prognosis. These included a classifying surgical performance exhibited on video recordings and prediction of biological sex from medical images.[25,26]

Characteristics of included studies

The characteristics of the 82 included studies are summarised in Fig. 2 and Supplementary Appendix S3. To our knowledge, AutoML first entered medical literature in 2018 and has been growing in impact ever since: 1 paper in 2018, 7 in 2019, 21 in 2020, 35 in 2021, and 18 by 11 July 2022. Use cases are diverse, but diagnostic tasks (53 studies) were more common than management (4 studies) or prognostic (25 studies) tasks. The most common specialties in which autoML was used were pulmonology and neurology. Structured (e.g. tabulated) and unstructured (e.g. imaging) data were used similarly commonly. Dataset size varied widely, from 31 to 2,185,920 for training; 8 to 2,185,920 for internal validation; and 27 to 34,128 for external validation.

Fig. 1. PRISMA flow chart depicting the search, screening and inclusion process of this review.



Quality of reporting is summarised in Fig. 2F, with individual scores reported in Supplementary Appendix S4. The median number of fulfilled DECIDE-AI criteria was 14 out of 27, with the highest score being 19 out of 27. Nine criteria were fulfilled by over 90% of included studies. Thirteen criteria were not fulfilled in over half of the included studies, their DECIDE-AI criteria item numbers and items are as follows: (III) Research governance, (3) Participants, (5) Implementation, (6) Safety and errors in the results, (7) Human factors, (8) Ethics, (VI) Patient involvement, (9) Participants, (10) Implementation, (11) Modifications, (13) Safety and errors in results, (14) Human factors, and (16) Safety and errors in the discussion. Of these, 3 criteria were not fulfilled by any of the 82 included studies: (8) Ethics; (VI) Patient involvement; and (13) Safety and errors in the results.

Risk of bias and concerns regarding applicability are summarised in Fig. 2G. The most common sources of bias were retrospective study design that often used publicly available datasets, rather than testing autoML models in prospective trials to validate clinical performance and establish generalisability; and failure to provide an appropriate bespoke computational or clinical benchmark to demonstrate the performance of autoML—conferring unclear or high risk of bias in PROBAST

Fig. 2. Collectively summarised characteristics of included studies: (A) Date of publication bar chart. (B) Continent of corresponding author bar chart. (C) Country of corresponding author bar chart. (D) Funding source bar chart. (E) Clinical specialty bar chart. (F) DECIDE-AI score histogram. (G) PROBAST evaluation bar chart. (H) Dataset size histogram with logarithmic X-axis. (I) Data nature bar chart. (J) Data source bar chart.

CoA: concerns of applicability; RoB: risk of bias

appraisal (Supplementary Appendix S5). In many cases, this was because autoML was used as a tool rather than the study being a trial of autoML technology. However, a statement was made in the discussion or conclusion regarding the effectiveness of autoML in 27 of 47 studies (57%) judged to have a high or unclear risk of bias.

AutoML performance relative to other modalities

The reporting of performance metrics varied widely between papers, likely representing the inherent limitations of applied autoML platforms. There were 79 studies (96%) that provided AUCROC (Fig. 3), F1-score (Fig. 4), or AUPRC (Supplementary Appendix S6) as a measure of performance. Of these, 35 studies (44%) reported a computational or clinical benchmark to compare autoML performance against, and 21 studies (27%) provided 95% confidence intervals for estimates of performance metrics. Of 12 studies (15%) with benchmark comparisons and confidence intervals, autoML exhibited statistically significantly superior AUCROC in 6 of 17 trials (35%); significantly superior F1-score in 0 of 1 trial; and significantly superior AUPRC in 0 of 2 trials. In studies with benchmark comparisons and confidence intervals, autoML did not exhibit the lowest AUCROC, F1-score, or AUCPR in any trial. In all studies comparing modalities, autoML exhibited the highest AUCROC in 28 of 37 trials (76%), the highest F1-score in 11 of 26 trials (42%), and the highest AUPRC in 10 of 12 trials (83%). AutoML exhibited the lowest AUCROC in 5 of 37 trials (14%); the lowest F1-score in 6 of 26 trials (23%); and the lowest AUPRC in 2 of 12 trials (17%). For autoML models, AUCROC ranged from 0.346–1.000 (scores of 0.5 are equivalent to chance; maximum score=1); F1-score ranged from 0.128–0.992 (maximum score=1); and AUPRC ranged from 0.280–1.000 (maximum score=1).

There were 57 studies (70%) that compared autoML to other conventional modelling methods in the prose of their discussion or conclusion. Of these, 28 suggested that autoML was superior to conventional methods; 29 suggested that autoML was comparable to conventional methods; and none suggested that autoML was inferior to conventional methods. Only 35 studies provided a quantitative comparison in their results, as described above (Fig. 3, Fig. 4 and Supplementary Appendix S6). Conclusions of comparable effectiveness were justified by congruence with reported performance metrics in 16 of 29 studies (55%); conclusions of superior effectiveness of autoML were justified in 11 of 28 studies (39%).

Comparative performance of autoML platforms

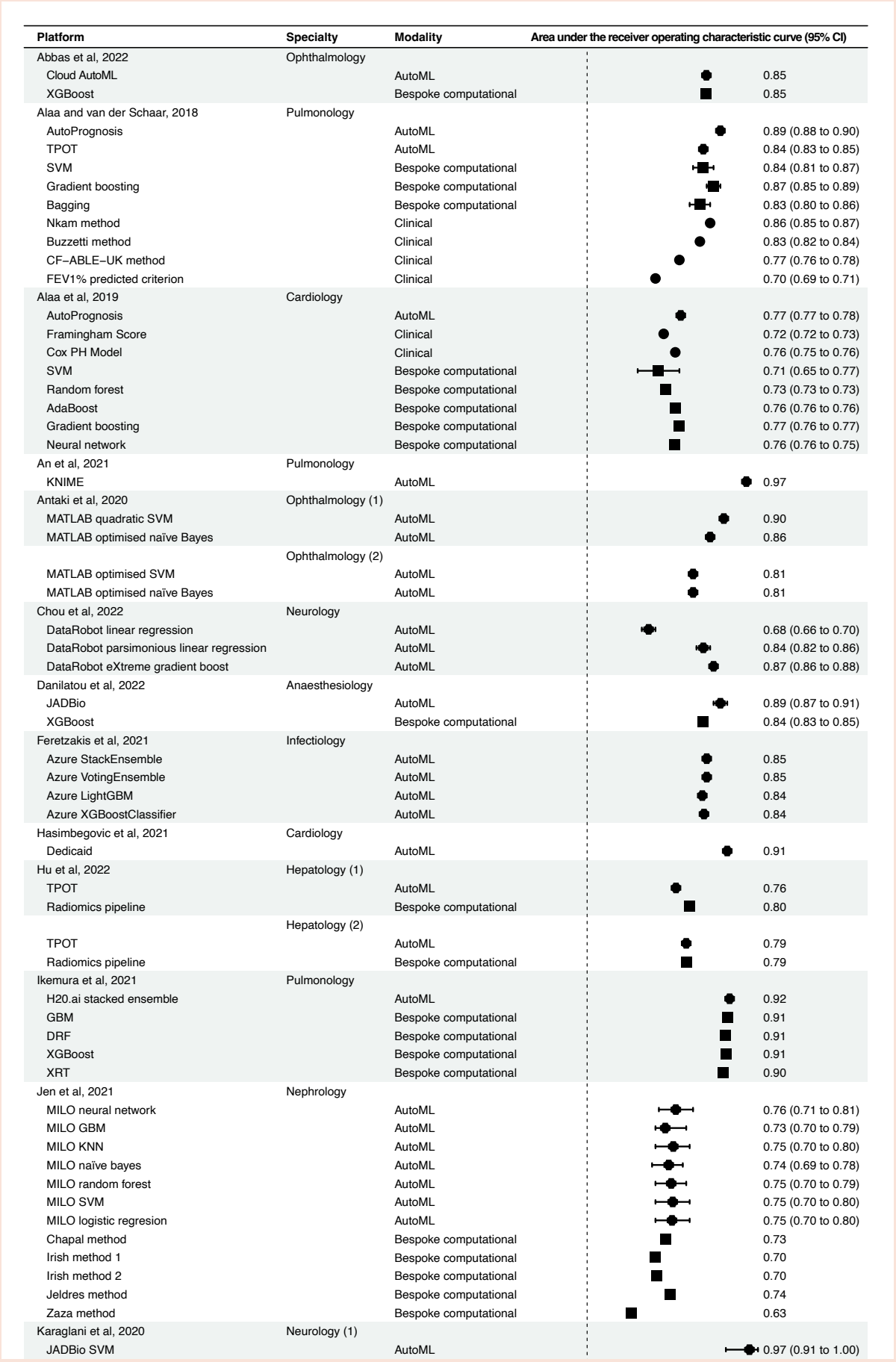
A comparative summary of the autoML platforms validated in the literature is presented in Table 1. Platforms vary greatly in their accessibility, technical features and portability. While performance in different tasks cannot be compared, 5 studies directly compared distinct autoML platforms in the same task. Of these, 1 study (20%) provided AUCROC metrics, which favoured AutoPrognosis over Tree-based Pipeline Optimization Tool (TPOT) to prognosticate mortality in cystic fibrosis.[27] Four studies (80%) provided F1-score metrics for a total of 9 trials (Fig. 5) on: prognosticating mortality in cystic fibrosis; predicting invasion depth of gastric neoplasms from endoscopic photography; diagnosing referrable diabetic retinopathy from fundus photography; diagnosing age-related macular degeneration, central serous retinopathy, macular hole and diabetic retinopathy from optical coherence tomography (OCT); diagnosing choroidal neovascularisation, diabetic macular oedema and drusen from OCT; and classifying spine implants from lumbar spine radiographs.[27-30]

AutoPrognosis (structured data) and Rekognition (unstructured data) exhibited the strongest performance as they were superior to every platform they were compared with, although this was only to TPOT for AutoPrognosis, and Rekognition was compared with fewer platforms than Cloud AutoML. Two studies (40%) reporting 5 trials provided AUPRC metrics for prognosticating mortality in cystic fibrosis and classifying electrocardiogram traces.[27,31] Here, performance favoured AutoPrognosis over TPOT; and AutoDAL-SOAR over USDM, AER, Auto-Weka, Auto-Sklearn and ASSL+US. While not all platforms can be compared against one another due to incompatibility with data structure, many possible combinations were not trialled and the number of comparative trials was small, making it difficult to establish comparative performance.

Confidence in conclusions

Confidence in conclusions is tempered by high risk of bias, particularly in retrospective study design and limited metrics facilitating statistical comparisons. However, as autoML did not exhibit statistically significantly worse performance than conventional techniques in any trial and exhibited lower performance metrics than conventional trials in a minority of studies, there is high confidence in the conclusion that autoML technology facilitates production of models with comparable performance to conventional techniques such as bespoke computational approaches. Given the low number of studies providing confidence intervals to enable statistical comparison of models' performance within trials, conclusions regarding the superiority of autoML relative to conventional techniques have low confidence. In addition, conclusions cannot be assumed to be generalisable to all use cases and datasets, given that performance is highly context-specific as demonstrated by the large variability observed in AUCROC (Fig. 3), F1-score (Fig. 4) and AUPRC (Supplementary Appendix S6). Confidence in the superior performance of AutoPrognosis with

Fig. 3. Forest plot depicting reported AUCROC metrics.



