OSONG PUBLIC HEALTH
AND RESEARCH PERSPECTIVES
https://ophrp.org

phrp
OSONG PUBLIC HEALTH
AND RESEARCH PERSPECTIVES

Vol. 16, No. 3, June 2025

Pages 193-307

phrp
OSONG PUBLIC HEALTH
AND RESEARCH PERSPECTIVES

Vol. 16, No. 3, June 2025

pISSN 2210-9099
eISSN 2233-6052

Vol. 16, No. 3, June 2025

Korea Disease Control and
Prevention Agency

https://ophrp.org

OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

https://ophrp.org

pISSN 2210-9099
eISSN 2233-6052

Vol. 16, No. 3, June 2025

Published on 30 June 2025

Aims & Scope

Osong Public Health and Research Perspectives (PHRP) is the international bimonthly (published at the end of February, April, June, August, October, and December) journal founded in 2010 by the Korea Disease Control and Prevention Agency (KDCA). With the mission of the KDCA, to create a disease-free world, PHRP encourages sharing medical information and knowledge in the areas of public health.
PHRP publishes original articles, review articles, guidelines, data profiles (including cohort profiles), special articles, short communications, viewpoints, editorials, commentaries, and correspondence, and book reviews, with a focus on the following areas of expertise: emerging infectious diseases, vaccinology, zoonotic diseases, non-communicable diseases, intractable and rare diseases, and human genomics.

Abstracting and Indexing Services

A part of articles, metadata, or full text is available from CrossRef metadata (2010–), PubMed Central (2010–), PubMed (2010–), Scopus (2010–), Embase (2010–), EBSCO Host (2010–), Directory of Open Access Journals (2018–), Korea Citation Index (2011–), and Emerging Sources Citation Index (2020–).

Open Access

PHRP is an open access journal. All contents of the journal are available immediately upon publication without embargo period (https://ophrp.org). All articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher: The Korea Disease Control and Prevention Agency

Editor-in-Chief: Jong-Koo Lee, MD, PhD

Editorial Office:

- **Korea Disease Control and Prevention Agency**
  National Center for Medical Information and Knowledge, 202 Osongsengmyung 2nd street, Osong-eup, Heungdeok-gu, Cheongju 28159, Republic of Korea. E-mail: ophrp@korea.kr

- **Infolumi**
  #245, 10 Jangmi-ro 48beon-gil, Bundang-gu, Seongnam 13496, Republic of Korea
  Tel: +82-31-709-0918, E-mail: osong.editorialoffice@gmail.com

© 2025 Korea Disease Control and Prevention Agency
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
♾ This paper meets the requirements of KS X ISO 9706, ISO 9706-1994 & ANSI/NISO Z39.48-1992 (Permanence of Paper).

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

https://ophrp.org
pISSN 2210-9099 eISSN 2233-6052

Editorial Board

Editor-in-Chief

Jong-Koo Lee, MD, PhD
National Academy of Medicine of Korea, Seoul, Korea

Editors

Curie Ahn, MD, PhD National Medical Center, Korea
Shilpa Buch, PhD University of Nebraska, USA
Young June Choe, MD, PhD Korea University, Korea
Thomas Grischott, MD University of Zurich, Switzerland
Sun Huh, MD, PhD Hallym University, Korea
Sun Ha Jee, PhD Yonsei University, Korea
Yong Seok Jeong, PhD Kyung Hee University, Korea
Joseph Kamgno, MD, PhD CRFilMT, Cameroon
Jae-Heon Kang, MD, PhD Sungkyunkwan University, Korea
Dong-Hyun Kim, MD, PhD Hallym University, Korea
Dong-Su Kim, PhD Ajou University, Korea
Jinhyun Kim, PhD Seoul National University, Korea
Kyehun Kim, MD, PhD Chonnam National University, Korea
Yongkuk Kim, PhD Kyungpook National University, Korea
Moo-Song Lee, MD, PhD University of Ulsan, Korea
Martin Meltzer, PhD CDC, USA
Hiroshi Nishiura, MD, PhD Kyoto University, Japan
Hyun Joo Pai, MD, PhD Hanyang University, Korea
Kholoud Porter, PhD Medical Research Council, UK
Tomoya Saito, MD, PhD National Institute of Infectious Diseases, Japan
Gabriel Sanfélix-Gimeno, PharmD Foundation for the Promotion of Health and Biomedical Research of Valencia Region (FISABIO), Spain
James Sneyd, PhD The University of Auckland, New Zealand
Aeree Sohn, PhD Sahmyook University, Korea
Byoung-Joon Song, PhD National Institutes of Health, USA
Kiyosu Taniguchi, MD, PhD National Institute of Infectious Diseases, Japan

Statistical Editors

Eunok Jung, PhD Konkuk University, Korea
Ho Kim, PhD Seoul National University, Korea
Nam-Kyoo Lim, PhD Korea Disease Control and Prevention Agency, Korea