Platform	Specialty	Modality	Area under the receiver operating characteristic curve (95% CI)
Abbas et al, 2022	Ophthalmology		
Cloud AutoML		AutoML	0.85
XGBoost		Bespoke computational	0.85
Alaa and van der Schaar, 2018	Pulmonology		
AutoPrognosis		AutoML	0.89 (0.88 to 0.90)
TPOT		AutoML	0.84 (0.83 to 0.85)
SVM		Bespoke computational	0.84 (0.81 to 0.87)
Gradient boosting		Bespoke computational	0.87 (0.85 to 0.89)
Bagging		Bespoke computational	0.83 (0.80 to 0.86)
Nkam method		Clinical	0.86 (0.85 to 0.87)
Buzzetti method		Clinical	0.83 (0.82 to 0.84)
CF–ABLE–UK method		Clinical	0.77 (0.76 to 0.78)
FEV1% predicted criterion		Clinical	0.70 (0.69 to 0.71)
Alaa et al, 2019	Cardiology		
AutoPrognosis		AutoML	0.77 (0.77 to 0.78)
Framingham Score		Clinical	0.72 (0.72 to 0.73)
Cox PH Model		Clinical	0.76 (0.75 to 0.76)
SVM		Bespoke computational	0.71 (0.65 to 0.77)
Random forest		Bespoke computational	0.73 (0.73 to 0.73)
AdaBoost		Bespoke computational	0.76 (0.76 to 0.76)
Gradient boosting		Bespoke computational	0.77 (0.76 to 0.77)
Neural network		Bespoke computational	0.76 (0.76 to 0.75)
An et al, 2021	Pulmonology		
KNIME		AutoML	0.97
Antaki et al, 2020	Ophthalmology (1)		
MATLAB quadratic SVM		AutoML	0.90
MATLAB optimised naïve Bayes		AutoML	0.86
	Ophthalmology (2)		
MATLAB optimised SVM		AutoML	0.81
MATLAB optimised naïve Bayes		AutoML	0.81
Chou et al, 2022	Neurology		
DataRobot linear regression		AutoML	0.68 (0.66 to 0.70)
DataRobot parsimonious linear regression		AutoML	0.84 (0.82 to 0.86)
DataRobot eXtreme gradient boost		AutoML	0.87 (0.86 to 0.88)
Danilatou et al, 2022	Anaesthesiology		
JADBio		AutoML	0.89 (0.87 to 0.91)
XGBoost		Bespoke computational	0.84 (0.83 to 0.85)
Feretzakis et al, 2021	Infectiology		
Azure StackEnsemble		AutoML	0.85
Azure VotingEnsemble		AutoML	0.85
Azure LightGBM		AutoML	0.84
Azure XGBoostClassifier		AutoML	0.84
Hasimbegovic et al, 2021	Cardiology		
Dedicaid		AutoML	0.91
Hu et al, 2022	Hepatology (1)		
TPOT		AutoML	0.76
Radiomics pipeline		Bespoke computational	0.80
	Hepatology (2)		
TPOT		AutoML	0.79
Radiomics pipeline		Bespoke computational	0.79
Ikemura et al, 2021	Pulmonology		
H20.ai stacked ensemble		AutoML	0.92
GBM		Bespoke computational	0.91
DRF		Bespoke computational	0.91
XGBoost		Bespoke computational	0.91
XRT		Bespoke computational	0.90
Jen et al, 2021	Nephrology		
MILO neural network		AutoML	0.76 (0.71 to 0.81)
MILO GBM		AutoML	0.73 (0.70 to 0.79)
MILO KNN		AutoML	0.75 (0.70 to 0.80)
MILO naïve bayes		AutoML	0.74 (0.69 to 0.78)
MILO random forest		AutoML	0.75 (0.70 to 0.79)
MILO SVM		AutoML	0.75 (0.70 to 0.80)
MILO logistic regresion		AutoML	0.75 (0.70 to 0.80)
Chapal method		Bespoke computational	0.73
Irish method 1		Bespoke computational	0.70
Irish method 2		Bespoke computational	0.70
Jeldres method		Bespoke computational	0.74
Zaza method		Bespoke computational	0.63
Karaglani et al, 2020	Neurology (1)		
JADBio SVM		AutoML	0.97 (0.91 to 1.00)

Fig. 3. Forest plot depicting reported AUCROC metrics. Cont'd

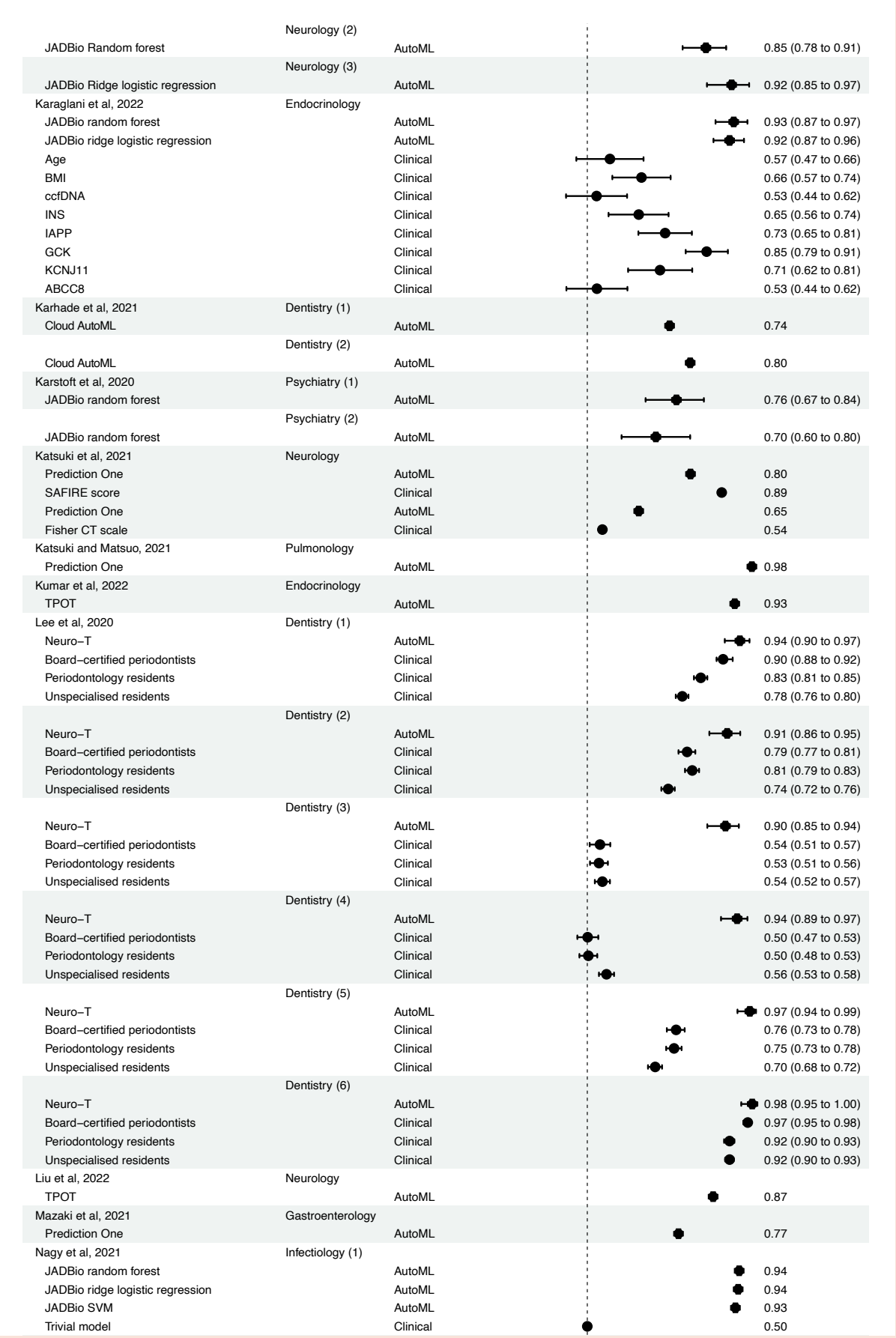


Fig. 3. Forest plot depicting reported AUCROC metrics. Cont'd

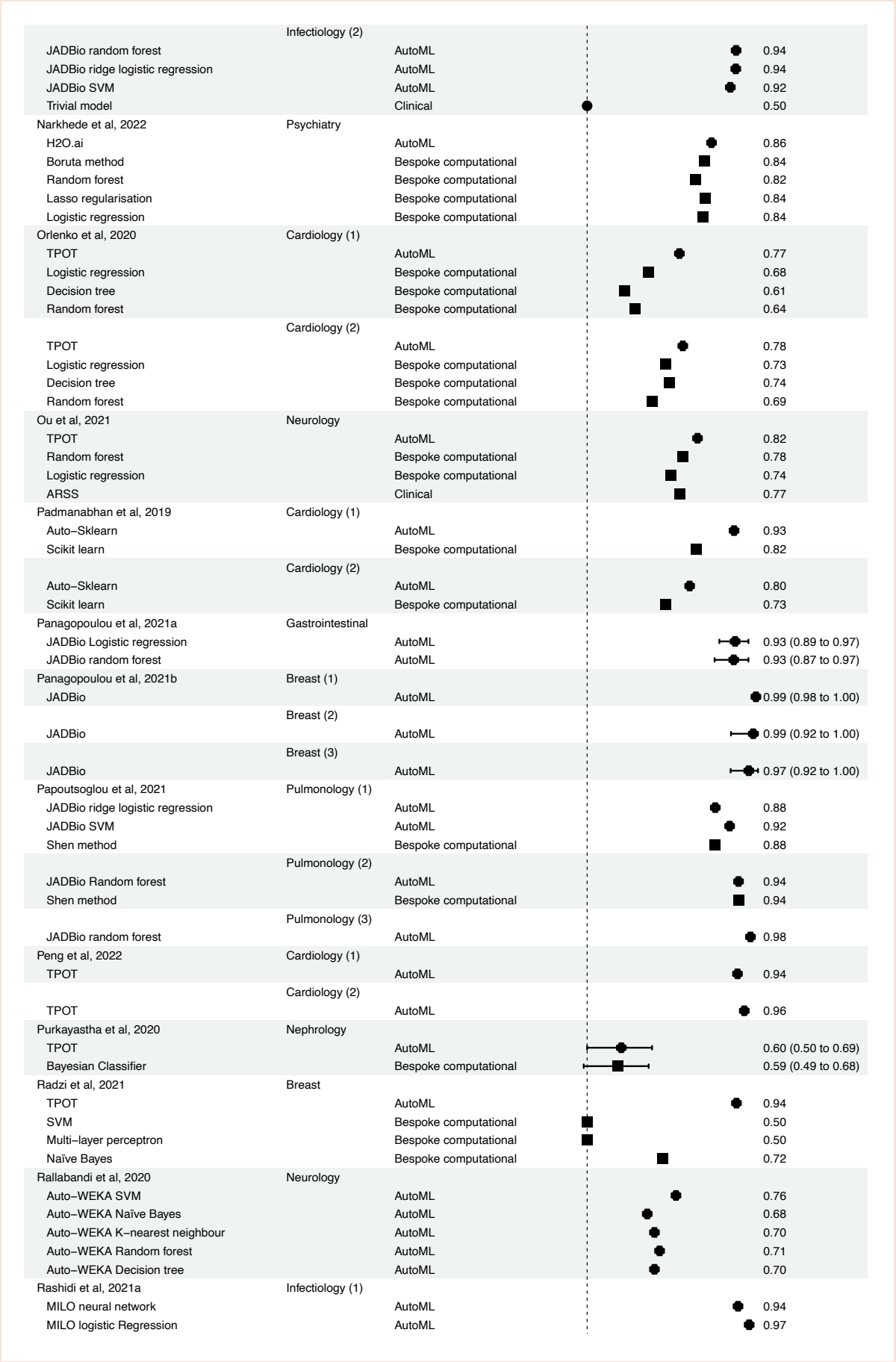


Fig. 3. Forest plot depicting reported AUCROC metrics. Cont'd

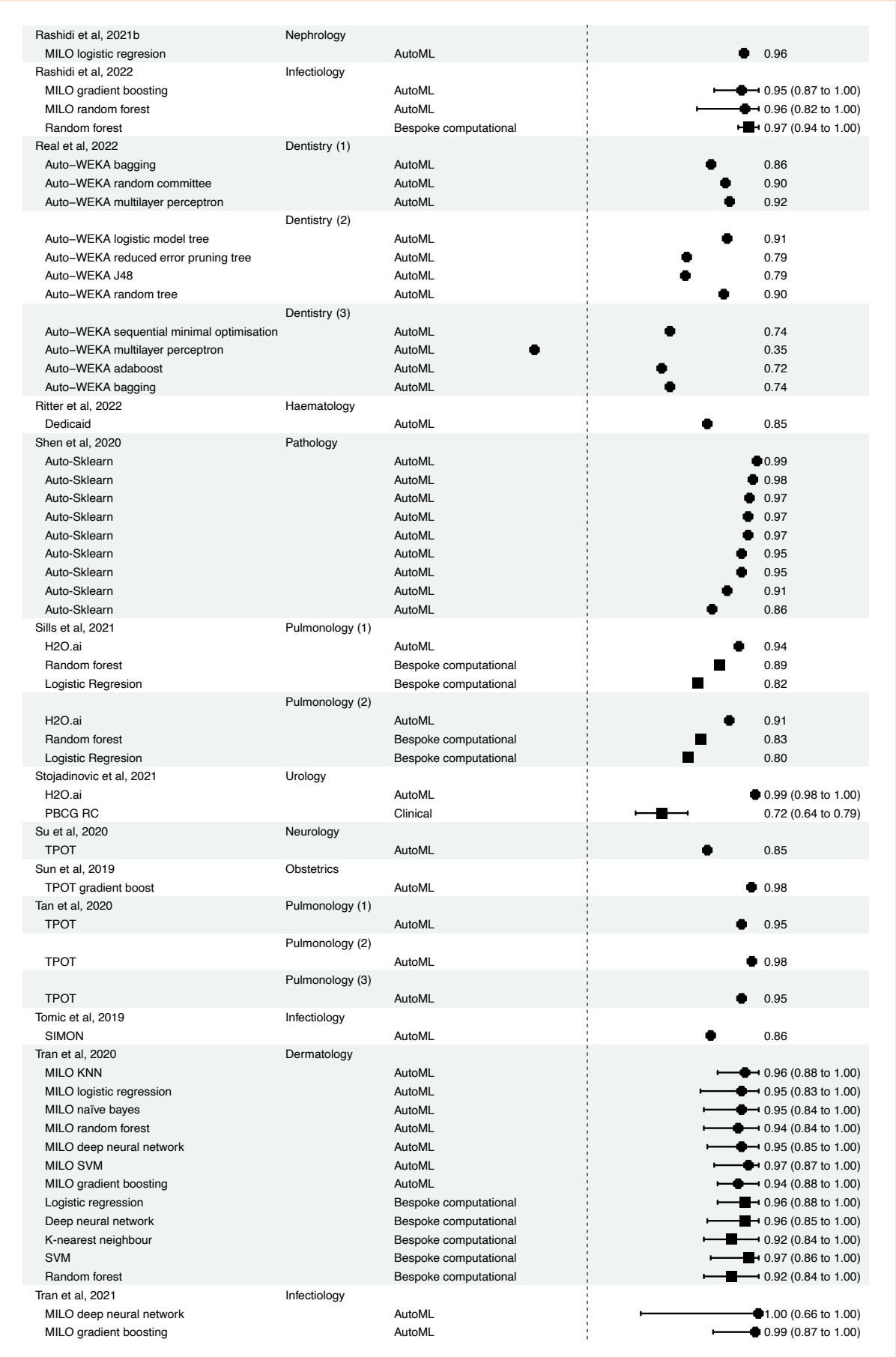
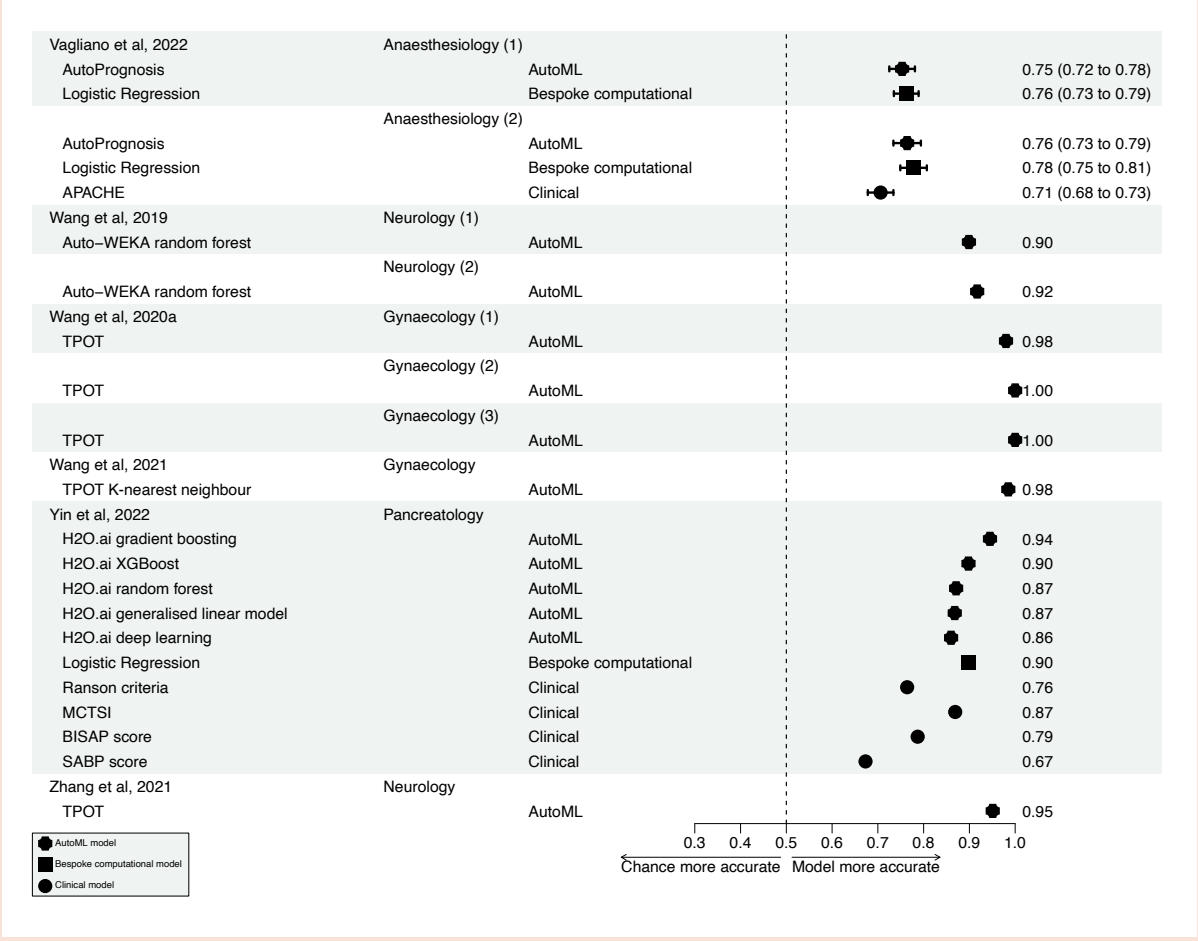


Fig. 3. Forest plot depicting reported AUCROC metrics. Cont'd



ARSS: Aneurysm Recanalization Stratification Scale; BMI: body-mass index; ccf-DNA: circulating cell-free DNA; CT: computerised tomography; DRF: distributed random forest; FEV1: forced expiratory volume in 1 second; GBM: gradient boosting machine; MCTSI: Modified Computed Tomography Severity Index; PH: proportional hazards; SVM: support vector machine; XRT: extremely randomised tree

structured data is very low, as there were very few comparative trials; and also low for the superior performance of Rekognition with unstructured data, as the number of comparative trials was low—though not as low as for structured data—and as there were no data for many possible platform comparisons.

DISCUSSION

This study shows that autoML has been trialled in a wide variety of diagnostic, patient management and prognostic tasks. AutoML has been used in many clinical specialties, most commonly in brain and lung imaging. Performance of autoML models generally compares well to bespoke computational and clinical benchmarks, often exhibiting superior performance. However, available studies and appraised risk of bias preclude conclusion that autoML provides universally superior performance to conventional modelling, as relative and absolute performance vary widely with the applied platform, use case, and data source. The strength of the evidence base supporting use of different autoML platforms is highly heterogenous, with some platforms exhibiting results more supportive of equivalence or superiority to conventional techniques than others. Few studies compared different autoML platforms to determine which provide optimal performance for a given task. Despite these knowledge gaps, a high number of non-comparative studies suggests that autoML is already being applied as a statistical tool, comparable to bespoke machine learning coding packages or statistical software.

There are 5 main deficiencies in the quality of the autoML evidence base. First, inconsistency in performance metrics may be a consequence of restrictions imposed by autoML platforms but

Fig. 4. Forest plot depicting reported F1-score metrics.