Journal Management Team

Director

Hee Youl Chai, PhD Korea Disease Control and Prevention Agency, Korea

Managing Editors

Se Hee Hwang, MS Korea Disease Control and Prevention Agency, Korea
Chaeshin Chu, PhD Korea Disease Control and Prevention Agency, Korea

Editorial Coordinator

Hye-Min Cho, MA Infolumi, Korea
Jisoo Yoon, MS Infolumi, Korea

Manuscript Editors

Sue Yeon Chung Infolumi, Korea
Yoon Joo Seo, ELS Infolumi, Korea

© 2025 Korea Disease Control and Prevention Agency

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

https://ophrp.org
pISSN 2210-9099 eISSN 2233-6052

Vol. 16, No. 3, June 2025

Editorial

193 The Republic of Korea's health system at a turning point: from infectious disease threats to comprehensive reform
Jong-Koo Lee

Review Articles

195 Effectiveness of non-pharmacological school-based therapies for cigarette smoking cessation among adolescents in South and Southeast Asian countries: a systematic review and meta-analysis
Fahad Ali Mangrio, Penpaktr Uthis, Suwimon Rojnawee, Alicia K Matthews

211 The role of vitamin D and the *VDR* gene polymorphism in sepsis risk and mortality: a systematic review and meta-analysis
Wiwi Jaya, Sumarno Reto Prawiro, Hani Susianti, Aswoco Andyk Asmoro

223 Effect of cognitive behavior therapy on abstinence among healthy smokers: a systematic review and meta-analysis of randomized controlled trials
Rajesh Kumar, Priyanka Malhotra, Vishwas AS, Kalpana Beniwal, Gopal Singh Charan

Original Articles

236 Design and evaluation of a multi-epitope subunit vaccine against human norovirus using an immunoinformatics approach
Zxcy L. Nonog, Edward C. Banico, Ma. Easter V. Sajo, Pablo V. Serrano, Fredmoore L. Orosco

252 Metabolic and systemic inflammation status in rheumatoid arthritis—fasting blood glucose as a primary predictor of rheumatoid arthritis risk: a cross-sectional study in Iran
Sorayya Kheirouri, Mohammad Alizadeh, Arash Tandorost

261 The level of household food insecurity is associated with the risk of infectious diseases among toddlers in Indonesia: a cross-sectional study
Siti Masitoh, Tri Wurisastuti, Woro Riyadina, Sudarto Ronoatmodjo

270 Analysis of influenza-like illness trends in Saudi Arabia: a comparative study of statistical and deep learning techniques
Fathelrhman EL Guma

Short Communications

285 Subacute sclerosing panencephalitis incidence following measles elimination efforts in the Republic of Korea
Young Hwa Lee, Jung Hye Byeon, Cho Ryok Kang, Young June Choe, Jong-Koo Lee

292 First report of human *Mycobacterium bovis* infection in a veterinary laboratory worker in the Republic of Korea
Ji-Yeon Lee, Seong Wook Pyo, Jieun Kim, Young-Joon Park

Brief Report

300 Standardizing the approach to clinical-based human microbiome research: from clinical information collection to microbiome profiling and human resource utilization
Jung Wook Kim, Eun Chae Choi, Kwang Jun Lee

© 2025 Korea Disease Control and Prevention Agency

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0235
Osong Public Health Res Perspect 2025;16(3):193-194

Editorial

The Republic of Korea's health system at a turning point: from infectious disease threats to comprehensive reform

Jong-Koo Lee

National Academy of Medicine of Korea, Seoul, Republic of Korea

This issue carries special significance in light of the 34th Technical Advisory Group on Immunization meeting of the World Health Organization Western Pacific Regional Office. Recent resurgences of measles and polio in the region have prompted renewed discussions on response strategies, including supplementary immunization activities. The Republic of Korea, which was certified as polio-free in 2000 and has adopted the inactivated poliovirus vaccine, now requires sustained surveillance and close cooperation with neighboring countries to guard against both wild poliovirus and vaccine-derived poliovirus. Similarly, following a major measles outbreak in 2000–2001, the Republic of Korea declared measles elimination in 2006. To maintain this status, the Republic of Korea has strengthened its outbreak response, strengthened contact tracing and antibody monitoring, implemented 2-dose vaccination and recall/reminder systems, and eliminated out-of-pocket costs for essential vaccines to reinforce universal health coverage. This issue features a study using insurance data to evaluate cases of subacute sclerosing panencephalitis, a severe measles complication projected to result from the 2000–2001 epidemic. The findings underscore the global urgency of measles eradication [1].

A new government was inaugurated on June 3. The new administration faces a healthcare crisis triggered by the previous government's decision to expand medical school enrollment by 2,000. This decision led to mass resignations among residents and boycotts by medical students, disrupting hospital staffing and halting physician graduation. The crisis highlights deeper challenges in medical education and regional healthcare inequality.