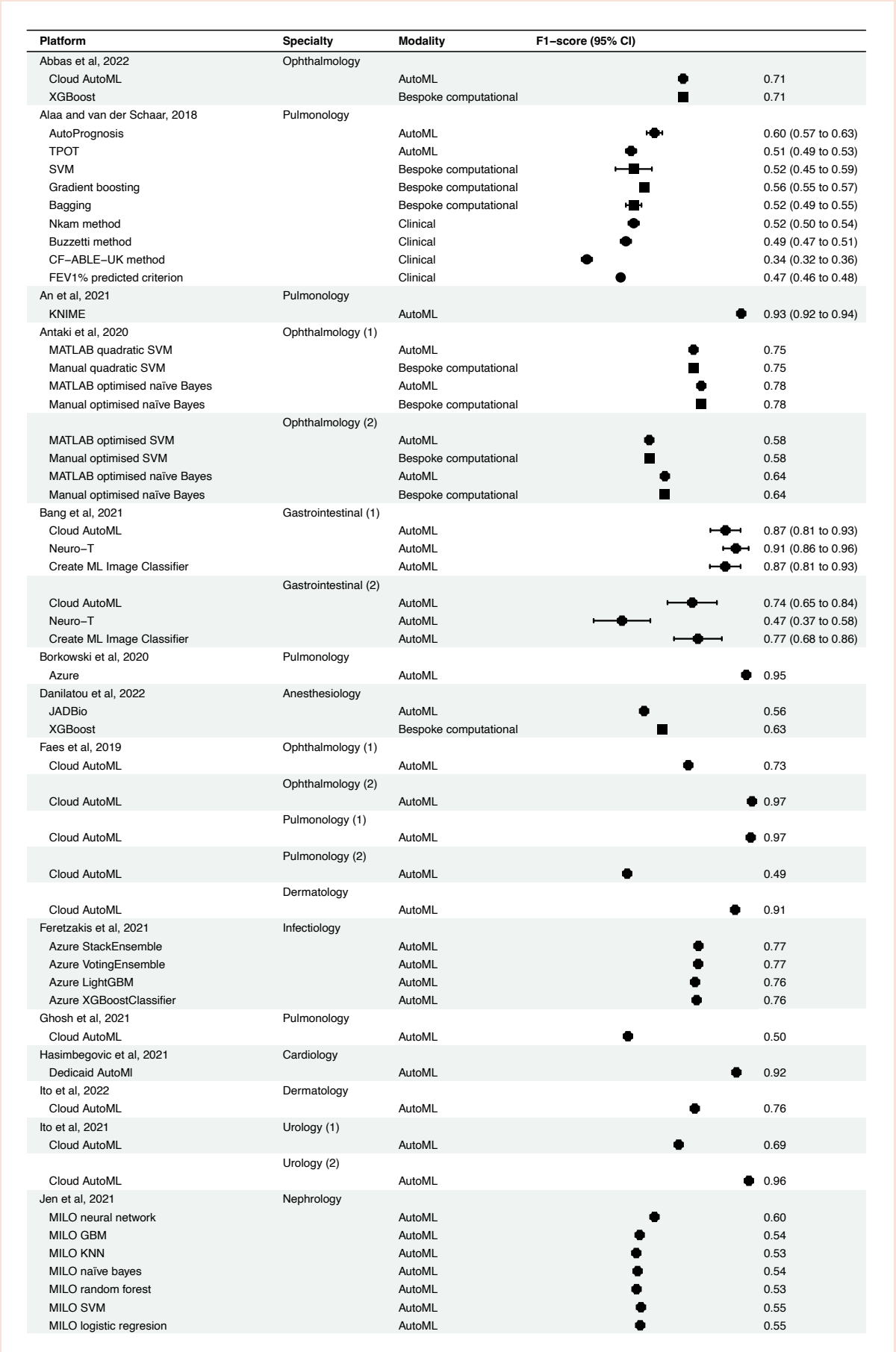


Fig. 4. Forest plot depicting reported F1-score metrics. Cont'd

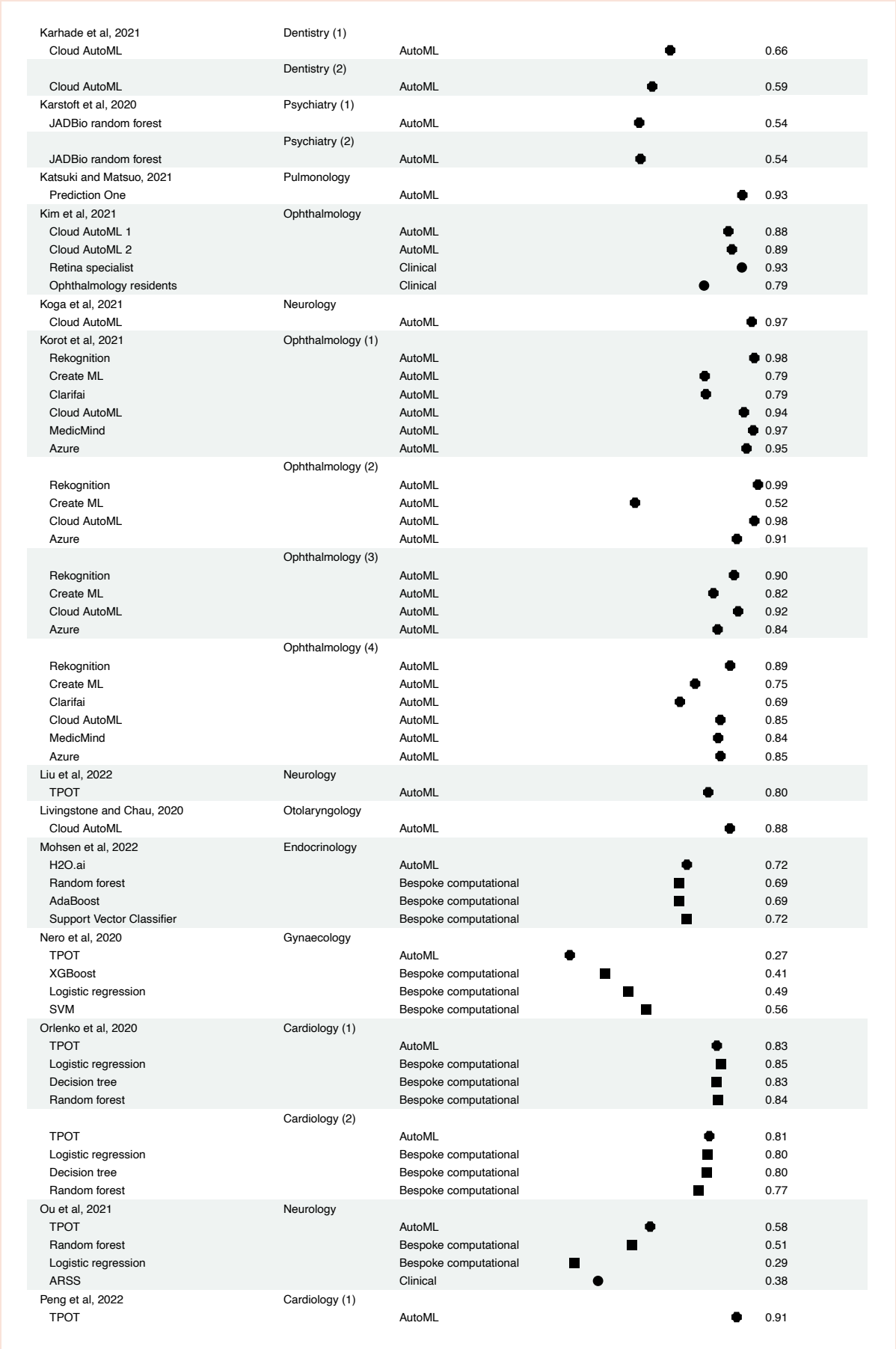


Fig. 4. Forest plot depicting reported F1-score metrics. Cont'd

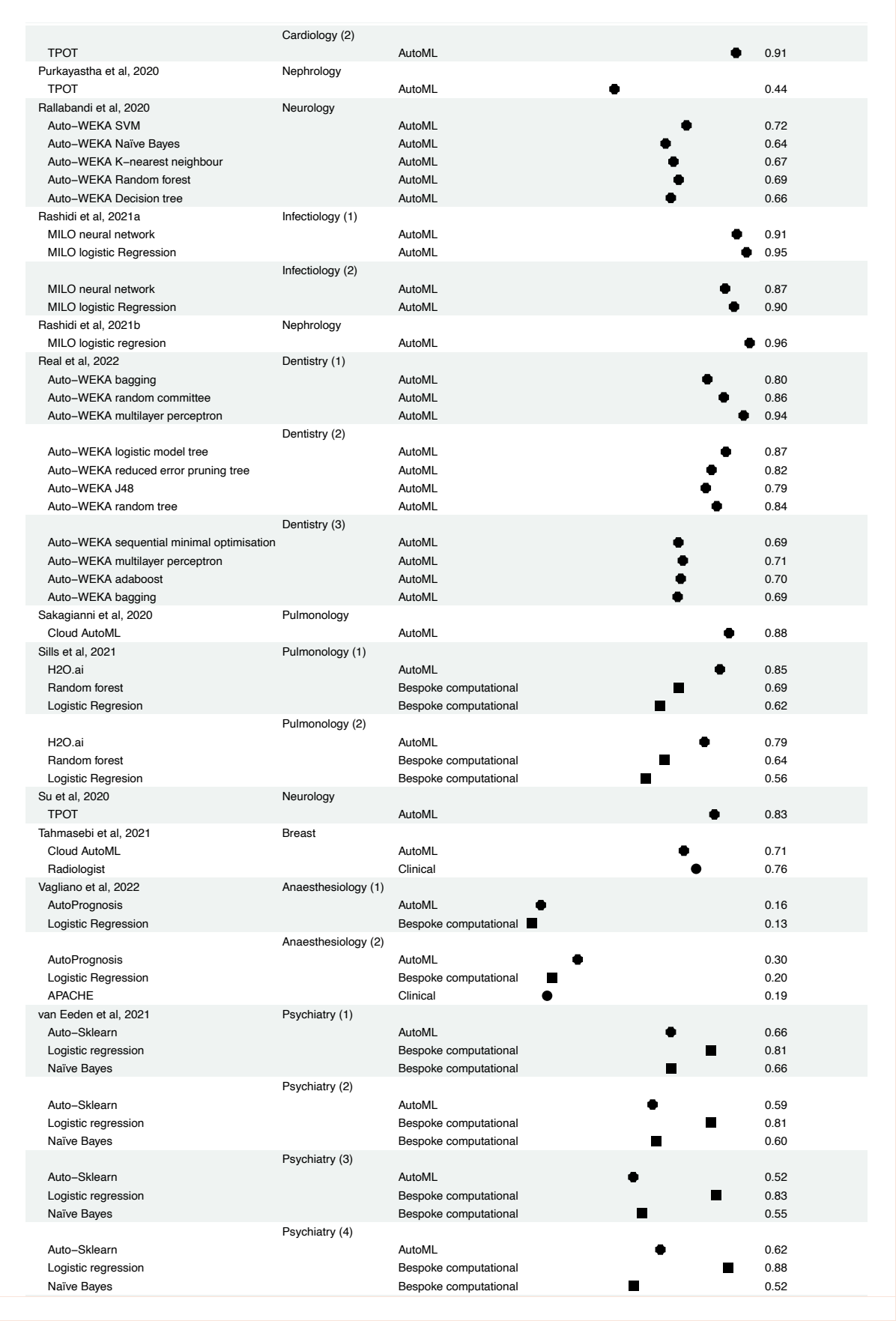
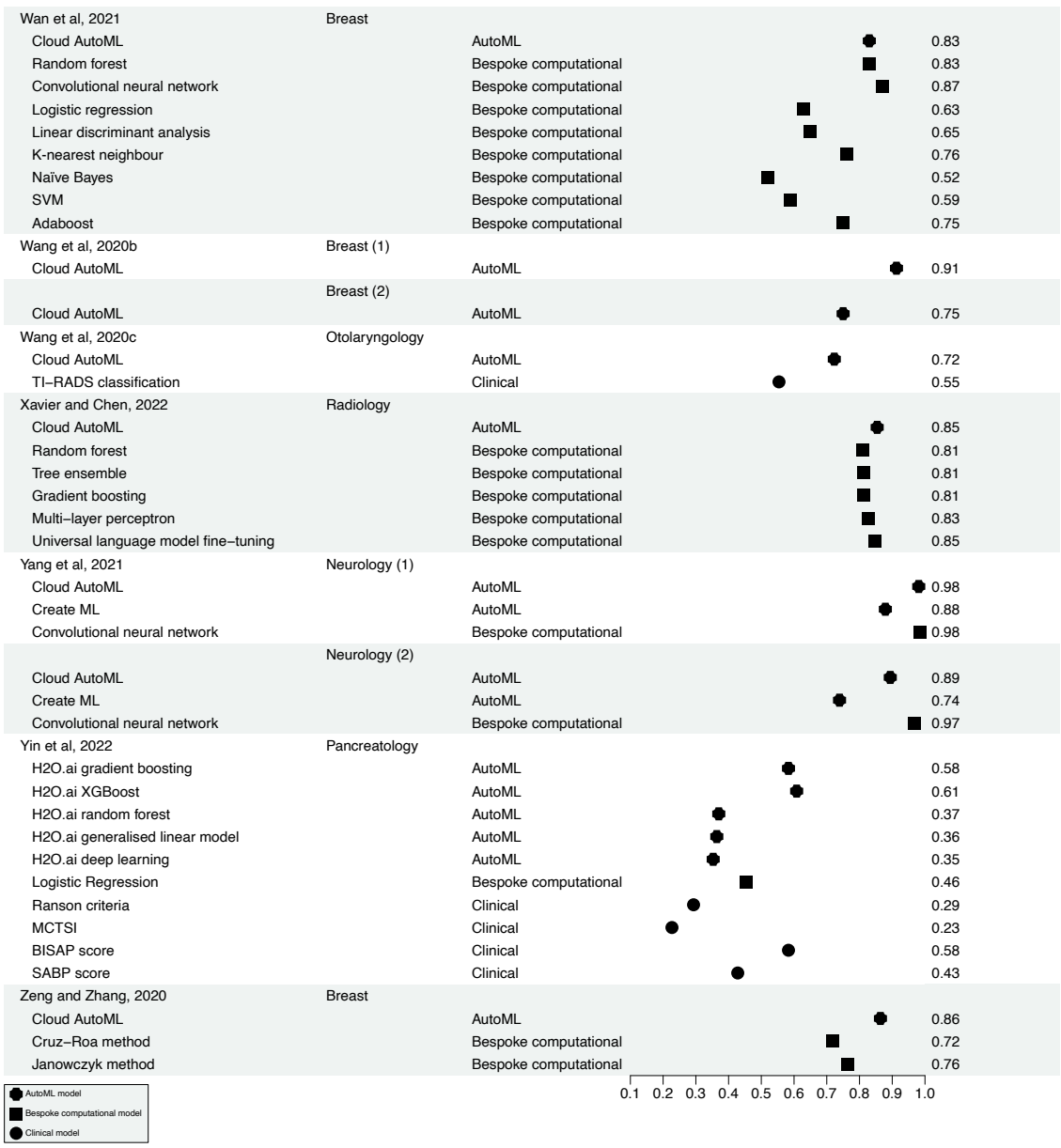


Fig. 4. Forest plot depicting reported F1-score metrics. Cont'd

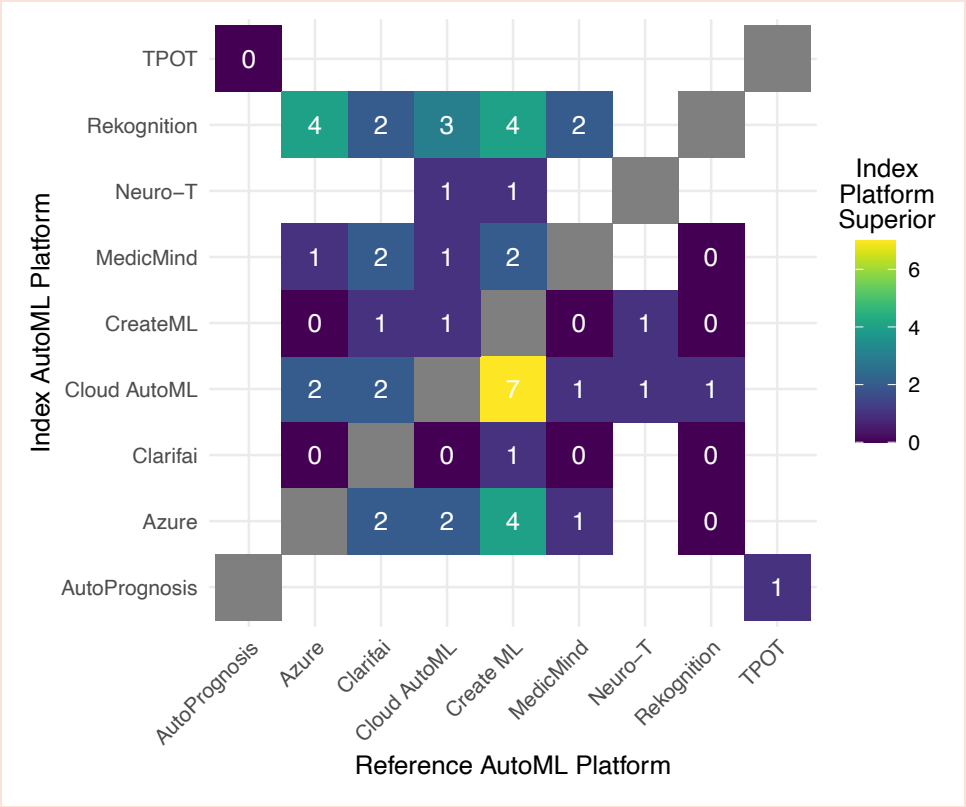


ARSS: Aneurysm Recanalization Stratification Scale; CI: confidence interval; FEV1: forced expiratory volume in 1 second; GBM: gradient boosting machine; SVM; support vector machine

observed variation between studies using similar platforms also suggests that selective reporting is common. Reporting comprehensive metrics is essential, particularly in the context of diagnostic algorithms, as some metrics are a function of prevalence or model threshold.[19] Second, explainability analysis is challenging for similar reasons to portability, but is possible with emerging technological solutions.[32] In addition, some platforms incorporate inbuilt explainability, such as by providing salience maps for deep learning models.[12] Issues regarding "black box" algorithms are accentuated in autoML research, leading to a third limitation: a lack of ethical consideration—such as regarding algorithmic fairness—by all the included studies.