The Republic of Korea now confronts three major health challenges: an aging population, increasing burdens of chronic disease, and the need for sustainable healthcare financing. Health policy must now be firmly grounded in science, long-term vision, and comprehensive institutional reform. Senior experts have proposed the following key policy directions [2]:

The first recommendation is to introduce or increase taxes on health-harming behaviors such as smoking, alcohol consumption, and the intake of sugary beverages. These behaviors drive preventable chronic diseases and strain the healthcare system. The increased revenues should be reinvested in preventive care and biohealth research and development.

Second, the digital transformation in healthcare should be accelerated, including the integration of artificial intelligence. Despite the Republic of Korea's reputation as a digital leader, its healthcare data infrastructure lacks full integration and interoperability. A centralized digital

Received: 23 June 2025
Accepted: 25 June 2025

Corresponding author:
Jong-Koo Lee
National Academy of Medicine of Korea, 51 Seochojungang-ro, Seocho-gu, Seoul 06654, Republic of Korea
E-mail: docmohw@snu.ac.kr

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

governance framework is necessary, alongside platforms such as My Healthway, legal reforms, and enhanced privacy protections.

Third, further innovation in biotechnology, particularly in cell therapy and xenotransplantation, is an important priority. The Republic of Korea's strong academic foundation has not yet translated into comparable clinical or commercial progress. Addressing this gap may require faster regulatory pathways, the expansion of public–private partnerships, improved ethical oversight, and enhanced global collaboration.

Fourth, reforms to healthcare R&D governance are essential. At present, research functions remain fragmented across various ministries. Greater efficiency and strategic alignment could be achieved through the creation of a presidential-level coordinating body and a unified agency modeled after United States' National Institutes of Health.

Fifth, the Korea National Institute of Health should be strengthened. This institution should serve as a strategic leader in national health research and emergency response, beginning with infectious diseases and expanding into chronic disease research and regional clinical research networks.

Sixth, academic medicine at university hospitals requires attention due to heavy clinical workloads, which have diminished research and teaching capacity. Policy measures that ensure protected research time for faculty, increased funding for investigator-led projects, and the development of a coordinated national strategy for academic medicine are necessary for revitalization.

Finally, national public health capacity should be enhanced by expanding the role of the Korea Disease Control and Prevention Agency beyond infectious disease control, to include chronic disease surveillance, community health performance assessment, and advanced workforce development. The establishment of a public health statistics center and new advanced training programs would further support these efforts.

These comprehensive proposals are designed to address both immediate concerns and longer-term shifts in global health. Effective collaboration among government, academia, and healthcare stakeholders, along with investment in structural reforms, will likely prove pivotal as the Republic of Korea's healthcare system faces this critical juncture. The recommendations presented here provide a roadmap toward a more equitable, resilient, and sustainable health future.

Notes

Ethics Approval

Not applicable.

Conflicts of Interest

Jong-Koo Lee has been the editor-in-chief of *Osong Public Health and Research Perspectives* since October 2021.

Funding

None.

References

1. Lee YH, Byeon JH, Kang CR, et al. Subacute sclerosing panencephalitis incidence following measles elimination efforts in the Republic of Korea. Osong Public Health Res Perspect 2025;16:285–91.
2. Korean Academy of Medical Sciences (KAMS). Strategic policy recommendations for a healthy Korea. KAMS. Forthcoming 2025. Korean.

https://doi.org/10.24171/j.phrp.2025.0235

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2024.0320
Osong Public Health Res Perspect 2025;16(3):195-210

Review Article

Effectiveness of non-pharmacological school-based therapies for cigarette smoking cessation among adolescents in South and Southeast Asian countries: a systematic review and meta-analysis

Fahad Ali Mangrio[1], Penpaktr Uthis[1,2], Suwimon Rojnawee[1], Alicia K Matthews[3]

[1]Faculty of Nursing, Chulalongkorn University, Bangkok, Thailand
[2]Research Unit for Enhancing Well-being in Vulnerable and Chronic Illness Populations, Faculty of Nursing, Chulalongkorn University, Bangkok, Thailand
[3]Columbia University School of Nursing, New York, NY, USA

ABSTRACT

Objectives: This review and meta-analysis examined the effectiveness of non-pharmacological therapies delivered through school-based interventions for smoking cessation among adolescents in South and Southeast Asian countries.

Methods: A systematic search was conducted across PubMed, Scopus, Science Direct, BioMed Central, the Cochrane Library, and ProQuest Dissertations & Theses Global from inception to October 2024. Eligible studies comprised randomized controlled trials and quasi-experimental studies that compared non-pharmacological smoking cessation interventions delivered in schools or other educational institutions. Data on smoking abstinence outcomes were extracted from published studies, and odds ratios (ORs) with 95% confidence intervals (CIs) were pooled using a random-effects model via the Mantel-Haenszel estimator.

Results: Seven studies involving 1,260 participants were included. The meta-analysis demonstrated that non-pharmacological school-based therapies significantly increased smoking abstinence compared to controls (OR, 2.83; 95% CI, 1.83–4.40; $p < 0.001$. Subgroup analyzes revealed benefits across both randomized controlled trials and quasi-experimental studies with varying abstinence rates. Studies utilizing biochemical verification showed significant positive effects despite substantial heterogeneity, and short-term (<3 months) abstinence was significantly higher in intervention groups compared to controls. Overall, no differences were found between subgroups regarding intervention effectiveness.