Fourth, inconsistent use of benchmarking represents a form of publication bias leading to erroneous conclusions of equivalent or superior autoML performance relative to conventional bespoke computational methods or clinicians. Many studies relied on historical controls or provided no benchmark at all. To confidently conclude that autoML performance compares well to bespoke

Fig. 5. Heat map depicting the comparative performance of autoML platforms as applied to the same clinical tasks in terms of F1-score. Shading and numbers correspond to the number of superior performances exhibited by the index platform with respect to the reference platform.



models—and particularly to state-of-the-art techniques—a researcher with computational aptitude should have an opportunity to maximise performance. Finally, models should be deployed on separate datasets which were not used in testing or training, for external validation to demonstrate generalisability, which is a critical component of clinical potential.[6] Without external validation, overfitting to the datasets provided may lead to inflated estimates of performance.[6,33] External validation is limited on many autoML platforms by a lack of ability to batch test on new data, or to export models for analysis and deployment.

Limitations

This systematic review was limited by 3 issues: (1) PROBAST had to be adapted to apply it in non-diagnostic applications of autoML—we employed DECIDE-AI as a domain-specific quality indicator to mitigate this limitation, and utilised PROBAST in the context of trialling autoML technology rather than in validating models for clinical application. Development of more domain-specific tools to optimise AI-related systematic reviews is underway and will be a welcome development.[34,35] (2) Confidence in conclusions was affected by high risk of bias, a common theme in AI research more broadly.[36] We provide comprehensive indicators of quality, risk of bias and concerns regarding applicability to facilitate contextualisation of performance metrics. (3) It is difficult to draw conclusions for autoML as a modality because platforms are variable in their features, performances and requirements—future reviews may focus on individual platforms, although the number of studies featuring most platforms is very small.

Implications

Researchers applying a platform without providing benchmark comparators for the purposes of primary research or clinical work should justify their decision with validation data demonstrating that their approach is acceptable. Evidence should be contextually relevant, preferably pertaining to the same clinical task. While it is apparent that autoML has already begun to be applied in clinical research as a statistical tool, it is important that these tools are demonstrated to produce accurate, reliable and fair models. Studies purported as evidence of validation of autoML are often limited by retrospective design, high risk of bias, and unfulfill-

Table 1. Technological comparison of autoML platforms applied in the studies included in this review.

AutoML platform name	Accessibility			Technical features					Portability	
	Cost	Code requirement	Computing location	Dataset format	Feature extraction or selection	Model selection or training	Hyperparameter optimisation	Evaluation	Model exportability	Explainability
AER	Free	Coding required	Local	Structured/ Unstructured	Yes	Yes	Yes	No	Yes	No
Amazon Rekognition	Chargeable	None	Cloud	Unstructured	Yes	Yes	Yes	Yes	No	No
Apple Create ML	Free on Apple devices	None	Local	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	No
ASSL	Free	Coding required	Local	Strucured/ Unstructured	Yes	Yes	Yes	No	Yes	No
Auto-Sklearn	Free	Coding required	Local	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	No
Auto-WEKA	Free	Coding required	Local	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	No
AutoDAL	Free	Coding required	Local	Structured	No	Yes	Yes	Yes	Yes	No
AutoDC	Free	Coding required	Local	Structured	Yes	Yes	Yes	No	Yes	No
AutoPrognosis	Free	Coding required	Local	Structured	Yes	Yes	Yes	Yes	Yes	Yes
Clarifai	Chargeable	None	Cloud	Unstructured	Yes	Yes	Yes	Yes	No	No
Google Cloud AutoML	Chargeable	None	Cloud	Structured/ Unstructured	Yes	Yes	Yes	Yes	No	No
DataRobot AutoML	Chargeable	None	Cloud	Structured	Yes	Yes	Yes	Yes	Yes	Yes
Dedicaid AutoML	Restricted to collaborators	None	External private server	Structured	Yes	Yes	Yes	Yes	No	No
H2O.ai R/Python Packages	Free	Coding required	Local/Cloud	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	Yes
H2O.ai Driverless AI	Chargeable	None	Local/Cloud	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	Yes

Table 1. Technological comparison of autoML platforms applied in the studies included in this review. (Cont'd)

AutoML Platform Name	Accessibility			Technical Features					Portability	
	Cost	Code requirement	Computing location	Dataset format	Feature extraction or selection	Model selection or training	Hyperparameter optimisation	Evaluation	Model exportability	Explainability
JADBio	Chargeable	None	Cloud	Structured	Yes	Yes	Yes	Yes	Yes	Yes
KNIME	Chargeable	None	Local	Structured	No	Yes	Yes	Yes	Yes	No
MATLAB	Chargeable	Coding required	Local	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	No
MedicMind	Free	None	Cloud	Unstructured	Yes	Yes	Yes	Yes	Yes	Yes
Microsoft Azure AutoML	Chargeable	None	Cloud	Structured	Yes	Yes	Yes	Yes	No	No
MILO	Chargeable	None	Local	Structured	No	Yes	Yes	Yes	Yes	No
Neuro-T	Chargeable	None	Local	Unstructured	Yes	Yes	Yes	Yes	Yes	No
SIMON	Free	None	Local	Strcutured	Yes	Yes	Yes	Yes	Yes	No
Sony Prediction One	Chargeable	None	Local	Structured/ Unstructured	Yes	Yes	Yes	Yes	Yes	No
TPOT	Free	Coding required	Local	Structured	Yes	Yes	Yes	Yes	Yes	No
USDM	Free	Coding required	Local	Structured/ Unstructured	Yes	Yes	Yes	No	Yes	No

AER: approximated error reduction; ASSL: automated semi-supervised learning; AutoDAL: automated distributed active learning; AutoDC: automated data-centric processing; JADBio: Just-Add-Data Bio; KNIME: Konstanz Information Miner; MILO: Machine Intelligence Learning Optimizer; ML: machine learning; TPOT: Tree-based Pipeline Optimization Tool; USDM: uncertainty sampling with diversity maximization; WEKA: Waikato Environment for Knowledge Analysis

ment of conventional reporting standards—comparable to research regarding other AI technologies.[37] Future comparative studies should address the limitations discussed above to convince researchers, clinicians and policy makers that autoML platforms may be applied in lieu of bespoke modelling.[38]

When reporting AI algorithms tasked with a certain clinical job, it would be helpful to avoid ambiguity in terminology. We would suggest a complete restriction of the terms "automated machine learning", or "autoML" to those algorithms built with technology that automates some or all parts of the engineering process—all conventional ML models process data without human guidance, so description of these technologies as automated is redundant. Similar terms such as "automated artificial intelligence", "automated machine learning" and "automated deep learning" are also redundant in the context of bespoke computational models. A simple alternative term for conventional ML projects would be "automatic"—these systems may automate a particular task, but their development is not automated, the defining feature of autoML.

The reduced barrier to entry in terms of computational expertise and hardware requirements conferred by many autoML platforms makes them a powerful contributor to democratisation of AI technology—a far greater number of clinicians and scientists are capable of ML development through use of these platforms. AutoML could be an invaluable resource for teaching, as individuals can more rapidly develop hands-on experience, learn by trial-and-error, and thereby develop intuitive understanding of the capabilities and limitations of ML.[39] AutoML could also be applied in pilot studies, enabling clinicians with domain-specific expertise to explore possibilities for ML research—facilitating prioritisation of allocation of scarce resources such as access to graphics processing units and expert computer scientists.[6] Validated platforms may be applied more broadly, including in patient care. Moreover, autoML is well placed to respond to calls to inculcate data-centric AI as opposed to model-centric development; focusing effort on curating high-quality data, which limits development more often than code or model infrastructure.[7] Acceleration in this process may be facilitated by large language models as their emerging capability to leverage plugins will allow autoML to facilitate AI building itself to fulfil user-defined aims.[40]

Further work is indicated to improve validation of autoML platforms, either by allowing models to be exported, or by providing more comprehensive internal metrics. Other work should focus on improving the functionality of autoML, specifically on reducing the trade-offs currently implicit in selecting a platform with a given code intensity and computing locus. Using automation to reduce human error to optimise engineering and improve performance is one ideal—this has been demonstrated with structured data by AutoPrognosis.[27,41] Increased functionality of code-free platforms while retaining the customisability of code-intense solutions is another ideal—H2O.ai Driverless AI offers the same functionality as the H2O.ai R and Python packages, but with a code-free graphical user interface.[42-48] Alternatively, maximising accessibility by automating the whole engineering process may be desirable—Dedicaid is a platform requiring just data, with no customisable parameters, but has an "ethical compass" which flags inappropriate datasets.[49,50]

CONCLUSION

AutoML performance is often comparable to bespoke ML and human performance. Many autoML platforms have been developed in academia and industry, with variable strengths and limitations. AutoML may prove especially useful in pilot studies and education, but potential use cases include primary research and clinical deployment if platforms are rigorously validated.[6] Future autoML research must be more transparently reported, adhere to reporting guidelines and provide appropriate benchmarks for performance comparisons. Further autoML development should seek to minimise the trade-offs currently inherent in selecting any given platform.

Data availability statement

The raw data from this review may be provided upon request.

Disclosure

AJT is supported by The Royal College of Surgeons in Edinburgh (RCSED Bursary 2022), Royal College of Physicians (MSEB 2022), and Corpus Christi College, University of Cambridge (Gordon Award 1083874682). DSWT is supported by the National Medical Research Council, Singapore (NMCR/HSRG/0087/2018; MOH-000655-00; MOH-001014-00), Duke-NUS Medical School (Duke-NUS/RSF/2021/0018; 05/FY2020/EX/15-A58), and Agency for Science, Technology and Research (A20H4g2141; H20C6a0032). These funders were not involved in the conception, execution or reporting of this study. All authors declare no competing interests.

Supplementary materials

Appendix S1. Systematic review search strategy.
Appendix S2. Inclusion and exclusion criteria, as provided to researchers conducting abstract and full-text screening.
Appendix S3. Tabulated study characteristics including citation details and study identifiers used elsewhere in the article.
Appendix S4. Study-level data exhibiting fulfilment of DECIDE-AI reporting standards.
Appendix S5. Study-level data exhibiting appraisal of risk of bias (RoB) and concerns regarding applicability (CrA) using PROBAST.
Appendix S6. Forest plot depicting reported AUPRC metrics.

REFERENCES

1. Pianykh OS, Guitron S, Parke D, et al. Improving healthcare operations management with machine learning. Nat Mach Intell 2020;2:266-73.
2. Park JY, Hsu TC, Hu JR, et al. Predicting Sepsis Mortality in a Population-Based National Database: Machine Learning Approach. J Med Internet Res 2022;24:e29982.
3. Car J, Sheikh A, Wicks P, et al. Beyond the hype of big data and artificial intelligence: building foundations for knowledge and wisdom. BMC Med 2019;17:143.
4. Dash S, Shakyawar SK, Sharma M, et al. Big data in healthcare: management, analysis and future prospects. Journal of Big Data 2019;6:54.
5. Tan TF, Thirunavukarasu AJ, Jin L, et al. Artificial intelligence and digital health in global eye health: opportunities and challenges. Lancet Glob Health 2023;11:e1432-43.
6. Thirunavukarasu AJ, Elangovan K, Gutierrez L, et al. Democratizing Artificial Intelligence Imaging Analysis With Automated Machine Learning: Tutorial. Journal of Medical Internet Research 2023;25:e49949.
7. Khang A, Rana G, Tailor RK, et al. Data-Centric AI Solutions and Emerging Technologies in the Healthcare Ecosystem. 1st Ed. Boca Raton: CRC Press; 2023.
8. Hutter F, Kotthoff L, Vanschoren J (Eds). Automated Machine Learning: Methods, Systems, Challenges. The Springer Series on Challenges in Machine Learning. Cham: Springer International Publishing; 2018.
9. Rawat T, Khemchandani V. Feature Engineering (FE) Tools and Techniques for Better Classification Performance. IJIET 2017;8:169-79.
10. Waring J, Lindvall C, Umeton R. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Artificial Intelligence in Medicine 2020;104:101822.
11. Faes L, Wagner SK, Fu DJ, et al. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. Lancet Digit Health 2019;1:e232-42.
12. Thirunavukarasu A, Elangovan K, Gutierrez L, et al. Comparative analysis of diagnostic imaging models built with automated machine learning. Future Healthc J 20231;10(Suppl 3):21-3.
13. Thirunavukarasu A, Gutierrez L, Elangovan K, et al. The applications of automated machine learning in clinical contexts. PROSPERO 2022 CRD42022344427. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022344427. Accessed 7 January 2024.
14. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
15. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews 2016;5:210.
16. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ 2022;377:e070904.
17. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med 2019;170:51-8.
18. Shen H, Liu T, Cui J, et al. A web-based automated machine learning platform to analyze liquid biopsy data. Lab Chip 2020;20:2166-74.
19. Erickson BJ, Kitamura F. Magician's Corner: 9. Performance Metrics for Machine Learning Models. Radiol Artif Intell 2021;3:e200126.
20. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. Research Synthesis Methods 2020;12:55-61.
21. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. Journal of Open Source Software 2019;4:1686.
22. Dayim A. forestploter. https://github.com/adayim/forestploter. Accessed 3 January 2023.
23. Cho BH, Kaji D, Cheung ZB, et al. Automated Measurement of Lumbar Lordosis on Radiographs Using Machine Learning and Computer Vision. Global Spine J 2020;10:611-8.
24. Adaszewski S, Dukart J, Kherif F, et al. How early can we predict Alzheimer's disease using computational anatomy?. Neurobiol Aging 2013;34:2815-26.
25. Smith R, Julian D, Dubin A. Deep neural networks are effective tools for assessing performance during surgical training. J Robot Surg 2022;16:559-62.
26. Korot E, Pontikos N, Liu X, et al. Predicting sex from retinal fundus photographs using automated deep learning. Sci Rep 2021;11:10286.
27. Alaa AM, van der Schaar M. Prognostication and Risk Factors for Cystic Fibrosis via Automated Machine Learning. Sci Rep 2018;8:11242.
28. Bang CS, Lim H, Jeong HM, et al. Use of Endoscopic Images in the Prediction of Submucosal Invasion of Gastric Neoplasms: Automated Deep Learning Model Development and Usability Study. J Med Internet Res 2021;23:e25167.
29. Korot E, Guan Z, Ferraz D, et al. Code-free deep learning for multi-modality medical image classification. Nature Machine Intelligence 2021;3:288-98.
30. Yang HS, Kim KR, Kim S, et al. Deep Learning Application in Spinal Implant Identification. Spine 2021;46:E318-24.
31. Chen X, Wujek B. A Unified Framework for Automatic Distributed Active Learning. IEEE Trans Pattern Anal Mach Intell 2022;44:9974-86.
32. Abbas A, O'Byrne C, Fu DJ, et al. Evaluating an automated machine learning model that predicts visual acuity outcomes in patients with neovascular age-related macular