Conclusion: This meta-analysis indicates that non-pharmacological school-based interventions positively impact smoking abstinence rates, although effectiveness may vary based on study design, follow-up duration, and use of biochemical verification. The findings underscore the need for further research with larger sample sizes, extended follow-up periods, and improved methodological rigor in these regions.

Keywords: Non-pharmacological; School-based; Smoking cessation interventions; Southeast Asia; Systematic review and meta-analysis; Tobacco control

Received: November 22, 2024
Revised: February 14, 2025
Accepted: March 6, 2025

Corresponding author:
Penpaktr Uthis
Faculty of Nursing, Chulalongkorn University, Borommaratchachonnani Srisataphat Building Rama1 Road, Pathumwan Bangkok 10330, Thailand
E-mail: penpaktr.u@chula.ac.th

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Tobacco smoking is a major global health issue, with an estimated 24 million adolescents starting to smoke each year [1]. Epidemiological models predict that smoking may cause 1 billion deaths in the 21st century [2]. Smokers face a mortality rate 2 to 3 times higher than non-smokers and lose an average of 10 years of life expectancy due to smoking [3]. In South and Southeast Asia, tobacco use has been linked to over 40 to 160 million all-cause deaths [4,5]. Adolescence is a critical period of vulnerability to addictive substances such as nicotine. In Southeast Asia, adolescent smoking prevalence ranges from 11.3% to 19.2% in countries including Thailand, Malaysia, Cambodia, the Philippines, Singapore, Indonesia, East Timor-Leste, Brunei, Lao People's Democratic Republic, Myanmar, and Vietnam [6]. In the South Asian Region (SAR), which comprises Afghanistan, Bangladesh, Nepal, Bhutan, Pakistan, Maldives, Sri Lanka, and India, adolescent smoking rates range from 9.0% to 47.2% [7–11]. Initiating smoking during adolescence often leads to continued use into adulthood, and early nicotine exposure can adversely affect brain development, impairing cognition, attention, and mood [12,13]. Therefore, addressing nicotine dependence and promoting smoking cessation among adolescents is a critical public health priority.

During their formative years, children spend a substantial amount of time in school, where health behaviors are shaped through education, role modeling, and access to health promotion resources [14]. As such, school-based interventions have effectively enhanced adolescent health promotion and disease prevention, including tobacco control [14]. A previous systematic review of 14 school-based smoking cessation programs conducted between 2007 and 2019 concluded that these interventions can be effective for adolescent smokers [15]. These interventions incorporated behavioral support, peer-led initiatives, comprehensive health education, and integration into broader health curricula, contributing to successful smoking cessation. Nevertheless, limitations were identified, such as inconsistent results, short follow-up durations, a focus on public schools only, and challenges in program implementation across diverse settings. The heterogeneity in program designs and outcomes further complicated pooled statistical analysis [15]. Therefore, non-pharmacological interventions continue to be the primary recommendation for adolescent smoking cessation, with pharmacotherapy considered an alternative that warrants further investigation regarding its long-term efficacy and safety in adolescents [16].

Additionally, 2 reviews have produced mixed findings. One systematic review, which included 6 interventions, reported mixed results for school-based smoking cessation programs and found insufficient evidence to support their efficacy [17]. The review attributed this to high heterogeneity in intervention delivery methods and outcome measurement durations, which hindered robust statistical analysis [17]. In contrast, another rapid review and meta-analysis of 32 non-pharmacological interventions concluded that the intervention group achieved a 30% higher abstinence rate than the controls [18]. However, most studies in that review were conducted in the United States and Europe, and the smoking cessation programs varied considerably in terms of setting and counseling type [18].

Therefore, the efficacy of school-based smoking cessation therapies may differ significantly by geographical region, influenced by cultural, social, and economic factors that affect adolescents' responses to these interventions. For instance, most studies included in previous reviews [15,16,18] were conducted in Western countries, and their findings may not be applicable to regions such as Asia. Furthermore, earlier systematic reviews predominantly focused on public schools, overlooking private and alternative educational settings [15]. This narrow focus may neglect the distinct challenges and opportunities present in various educational environments.

Additionally, a Cochrane review analyzing randomized controlled trials (RCTs) demonstrated the effectiveness of both pharmacological and non-pharmacological smoking cessation therapies [16]. Non-pharmacological approaches, such as counseling, cognitive-behavioral therapy, and motivational interviewing, aim to modify behavior and thought patterns without the use of drugs [19]. This minimizes the risk of side effects commonly associated with pharmacological treatments, including nicotine replacement

HIGHLIGHTS

- Non-pharmacological therapies delivered in educational settings significantly improved smoking abstinence rates compared to control groups, with moderate heterogeneity (I^2=55%).
- Studies employing biochemical verification yielded significant effects, whereas those without verification showed higher quit rates that were not statistically significant.
- Five studies with follow-ups of ≤3 months reported increased abstinence rates and 2 studies with follow-ups of 4 to 6 months showed a positive trend, but without reaching statistical significance.

https://doi.org/10.24171/j.phrp.2024.0320

therapy, varenicline, and bupropion [19,20]. However, the review did not disaggregate data to examine outcomes specific to school-based interventions or to assess region-specific results globally [16].

Given the limitations of previous reviews and the increasing prevalence and consequences of adolescent smoking, it is essential to focus on smoking cessation interventions tailored to adolescents in South and Southeast Asia. Therefore, this review aims to statistically synthesize evidence on non-pharmacological therapies in various educational settings, including both public and private institutions, for adolescent smoking cessation. The study concentrates on South and Southeast Asian regions to identify regional trends and explore how cultural, social, and economic factors may influence intervention efficacy. Moreover, it highlights a gap in the literature due to the limited number of studies available in these settings. Ultimately, this review seeks to provide evidence-based recommendations for enhancing smoking cessation programs for adolescents in these regions, thereby reducing the prevalence and adverse effects of smoking.

Materials and Methods

Eligibility Criteria

This systematic review and meta-analysis followed the methodologies outlined in the Cochrane systematic review [21] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [22]. The review is registered with PROSPERO (CRD42023431655). Studies published in any language from school-based or educational institutions were included, although searches were conducted in English. Articles published from the inception of the databases until October 30, 2024, were considered.

Population

Studies were eligible if they provided data on outcomes for adolescent school-based smoking cessation programs conducted in the SAR, which comprises Afghanistan, India, Maldives, Bangladesh, Nepal, Pakistan, Sri Lanka, and Bhutan [23], or in Southeast Asia, which includes Thailand, the Philippines, Cambodia, Myanmar (Burma), Vietnam, Singapore, Malaysia, Indonesia, Brunei, Timor-Leste, and the Lao People's Democratic Republic [24]. These countries were selected due to their higher rates of adolescent smoking.

Interventions

Non-pharmacological therapy studies were defined as those that provided participants with non-drug treatments aimed at evaluating their effects on smoking cessation [19]. Examples include (1) psychological interventions (e.g., cognitive-behavioral education, counseling), (2) behavioral interventions (e.g., exercise), and (3) other interventions (e.g., meditation).

Control/comparator

Any type of control group was eligible, including waitlist controls, treatment-as-usual, or comparisons of the new experimental intervention with an existing intervention.

Outcome

The primary outcome of interest was the quit rate following exposure to the smoking cessation intervention. Given the heterogeneity of intervention approaches and follow-up assessment time points, the review included all reported quit outcomes, regardless of the definition of quit status (e.g., 7-day point prevalence versus continuous abstinence) and assessment timing (e.g., 1, 3, or 6 months post-treatment).

Study Design

Only RCTs and quasi-RCTs were included.

Exclusion Criteria

Interventions involving (1) drug-related treatments, (2) e-cigarettes containing nicotine, (3) studies that did not report any smoking cessation outcomes, and (4) review articles, protocols, case reports, letters, abstracts, or studies lacking data were excluded.

Search Strategy

In consultation with a librarian, a search strategy was developed using key terms—including Medical Subject Headings (MeSH), Boolean operators, and wildcards—based on previous reviews [19] and the guidelines for each database. We searched 6 databases—PubMed, Scopus, Science Direct, BioMed Central, the Cochrane Library, and ProQuest Dissertations & Theses Global—for relevant articles. Additionally, a manual search was conducted using reference lists from the included studies and publication profiles of smoking cessation experts (e.g., ResearchGate or university faculty pages).

Data Collection and Processing

Two reviewers (F.A.M. and S.R.) independently screened titles, abstracts, and full texts using the Joanna Briggs Institute Summary platform [25]. Any discrepancies were resolved through discussion among the reviewers or by consulting a third reviewer (P.U.) to reach consensus. Data extracted by F.A.M. and S.R. included basic study information (first author, year, country, and setting), participant characteristics (age, sex, sample size), intervention details (intervention fidelity, study

design, outcome measurement, intervention duration, follow-up period, type of intervention, intention-to-treat analysis, and main findings), and risk of bias ratings. A third reviewer (S.R.) verified all extracted information. Cohen kappa statistics were calculated to assess inter-rater agreement; the kappa statistic was 0.78 for title/abstract screening (indicating substantial agreement) and 0.85 for full-text screening (indicating almost perfect agreement), according to Landis and Koch's criteria [26].