degeneration. Graefes Arch Clin Exp Ophthalmol 2022; 260:2461-73.

33. Ying X. An Overview of Overfitting and its Solutions. J Phys: Conf Ser 2019;1168:022022.
34. Cacciamani GE, Chu TN, Sanford DI, et al. PRISMA AI reporting guidelines for systematic reviews and meta-analyses on AI in healthcare. Nat Med 2023;29:14-5.
35. Collins GS, Dhiman P, Navarro CLA, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open 20211;11:e048008.
36. Navarro CLA, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ 2021;375:n2281.
37. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020;368:m689.
38. Thirunavukarasu AJ. How Can the Clinical Aptitude of AI Assistants Be Assayed? Journal of Medical Internet Research 2023;25:e51603.
39. Ng FYC, Thirunavukarasu AJ, Cheng H, et al. Artificial intelligence education: An evidence-based medicine approach for consumers, translators, and developers. CR Med 2023;4:101230.
40. Thirunavukarasu AJ, Ting DSJ, Elangovan K, et al. Large language models in medicine. Nat Med 2023;29:1930-40.
41. Alaa AM, Bolton T, Di Angelantonio E, et al. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLoS One 2019;14:e0213653.
42. Ikemura K, Bellin E, Yagi Y, et al. Using Automated Machine Learning to Predict the Mortality of Patients With COVID-19: Prediction Model Development Study. J Med Internet Res 2021;23:e23458.
43. Luna A, Bernanke J, Kim K, et al. Maturity of gray matter structures and white matter connectomes, and their relationship with psychiatric symptoms in youth. Hum Brain Mapp 2021;42:4568-79.
44. Mohsen F, Biswas MR, Ali H, et al. Customized and Automated Machine Learning-Based Models for Diabetes Type 2 Classification. Stud Health Technol Inform 2022; 295:517-20.
45. Narkhede SM, Luther L, Raugh IM, et al. Machine Learning Identifies Digital Phenotyping Measures Most Relevant to Negative Symptoms in Psychotic Disorders: Implications for Clinical Trials. Schizophr Bull 2022;48:425-36.
46. Sills MR, Ozkaynak M, Jang H. Predicting hospitalization of pediatric asthma patients in emergency departments using machine learning. Int J Med Inf 2021;151:104468.
47. Stojadinovic M, Milicevic B, Jankovic S. Improved predictive performance of prostate biopsy collaborative group risk calculator when based on automated machine learning. Comput Biol Med 2021;138:104903.
48. Yin M, Zhang R, Zhou Z, et al. Automated Machine Learning for the Early Prediction of the Severity of Acute Pancreatitis in Hospitals. Front Cell Infect Microbiol 2022;12:886935.
49. Ritter Z, Papp L, Zambo K, et al. Two-Year Event-Free Survival Prediction in DLBCL Patients Based on In Vivo Radiomics and Clinical Parameters. Front Oncol 2022;12:820136.
50. Hasimbegovic E, Papp L, Grahovac M, et al. A Sneak-Peek into the Physician's Brain: A Retrospective Machine Learning-Driven Investigation of Decision-Making in TAVR versus SAVR for Young High-Risk Patients with Severe Symptomatic Aortic Stenosis. J Pers Med 2021;11:1062.

Ann Acad Med Singap 2024;53:208-10
https://doi.org/10.47102/annals-acadmedsg.2023186

LETTER TO THE EDITOR

Validating two international warfarin pharmacogenetic dosing algorithms for estimating the maintenance dose for patients in Singapore

Dear Editor,

Predicting optimal warfarin dosing is difficult due to complex pharmacodynamics and pharmacokinetics, narrow therapeutic index and susceptibility to many factors.[1] Genetic variations of the CYP2C9 and VKORC1 enzymes, occurring in different frequencies in different populations, play a significant role in determining warfarin dosing.[1-4] Using pharmacogenetic dosing algorithms to predict warfarin doses may shorten the time to achieve target International Normalised Ratio (INR) and stable dose.[2,5] The Clinical Pharmacogenetics Implementation Consortium Guidelines 2017 Update[4] recommends the Gage (WarfarinDosing.org[7]) and International Warfarin Pharmacogenetics Consortium (IWPC)[8] pharmacogenetic algorithms.

Singapore's genetic make-up is diverse due to its multi-ethnic community that consists of Chinese, Malays, Indians and other ethnicities. We present our evaluation of the Gage and IWPC algorithms in estimating warfarin maintenance doses i n Singapore.

Patients receiving warfarin at a large hospital in Singapore's outpatient anticoagulation clinics were prospectively recruited between March 2020 and June 2021. Patients were ≥21 years, have an INR target between 2 and 3, and are stable on therapeutic warfarin doses beyond a month, derived with the hospital's standardised clinical dosing algorithm. We excluded patients with unstable liver or renal disease, end-stage renal failure not on regular dialysis, organ transplant within 2 months, or have received short-term drugs that interact with warfarin.

Consented patients' blood or buccal swab samples were analysed for CYP2C9 (rs1799853), VKORC1 (rs9923231) and CYP4F2 (rs2108622) polymorphisms using Taqman SNP Genotyping Assays on real-time Polymerase Chain Reaction systems.

Patients' demographics, warfarin profile, concomitant interacting drugs, smoking/alcohol history, vitamin K intake, health/herbal supplement use and physical activity were collected. Clinical and genotype data (CYP2C9 [*1, *2 and *3]), VKORC1 c.-1639G>A allele and *CYP4F2*3* were entered into the Gage and IWPC pharmacogenetic dosing algorithms to obtain the predicted warfarin doses. Warfarin doses were expressed as total dosage per week. Descriptive analyses were performed using SPSS Statistics 28.

The proportion of patients whose predicted weekly warfarin maintenance dose differ ±20% of the actual weekly maintenance dose was determined. Bland and Altman analysis was used to study the agreement between the observed and the predicted doses, limits of agreement set as 95% of the data.

Ninety patients were recruited. The actual, Gage-predicted and IWPC-predicted mean warfarin maintenance doses of these patients were 27.2 ± 14.9 mg/week, 28.5 ± 10.9 mg/week and 27.4 ± 10.5 mg/week, respectively. The mean differences between actual and algorithm-predicted doses were -1.3 ± 10.5 mg/week (95% confidence interval [CI] -3.45 to 0.93) with Gage and -0.1 ± 9.7 mg/week (95% CI -2.16 to 1.92) with IWPC, respectively. Though not statistically different, the IWPC performed marginally better, with 45.6% of patients falling within ±20% of their actual doses than the 41.1% with the Gage algorithm, despite not requiring CYP4F2 genetic variation data, a contrast with Chan et al. who proposed that polymorphisms in this gene have a bearing on warfarin dose variation.[6] It is noteworthy that IWPC algorithm included Asian population data.[8,9]

Seven outliers were identified. Five required very high warfarin doses (includes 1 with 28 vegetable servings/week [versus average 11]), and 3 Indians on carbamazepine or azathioprine. Remaining 2 have VKORC1 G/G and at least 1 CYP4F2 *3 variant but required significantly lower warfarin doses. Table 1 shows genotypic variants in different ethnic groups.

Of the 31 Chinese patients, 11 (35.5%) and 15 (48.4%) of them were within ±20% difference of their Gage and IWPC predictions, respectively. The mean dose differences between actual and algorithm-predicted doses were -1.9 ± 7.4 mg/week (95% CI -4.58 to 0.82) for Gage, and -0.1 ± 7.6 mg/week (95% CI -2.82 to 2.72) for IWPC.

Fourteen (50.0%) of the 28 Indian patients' actual warfarin doses were within ±20% difference of their Gage or IWPC predictions. Mean dose differences between actual and algorithm-predicted doses were 1.0 ± 14.8 mg/week (95% CI -4.76 to 6.72) for Gage, and 2.6 ± 12.2 mg/week (95% CI -2.08 to 7.37) for IWPC, respectively.

Table 1. Frequency of the genotypic variants and their respective weekly warfarin maintenance dose requirements in different ethnic groups.

		Chinese (n=31) no. (%)	Malay (n=31) no. (%)	Indian (n=28) no. (%)	Total (n=90) no. (%)
	Warfarin maintenance dose, mean ± SD (mg/week)	20.8 ± 9.5	24.8 ± 12.6	37.1 ± 17.4	27.2 ± 14.9
VKORC1					
-1639 A/A	18.9 ± 6.5	24 (77.4)	16 (51.6)	2 (7.1)	42 (46.7)
-1639 G/A	29.9 ± 12.5	7 (22.6)	13 (41.9)	8 (28.6)	28 (31.1)
-1639 G/G	41.1 ± 19.2	0	2 (6.5)	18 (64.3)	20 (22.2)
CYP2C9					
*1/*1	27.4 ± 15.6	31 (100)	28 (90.3)	17 (60.7)	76 (84.4)
*1/*2	-	0	0	1 (3.6)	1 (1.1)
*1/*3	25.3 ± 11.1	0	3 (9.7)	9 (32.1)	12 (13.3)
*2/*2	-	0	0	0	0
*2/*3	-	0	0	1 (3.6)	1 (1.1)
*3/*3	-	0	0	0	0
CYP4F2					
*1/*1 (V433M CC)	24.9 ± 12.6	14 (45.2)	17 (54.8)	8 (28.6)	39 (43.3)
*1/*3 (V433M CT)	29.4 ± 17.5	14 (45.2)	14 (45.2)	15 (53.6)	43 (47.7)
*3/*3 (V433M TT)	26.9 ± 8.55	3 (9.7)	0	5 (17.9)	8 (8.9)

Chinese patients require the lowest maintenance dose, with 100% of the group possessing the CYP2C9*1/*1 allele, and 77.4% the VKORC1 -1639 A/A allele. Indian patients show heterogeneity in the CYP2C9 and VKORC1 alleles, predicating the largest spread of warfarin doses and have the highest maintenance doses (64.3% have the VKORC1 -1639 G/G allele).
SD: standard deviation

Twelve (38.7%) of the 31 Malay patients' actual warfarin doses were within ±20% difference of their Gage or IWPC predictions. Mean dose differences between actual and predicted were -2.7 ± 8.1 mg/week (95% CI -5.64 to 0.32) for Gage, -2.7 ± 8.7 mg/week (95% CI -5.88 to 0.53) for IWPC, respectively.

Dose groups were categorised into low-dose (<21 mg/week), moderate-dose (21–34.5 mg/week) and high-dose (≥35 mg/week), based on our clinical experience.

Among the 37 patients in the low-dose group, 7 (18.9%) and 12 (32.4%) actual doses were within ±20% difference of their Gage and IWPC predictions, respectively. Mean actual doses were significantly different from Gage-predicted -6.3 ± 4.5 mg/week (95% CI -7.84 to -4.86) and IWPC-predicted -5.5 ± 5.1 mg/week (95% CI -7.15 to -3.77) doses. The algorithms overestimated the doses by ≥5–6 mg/week.