Study Risk of Bias Assessment and Quality

Two reviewers (F.A.M. and S.R.) independently assessed the risk of bias using guidelines from the Cochrane Handbook for Systematic Reviews of Interventions and the Cochrane Tobacco Addiction Group [27]. Disagreements were resolved by consensus. The risk of bias for RCTs was evaluated using the RoB2 tool, which examines 5 domains: (1) randomization process, (2) deviations from intended interventions, (3) missing outcome data, (4) outcome measurement, and (5) selection of reported results. RCTs were categorized as having low risk, some concerns, or high risk of bias in accordance with Cochrane guidelines [23]. Additionally, the ROBINS-I tool was employed for quasi-experimental non-randomized studies, assessing 7 domains: (1) confounding variables, (2) participant selection, (3) intervention classification, (4) deviations from intended interventions, (5) missing data, (6) outcome measurement, and (7) selective reporting. Quasi-experimental studies were rated as having low, moderate, serious, or critical risk of bias, or as indeterminate [28]. The overall quality of evidence for each reported outcome was evaluated using the GRADE approach (Table S1) [29].

Data Synthesis

Cochrane Review Manager software version 5.4 was used to calculate odds ratios (ORs) with 95% confidence intervals (CI) for abstinence outcomes at all reported follow-up durations and the earliest available endpoints in each study. The Mantel-Haenszel method, fitted with a random-effects model, was employed to account for expected cross-trial variations in study design and intervention characteristics. This pooling method was also applied in all subsequent subgroup analyzes [30]. Heterogeneity between studies was assessed using the I^2 statistic according to the Cochrane guidelines for interpretation [31]. Publication bias was assessed using funnel plots and the Begg, Egger, and trim-and-fill tests [30], in Jamovi ver. 2.6. In addition, we conducted sensitivity and subgroup analyzes to examine whether study characteristics (e.g., design, sample size, outcome measurement report, type of abstinence measure, and follow-up period) affected the results [31]. Intention-to-treat analyzes were conducted, considering individuals lost to follow-up as current smokers [32].

Results

The initial search yielded 8,640 articles, including 3,610 from PubMed, 205 from Scopus, 999 from Science Direct, 515 from the Cochrane Library, 3,310 from ProQuest, and 1 from citation searching [18]. Most studies were published in English; no eligible non-English language trials were identified in the abstracts. After excluding ineligible automated records and duplicates, 4,643 articles were removed. Specifically, 100 observational studies, 3,713 articles unrelated to the research question and outcomes, and 10 articles based on abstract reviews were excluded. A full-text review and eligibility assessment were then performed on the remaining 171 articles by 2 independent reviewers, with discrepancies resolved by a third reviewer. Ultimately, 7 articles were deemed appropriate for inclusion (Figure 1).

Study Characteristics

Four of the 7 studies were conducted in Thailand and 3 in Malaysia [33–39]. Participants ranged from early adolescence to young adulthood, with an average age of 17.06 years. Two studies [37,39] included only male participants, while the remaining studies included both sexes [33–36,38]. Regarding study design, 5 studies were 2-arm RCTs (with 1 control and 1 intervention group) [33,35,36,38,39], and 2 studies employed a quasi-experimental 2-arm design without randomization [34,37] .

Intervention durations varied from 1 to 12 weeks; 71.4% of studies had short-term durations (1 to 6 weeks), while 28.6% had medium-term durations (10 to 12 weeks). Smoking cessation outcomes in 5 studies (71.4%) were confirmed via biological verification methods, such as expelled carbon monoxide, salivary cotinine, or urine cotinine levels. In contrast, 2 studies [34,38] relied solely on self-reported quit rates (Table 1) [33–39].

All studies utilized non-pharmacological smoking cessation interventions within school and university settings. For instance, 3 studies incorporated group-based cognitive-behavioral counseling [33,34,36], while 2 studies offered one-on-one counseling sessions [35,38]. Additionally, 2 studies implemented both group-based and individual face-to-face as well as mobile application–based counseling [37,39]. The interventions included the Fit and Smart Adolescent Smoking Cessation Program [33], the Project EX smoking cessation program [34], a smartphone application named "Quit with US" along with smoking cessation counseling [35], a group counseling smoking cessation package [36], a smoking

https://doi.org/10.24171/j.phrp.2024.0320

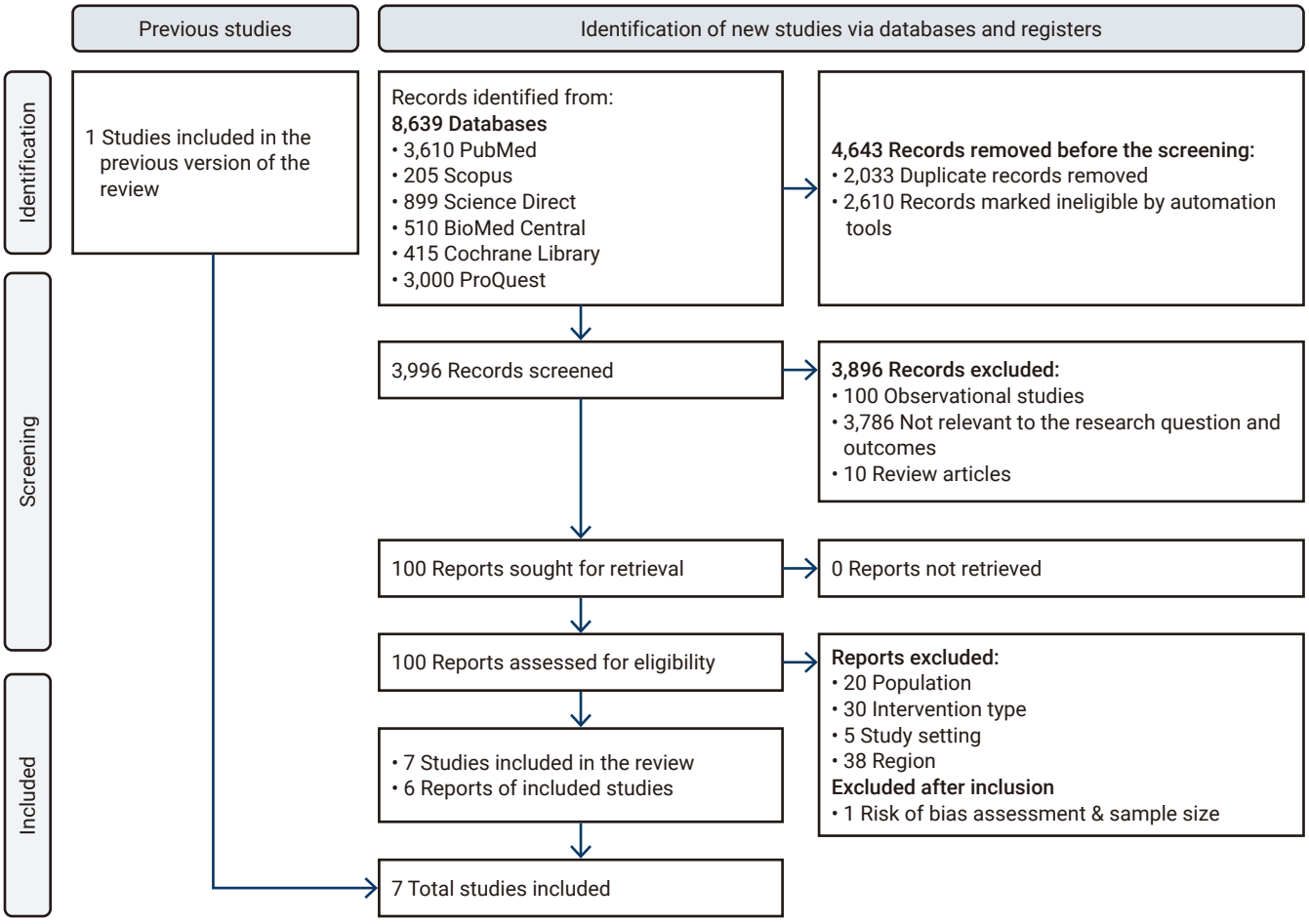


Figure 1. Inclusion of studies following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.

cessation program based on information-motivation-behavioral skills and stages of change [37], brief advice for smoking cessation [38], and a smoking cessation program [39].

Intervention facilitators varied considerably. Facilitators included a school teacher [33,34], a school counselor [33,36], a buddy [33], a pharmacist [35], a medical officer [38], and researchers from public health and sports science faculties [37,39]. Three studies described the training and supervision provided to the facilitators [33,34,36], whereas the remaining studies did not include any details regarding facilitator training [35,37–39] (Table 2).

Main Outcomes

A total of 1,260 adolescents and young adults participated in all 7 studies. All studies were theory-based and employed an intention-to-treat approach for outcome analysis. Three studies reported a 7-day point prevalence abstinence rate [33,35,36], 1 study reported a 30-day point prevalence abstinence [34], and 1 reported continuous 6-month abstinence [38]. Two studies did not report smoking cessation outcomes in terms of point prevalence or continuous abstinence [37,39]; however, they did document smoking abstinence immediately following intervention completion.

Due to the varied measures of smoking abstinence outcomes, all available results were pooled to provide a comprehensive analysis, ensuring that neither point prevalence nor continuous abstinence rates were disproportionately emphasized [30]. The analysis indicated that non-pharmacological school-based interventions were associated with significantly higher smoking abstinence rates compared to controls (OR, 2.83; 95% CI, 1.83–4.40; $p < 0.001$). The overall analysis showed moderate heterogeneity ($I^2 = 55\%$), suggesting some variability among study results that could affect the consistency and reliability of the overall conclusion (Figure 2).