Among the 23 patients in the high-dose group, 10 (43.5%) and 9 (39.1%) actual doses were within ±20% difference of their Gage- and IWPC-predicted, respectively. Mean actual doses were significantly different when compared with both Gage-predicted 8.7 ± 12.0 mg/week (95% CI 3.52–13.88) and IWPC-predicted 9.7 ± 10.2 mg/week (95% CI 5.28–14.07) doses. The dosing algorithms tend to underestimate doses by a mean of 8–9 mg/week, with 8 patients having high warfarin doses of ≥7 mg/day. A systematic review also showed that 22 dosing algorithms (including Gage) tend to underestimate when doses exceed 7 mg/day.[10]

For the 30 patients in the moderate dose group, 20 (66.7%) of them were within ±20% difference compared with their predicted doses. The mean differences were -2.6 ± 9.4 mg/week (95% CI -6.13 to 0.90) when compared with Gage-predicted, and -1.0 ± 8.2 mg/week (95% CI -4.10 to 2.03) with IWPC-predicted doses.

The pharmacogenetic dosing algorithms were originally developed to predict warfarin initiation doses. However, as we cannot take into account the full variability determining the initiation doses, we compared the predicted doses to maintenance doses and allowed for 20% difference.[8] Patients with unstable renal and liver conditions (and inherently higher bleeding risk) were excluded from the study, hence a more conservative approach is needed if these dosing algorithms are used.

Overall, there is no significant difference between the Gage- or IWPC-predicted doses compared with the observed maintenance doses.

Integrating pharmacogenetic algorithms into local clinical practice is promising, as it can potentially enable faster and safer attainment of therapeutic anticoagulation. Further studies can be done to identify patient groups who will benefit most from genetic testing.

Acknowledgements

We wish to express our sincere thanks to the Tan Tock Seng Hospital Anticoagulation Pharmacist team, and our research coordinators Ms Felicia Lee Si Min and Ms Stephanie Wong Li Ling from Clinical Research & Innovation Office (CRIO), for their support and guidance in the study process and patient recruitment.

Funding and ethics

This study was approved by the NHG Domain Specific Review Board (2019/01166) and supported by the Personalised Medicine Translational Research Seed Funding Programme (PMTRSFP19-03).

Keywords: *algorithm, pharmacogenetics, pharmacogenomics, Singapore, warfarin*

REFERENCES

1. Kaye JB, Schultz LE, Steiner HE, et al. Warfarin Pharmacogenomics in Diverse Populations. Pharmacotherapy 2017;37:1150-63.
2. Pavani A, Naushad SM, Rupasree Y, et al. Optimization of warfarin dose by population-specific pharmacogenomic algorithm. Pharmacogenomics J 2012;12:306-11.
3. Tham LS, Goh BC, Nafziger A, et al. A warfarin-dosing model in Asians that uses single-nucleotide polymorphisms in vitamin K epoxide reductase complex and cytochrome P450 2C9. Clin Pharmacol Ther 2006;80:346-55.
4. Johnson JA, Caudle KE, Gong L, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. Clin Pharmacol Ther 2017;102:397-404.
5. Goh LL, Lim CW, Sim WC, et al. Analysis of Genetic Variation in CYP450 Genes for Clinical Implementation. PLOS One 2017;12:e0169233.
6. Chan SL, Suo C, Lee SC, et al. Translational aspects of genetic factors in the prediction of drug response variability: a case study of warfarin pharmacogenomics in a multi-ethnic cohort from Asia. Pharmacogenomics J 2012;12:312-8.
7. Gage BF, Eby C, Johnson JA, et al. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clin Pharmacol Ther 2008;84:326-31.
8. International Warfarin Pharmacogenetics Consortium; Klein TE, Altman RB, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med 2009;360:753-64.
9. Sasano M, Ohno M, Fukuda Y, et al. Verification of pharmacogenomics-based algorithms to predict warfarin maintenance dose using registered data of Japanese patients. Eur J Clin Pharmacol 2019;75:901-11.
10. Saffian SM, Duffull SB, Wright D. Warfarin Dosing Algorithms Underpredict Dose Requirements in Patients Requiring ≥7 mg Daily: A Systematic Review and Meta-analysis. Clin Pharmacol Ther 2017;102:297-304.

Stephanie Pei Yun Soh[1] *MSc(Pharm)*,
Wei Yann See Toh[1] *PharmD*,
Wei Qing Ten[1] *BSc(Pharm)(Hons)*,
Khai Pang Leong[2,3] *FAMS*, Liuh Ling Goh[3] *PhD*

[1] Department of Pharmacy, Tan Tock Seng Hospital, Singapore
[2] Department of Rheumatology, Allergy and Immunology, Tan Tock Seng Hospital, Singapore
[3] Molecular Diagnostic Laboratory, Personalised Medicine Service, Tan Tock Seng Hospital, Singapore

Correspondence: Miss Soh Pei Yun Stephanie, Department of Pharmacy, Tan Tock Seng Hospital
11 Jalan Tan Tock Seng, Singapore 308433.
Email: Stephanie_py_soh@ttsh.com.sg

Ann Acad Med Singap 2024;53:211-2
https://doi.org/10.47102/annals-acadmedsg.2023357

LETTER TO THE EDITOR

HLA-B*5801 testing: Is it time to consider mandatory testing prior to prescribing allopurinol in Singapore?

Dear Editor,

Stevens-Johnsons Syndrome (SJS) and Toxic Epidermal Necrolysis (TEN) are severe, life-threatening mucocutaneous reactions that most commonly occur as drug-related reactions.[1] In recent years, several risk factors for the development of SJS/TEN, such as genetic factors, have been identified. Notably, carriers of the HLA-B*5801 and HLA-B*1502 alleles have an increased risk of SJS/TEN with the use of allopurinol and carbamazepine, respectively.[2] Consequently, much debate has arisen over the utility of pharmacogenomics in preventing SJS/TEN. We will discuss the evidence surrounding genetic testing in the prevention of allopurinol-induced SJS/TEN, with the aim of highlighting the potential value of pre-testing.

In Singapore, the number of cases of allopurinol-induced severe cutaneous adverse reactions (SCAR) remains high. A total of 131 cases of allopurinol-related SCAR were reported across the past decade (2013–2022), of which 63 were SJS/TEN cases. These incidents were in contrast to carbamazepine-related SCAR, which had 31 cases in the same time period (2013–2022), of which 19 were SJS/TEN. A clear decline in the number of carbamazepine-related SJS/TEN cases was seen since the mandatory HLA-B*1502 testing prior to carbamazepine initiation was introduced in 2013. Currently, the HLA-B*5801 testing prior to allopurinol initiation is not mandatory in Singapore. Reasons cited for this decision include the low positive predictive value (1.52%) compared to the high prevalence of the allele in our population (18.5%) and the lack of cost-effectiveness of testing.[3] Others have argued that although the HLA allele plays a central role in allopurinol-related SJS/TEN, it may not be the only factor required for its occurrence.[4]

Nonetheless, we believe there is value in HLA pre-testing. First, there is a need to relook at the data on the cost-effectiveness of testing. The current data cite the cost of testing and the higher cost of alternative urate-lowering therapies as the main drivers of the increased cost of universal HLA-B*5801 testing. However, most cost-effectiveness models are unable to account for the potentially significant morbidity and long-term multi-systemic sequelae of SJS/TEN. While local studies have attempted to examine long-term costs of SJS/TEN, these were limited to ocular complications of dry eye syndrome in both studies.[3,5] Statistics used in these studies were also based on incidence data in other countries, which may not be an accurate representation of the Singapore context. These are important considerations as such complications can lead to a substantial increase in costs. For example, a study on healthcare costs of the disease found that SJS and TEN resulted in significantly prolonged lengths of stay at approximately 9.8 days and 16.5 days, respectively, as well as higher costs of care at approximately USD21,437 and USD53,695 (approx. SGD28,718 and SGD71,932), respectively, in comparison to all other admissions (average length of stay at 4.7 days and cost at USD11,281 [approx. SGD15,113]).[6] Chiu et al. also found that patients with SJS/TEN tend to suffer a substantial loss of life expectancy and high lifetime healthcare expenditure compared to the general population.[7]

Second, cost-effectiveness studies are based on the assumption of appropriate prescription of allopurinol for the treatment of gout. However, inappropriate use of allopurinol, such as for infrequent gout attacks and asymptomatic hyperuricemia, remains prevalent in Singapore.[8] This is a universal issue where poor compliance to these standards of care are commonly seen in other published studies as well.[9] Although potentially remediable, allopurinol continues to be prescribed inappropriately despite efforts from the Health Sciences Authority, who have released numerous newsletters and guidelines on allopurinol prescription. This calls for a stronger safeguard in place to help reduce prescription of allopurinol, especially in cases where it is not indicated.

In addition, there is promising evidence on the utility of HLA testing. A study by Ko et al. demonstrated a reduction in the incidence of allopurinol-induced SCAR with the implementation of pre-testing in Taiwan, a country with a high incidence of the allele (reported to be up to 20%). The study results suggested that HLA-B*5801 pre-testing could potentially prevent an estimated number of 330 cases of allopurinol-induced SCAR per year.[10] Similarly, studies in Korea, China and Thailand also demonstrated a reduced incidence of SCAR with pre-testing in patients with chronic

kidney disease (CKD).[1,2,4] Furthermore, these studies also demonstrated the cost-effectiveness of pre-testing in such patients. For example, Wong et al. found that pre-testing in patients with renal impairment could potentially reduce healthcare costs by up to 67.1%, where the total cost of hospitalisation of SCAR patients amounted up to 3 times higher than performing HLA pre-testing itself.[4] Park et al. also found that treatment informed by HLA-typing in CKD patients cost USD138 (approx. SGD185) less than the conventional treatment and was also associated with a higher likelihood of continued gout treatment without SCARs.[2] Given that this patient group has a 5-fold risk of developing SCARs, effective preventative measures are of paramount importance. Hence, as a first step towards universal HLA pre-testing, it may be prudent to consider implementing testing in selected patient groups at higher risk of developing SCARs, such as CKD patients.

Although HLA-B*5801 screening will limit the use of allopurinol as a first-line medication in patients with the allele, several alternatives with a lower SCAR risk profile remain available, such as probenecid, benzbromarone and febuxostat. While febuxostat had been previously plagued with concerns over cardiovascular safety, the Febuxostat versus Allopurinol Streamlined Trial (FAST) published in 2020 demonstrated no increased risk of cardiovascular events or mortality compared to allopurinol. Greater awareness on the use of these alternative therapeutics will increase clinicians' confidence in HLA pre-testing.

In conclusion, allopurinol remains an important and potentially preventable cause of SCAR. Given the significant morbidity and potential mortality of allopurinol-related SCAR, it may be time to re-examine current data and reconsider the amendment of recommendations on HLA-B*5801 testing. Genetic testing before allopurinol initiation may be able to potentially act as a safeguard and prevent unnecessary prescription. On the other hand, future studies could consider exploring the development of guidelines regarding treatment alternatives for patients at high risk of allopurinol-related SCAR, and to personalise the treatment of gout based on test results and patient factors.

Declarations

The authors do not have any affiliations or financial involvement with any commercial organisations with a direct financial interest in the subject or materials discussed in the manuscript.

Acknowledgement

The authors would like to acknowledge Ms Celine Loke (Health Sciences Authority, Singapore) for providing statistics on cases of severe cutaneous adverse reactions for this manuscript.

Keywords: *dermatology, drug reactions, internal medicine, pharmacogenomics, public health*

REFERENCES

1. Saokaew S, Tassaneeyakul W, Maenthaisong R, et al. Cost-effectiveness analysis of HLA-B*5801 testing in preventing allopurinol-induced SJS/TEN in Thai population. PLoS One 2014;9:e94294.
2. Park DJ, Kang JH, Lee JW, et al. Cost-effectiveness analysis of HLA-B5801 genotyping in the treatment of gout patients with chronic renal insufficiency in Korea. Arthritis Care Res (Hoboken) 2015;67:280-7.
3. Dong D, Tan-Koi WC, Teng GG, et al. Cost-effectiveness analysis of genotyping for HLA-B*5801 and an enhanced safety program in gout patients starting allopurinol in Singapore. Pharmacogenomics 2015;16:1781-93.
4. Wong CSM, Yeung CK, Chan CY, et al. HLA-B*58:01 screening to prevent allopurinol-induced severe cutaneous adverse reactions in Chinese patients with chronic kidney disease. Arch Dermatol Res 2022;314:651-9.
5. Pruis SL, Jeon YK, Pearce F, et al. Cost-effectiveness of sequential urate lowering therapies for the management of gout in Singapore. J Med Econ 2020;23:838-47.
6. Hsu DY, Brieva J, Silverberg NB, et al. Morbidity and Mortality of Stevens-Johnson Syndrome and Toxic Epidermal Necrolysis in United States Adults. J Invest Dermatol 2016;136:1387-97.
7. Chiu YM, Chiu HY. Lifetime risk, life expectancy, loss-of-life expectancy, and lifetime healthcare expenditure for Stevens-Johnson syndrome/toxic epidermal necrolysis in Taiwan: follow-up of a nationwide cohort from 2008 to 2019. Br J Dermatol 2023;189:553-60.
8. Lee HY, Tay LK, Thirumoorthy T, et al. Cutaneous adverse drug reactions in hospitalised patients. Singapore Med J 2010;51:767-74.
9. Pasina L, Brucato AL, Djade CD, et al. Inappropriate prescription of allopurinol and febuxostat and risk of adverse events in the elderly: results from the REPOSI registry. Eur J Clin Pharmacol 2014;70:1495-503.
10. Ko TM, Tsai CY, Chen SY, et al. Use of HLA-B*58:01 genotyping to prevent allopurinol induced severe cutaneous adverse reactions in Taiwan: national prospective cohort study. BMJ 2015;351:h4848.

Felicia Li Ling Ong[1] *MBBS*,
Shiu Ming Pang[1,2] *MBBS*

[1] Department of Dermatology, Singapore General Hospital
[2] Duke-NUS Medical School

Correspondence: Clin. A/Prof Shiu Ming Pang, 20 College Road, Department of Dermatology, Level 4, Academia, Singapore General Hospital, Singapore 169856.
Email: pang.shiu.ming@singhealth.com.sg

Ann Acad Med Singap 2024;53:213-5
https://doi.org/10.47102/annals-acadmedsg.2023306

LETTER TO THE EDITOR

Patient's degree of adherence, challenges & preferences towards medicine taking (PACT) in Singapore

Dear Editor,

Medication nonadherence is a prevalent public health problem that compromises patients' health outcomes and increases healthcare expenditures.[1] Studies in Singapore showed that 25.7%–38.9% of patients are nonadherent.[2,3] Studies investigating the association between patients' reasons for nonadherence and their preferences towards adherence enablers are limited. We aimed to (1) examine the prevalence and reasons of medication nonadherence among patients with different clinical conditions and settings and (2) investigate possible associations with their preferred intervention for improving adherence.

A multicentre, cross-sectional and self-administered survey was conducted at 6 primary, intermediate and tertiary healthcare institutions in Singapore from March 2022 to November 2022. Participants were identified via convenience sampling at their routine care encounters (e.g. if they were aged 21 years and above, on any long-term medication and could understand the survey language [i.e. English, Chinese or Malay]). Eligible participants completed an anonymous online survey and received remuneration. The survey consisted of participant's profile, Domains of Subjective Extent of Nonadherence tool to examine the extent and reasons of nonadherence, and preferred interventions.[3] Open-ended survey responses were analysed thematically. Descriptive results were summarised by mean ± standard deviation or number (percentage). Multiple logistic regression and tetrachoric correlations were used to statistically test for associations between categorical variables. This study was approved by the National Healthcare Group Domain Specific Review Board (2021/01168).

Among 6646 potential participants approached, 1248 (19.0%) completed the survey. Participants were mostly from the ambulatory care settings—specialist outpatient centres (SOC) 553 (44.0%), polyclinics 497 (40.0%) and inpatients 198 (16.0%)—who responded unassisted (83.0%) in English (89.2%). The mean age was 54.3 ± 16.3 years and 49.8% were female. Most participants were Chinese (77.7%), had at least secondary school education (80.8%), married/partnered (63.2%), working full-time (41.9%) and were not trained healthcare workers (92.6%). Participants followed up with 1.6 ± 1.8 healthcare institutions, had 2.1 ± 1.3 chronic medical conditions, and received 3.6 ± 3.0 long-term medications and supplements. Common medical conditions were hypertension (50.2%), hyperlipidaemia (50.0%), diabetes (26.7%), depression (9.0%) and anxiety (8.8%). Majority (91.0%) of participants self-managed their medications. Moreover, 704 (56.4%) were nonadherent and this was comparable across the different settings: SOC (59.7%), polyclinics (51.3%) and inpatient (60.1%). Age (adjusted odds ratio [AOR] 0.98, 95% confidence interval [CI] 0.97–0.99, P=0.006), homemaker compared to full-time worker (AOR 0.57, 95% CI 0.36–0.91, P=0.018), number of medications/supplements (AOR 1.08, 95% CI 1.01–1.15, P=0.025) and anxiety (AOR 1.74, 95% CI 1.02–2.96, P=0.041) were independently associated with nonadherence. More than a third of the participants cited the following reasons: "I forgot" (75.6%), "I was out of my routine" (61.5%), "I was too late with my dose" (51.7%), "I did not have my medicines with me" (49.7%), "I was asleep" (49.7%), "The medication caused side effects" (40.9%), "I could not meet the food requirements" (36.2%), and "I was afraid the medications would interact with other medications I take" (35.9%). For open-ended responses, majority of the themes included forgetfulness, competing commitments, medication administration requirements with respect to food, polypharmacy and personal experiences with medication side effects. Notably, some participants expressed the lack of motivation. Most participants felt that pillboxes (43.1%), learning how to maintain their medication list (45.2%) and having access to medication information (42.6%) would improve their medication adherence, while 35.7% felt that one-on-one sessions with a healthcare professional would help. Open-ended responses listed pillboxes and various reminders (e.g. alarm, phone calls and family) as their top choices. Some participants recognised that they need more self-discipline and that they would not miss their medications if they view it as important. Participants, who forgot their medication, preferred mobile application reminders (r=0.158, P=0.026) and alarms (r=0.286, P<0.001). Similarly, participants, who were out of their routine, felt that mobile application reminders (r=0.133, P=0.040) and alarms (r=0.134, P=0.047) would be helpful. Participants who missed their medications due to being asleep were interested in discussing with healthcare providers on their regimen (r=0.215, P=0.001) (Table 1).

The study's prevalence of medication nonadherence (56.4%) is higher than the 38.9% reported

Table 1. Tetrachoric correlation coefficient (r) of main reasons for non-adherence and preferred interventions.

Reasons for nonadherence	I forgot	I was out of my routine	I was too late with my dose	I did not have my medicines with me	I was asleep
Preferred interventions					
Tools					
Pillbox	0.006	0.056	-0.015	-0.051	-0.042
Mobile app reminders	0.158[c]	0.133[c]	0.126[c]	0.188[b]	0.084
Specialised packaging	0.099	0.087	0.017	0.044	0.135[c]
Medicine labels	-0.029	-0.129	-0.030	0.018	0.083
Alarm	0.286[a]	0.134[c]	0.156[c]	0.103	0.160[c]
Services					
Learning how to maintain my medicine list	-0.046	-0.018	0.062	0.022	-0.018
Having access to medicine information	0.065	0.038	0.019	0.058	0.060
Getting my questions answered through interactive channels	-0.021	-0.061	0.071	0.038	0.109
Deciding on my medicines with my healthcare providers	-0.053	0.028	0.066	0.066	0.215[b]

[a] $P<0.001$, [b] $P<0.01$, [c] $P<0.05$

in another study using the same adherence measure in diabetic patients in Singapore.[3] Younger patients and patients with more medications/supplements tend to be more nonadherent. This could be due to multiple competing commitments as revealed in participants' open-ended responses and increased daily doses or complex dosing regimen. An overview of systematic reviews proposed that age has a concave relationship with adherence, where poor adherence was observed in the very young and very old.[4] Those with more medications/supplements tend to be more nonadherent, which highlights the importance of medication regimen simplification.

Anxiety was a significant predictor of nonadherence, and our study suggests more support to be given to these patients. Patients with more negative emotions could be less motivated to care for themselves and adhere to complex medication regimens. However, this association varied in literature.[5]

Forgetfulness was a major reason for nonadherence. In our study, participants also shared the lack of motivation and self-discipline. This suggests that forgetfulness may have both a cognitive and motivational component.[6] Medication adherence is a complex behaviour; hence, further studies should systematically explore beliefs, motivation and other variables affecting adherence.[7] Although participants who forgot their medications preferred interventions such as mobile application reminders and alarms, these reminders only target the cognitive aspect of forgetfulness and may not be effective alone.[6] Pillboxes, learning to maintain medication list and having access to medication information were well-perceived by participants and may be useful in a multimodal intervention strategy. One-third of the participants responded that having one-on-one sessions with healthcare professionals will be useful to improve their medication adherence. One-on-one medication therapy management by pharmacists has shown effectiveness in improving medication adherence and resolving drug-related problems.[8]

The online survey could have excluded the less digitally-savvy participants, even though Singapore has a high mobile internet user penetration rate of 93.7%.[9] Future studies should match specific interventions to participants' specific reasons for nonadherence for more targeted investigations.[10] Validation of a medication adherence intervention questionnaire would also add robustness to future work in this field.

Our study reported a significant burden of medication nonadherence in Singapore. The common factors of nonadherence and patients'

preference on possible interventions will provide greater insights in planning programmes to address medication nonadherence among Singapore's ageing population and increasing multimorbidity patient population.

Acknowledgements

We would like to thank Dr Corrine Voils for the permission to use the Domains of Subjective Extent of Nonadherence (DOSE-Nonadherence) Measure. We would also like to thank and acknowledge all pharmacists who have assisted us in the administration of the survey.

Funding

This study was supported by Pharmacy Transformation Office, National Healthcare Group.

Keywords: *DOSE-nonadherence, epidemiology, medication adherence, medication management, PACT*

REFERENCES

1. Benjamin RM. Medication Adherence: Helping Patients Take Their Medicines As Directed. Public Health Rep 2012;127:2-3.
2. Lum ZK, Suministrado MSP, Venketasubramanian N, et al. Medication compliance in Singaporean patients with Alzheimer's disease. Singapore Med J 2019;60:154-60.
3. Liau YW, Cheow C, Leung KTY, et al. A cultural adaptation and validation study of a self-report measure of the extent of and reasons for medication nonadherence among patients with diabetes in Singapore. Patient Prefer Adherence 2019; 13:1241-52.
4. Gast A, Mathes T. Medication adherence influencing factors—an (updated) overview of systematic reviews. Syst Rev 2019;8:112.
5. DiMatteo MR, Lepper HS, Croghan TW. Depression Is a Risk Factor for Noncompliance With Medical Treatment: Meta-analysis of the Effects of Anxiety and Depression on Patient Adherence. Arch Intern Med 2000;160:2101-7.
6. Stone VE. Strategies for optimizing adherence to highly active antiretroviral therapy: lessons from research and clinical practice. Clin Infect Dis 2001;33:865-72.
7. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42.
8. Ai AL, Carretta H, Beitsch LM, et al. Medication therapy management programs: promises and pitfalls. J Manag Care Spec Pharm 2014;20:1162-82.
9. Statista. Mobile internet user penetration rate in selected regions in 2022. Published 2023. https://www.statista.com/statistics/239114/global-mobile-internet-penetration/. Accessed 16 May 2023.
10. Allemann SS, Nieuwlaat R, van den Bemt BJF, et al. Matching adherence interventions to patient determinants using the Theoretical Domains Framework. Front Pharmacol 2016;7:429.

Vivien Teo[1] *MSc (Clin Pharm)*,
Benny Kaihui Li[2] *MSc (Healthcare Mgmt)*,
Suat Fern Low[3] *BSc Pharmacy*,
Jing Wen Oh[4] *BSc Pharmacy (Hons)*,
Kwok Foong Lam[5] *BSc Pharmacy (Hons)*,
Eng Sing Lee[6,7] *PhD*,
Cheryl Wei Yan Tan[8] *PharmD*,
Poh Ching Tan[8] *MSc (Clin Pharm)*,
Hui Shan Chia[8] *MSc (Clin Pharm)*,
Su Yin Pang[9] *BSc Pharmacy (Hons)*,
Fu Xun Marcus Heng[9] *BSc Pharmacy (Hons)*,
Sin Hui Neo[10] *MSc (Epidemiology & Biostats)*,
Tat Ming Ng[1] *PharmD*

[1] Pharmacy, Tan Tock Seng Hospital, Singapore
[2] Pharmacy, National Skin Centre, Singapore
[3] Pharmacy, Yishun Community Hospital, Singapore
[4] Pharmacy, Khoo Teck Puat Hospital, Singapore
[5] Pharmacy, Institute of Mental Health, Singapore
[6] Clinical Research Unit, National Healthcare Group Polyclinics, Singapore
[7] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
[8] National Healthcare Group Pharmacy, Singapore
[9] Group Pharmacy Office, National Healthcare Group, Singapore
[10] Research, National Skin Centre, Singapore

Correspondence: Mr Heng Fu Xun Marcus,
3 Fusionopolis Link #03-08, Nexus@one-north (South Lobby),
Singapore 138543.
Email: Marcus_fx_heng@nhg.com.sg

ANNALS, ACADEMY OF MEDICINE, SINGAPORE
81 Kim Keat Road, #11-00 & #12-00 NKF Centre, Singapore 328836
Email: annals@ams.edu.sg | Website: https://www.annals.edu.sg