Subgroup Comparisons

Subgroup analyzes were conducted to assess clinical

Table 1. Summary of findings from the included studies

No.	Study	Year	Country & setting	Mean age (y)	Sex	Total (*n*)	Fidelity	Study design	Outcome measurement	Treatment duration (wk)	Follow-up period (mo)	Intervention type		IT	Main findings	Risk of bias
												Control	Experimental			
1	Abdul Halim et al. [33]	2022	Malaysia, schools	15.3	Male & female	266	Yes	RCT	Self-reported biochemical (CO and salivary cotinine levels) 7-day point prevalence abstinence	4	3	Placebo	Fit and Smart Adolescent Smoking Cessation Program	Yes	At the 3-month follow-up, the quit rate in the intervention group (41.8%, 70/168) was significantly higher than that in the control group (24.5%, 24/98)	L
2	Chansatitporn et al. [34]	2016	Thailand, schools	15.12	Male & female	185	Yes	Quasi-exp	Self-reported 30-day point prevalence abstinence rate	6	3	Standard health education	Project EX: smoking cessation program	Yes	At the 3-month follow-up, the quit rate in the intervention group (23.0%, 21/91) was significantly higher than that in the control group (11.0%, 10/94)	M
3	Chulasai et al. [35]	2022	Thailand, universities	21.06	Male & female	273	Not mentioned	RCT	Self-reported biochemical (CO level), 7-day point prevalence abstinence	4	3	Smoking cessation counseling	A smartphone application named "Quit with US"	Yes	At the 3-month follow-up, the quit rate in the intervention group (58.4%, 80/137) was significantly higher than that in the control group (30.9%, 42/136)	L
4	Ismail et al. [36]	2010	Malaysia, schools	14	Male & female	346	Not mentioned	RCT	Self-reported, biochemical (CO level), 7-day point prevalence abstinence	10	4	Standard school programs for smoking cessation	Group Counseling; Smoking cessation package	Yes	At the 4-month follow-up, the quit rate in the intervention group (45.0%, 71/158) was significantly higher than that in the control group (32.0%, 60/188)	S

(Continued to the next page)

https://doi.org/10.24171/j.phrp.2024.0320

Table 1. Continued

No.	Study	Year	Country & setting	Mean age (y)	Sex	Total (*n*)	Fidelity	Study design	Outcome measurement	Treatment duration (wk)	Follow-up period (mo)	Intervention type		IT	Main findings	Risk of bias
												Control	Experimental			
5	Junnual et al. [37]	2019	Thailand, schools	16.8	Male	70	Not mentioned	Quasi-exp	Self-reported and biochemical (urinary cotinine) point prevalence abstinence or continuous abstinence rate not mentioned	12	3	Placebo	Information-Motivation-Behavioral-Stages of Change Smoking Cessation Program	Yes	At the 3-month follow-up, the quit rate in the intervention group (60.0%, 21/35) was significantly higher than that in the control group (35.0%, 12/35)	M
6	De Silva et al. [38]	2016	Malaysia, university	21.20	Male & female	80	Not mentioned	RCT	Self-reported continuous abstinence rate	1	6	Usual care (paper-based basic information on smoking cessation)	Brief advice	Yes	At the 6-month follow-up, the quit rate in the intervention group (15.0%, 6/40) was significantly higher than that in the control group (0%, 0/40)	H
7	Sarayuthpitak et al. [39]	2011	Thailand, schools	17.05	Male	40	Not mentioned	RCT	Self-reported biochemical (urinary cotinine) point prevalence abstinence or continuous abstinence rate not mentioned	6	Post-test	Usual care (information on smoking cessation)	Smoking cessation program	Yes	At the post-test, the quit rate in the intervention group (60.0%, 16/20) was significantly higher than that in the control group (10.0%, 2/20)	S

IT, intention to treat; RCT, randomized controlled trial; L, low risk bias; M, moderate risk bias; S, some concerns; H, high risk.

https://doi.org/10.24171/j.phrp.2024.0320

Table 2. Summary of study intervention description

No.	Study	Year	Intervention type	Theory	Intervention delivery and training provided	Description
1	Abdul Halim et al. [33]	2022	FSSCP	Social cognitive theory	Teachers, buddies, and trained school counselors received smoking cessation training from an expert counselor at the National Cancer Society of Malaysia. The training focused on program planning and the use of the FSSCP module. Group-based sessions delivered.	The FSSCP is a school-based, multi-component intervention designed to assist adolescents in quitting smoking conventional cigarettes. The intervention included 4 key components: (1) counseling based on social cognitive theory, (2) peer influence through a buddy system, (3) community involvement, and (4) implementation of a tobacco-free school policy. Each session lasts 45 to 60 minutes, focusing on motivation, social support, and maintaining smoke-free conditions.
2	Chansatitporn et al. [34]	2016	Project EX: smoking cessation program	Social cognitive theory	Through group-based sessions, the Thai public health research team trained teachers at Mahidol University on the history of teen tobacco cessation, recruitment, data collection, and curriculum delivery.	Project EX is a comprehensive smoking cessation program tailored explicitly for Thai adolescents. It involves 8 sessions, each lasting 40 to 45 minutes, focusing on motivational factors, personal skills, and coping strategies to help teens quit smoking. The program includes activities such as talk shows, alternative medicine techniques (like yoga and meditation), competitive games, and homework assignments to reinforce the quit attempt.
3	Chulasai et al. [35]	2022	A smartphone application named "Quit with US"	Self-efficacy theory and 5A's model	Pharmacists did not receive any training. Face-to-face individual session.	The "Quit with US" intervention involved a smartphone application designed to assist young adult smokers in quitting smoking. Participants in the intervention group received smoking cessation counseling from pharmacists at baseline and follow-ups, along with guidance to use the Quit with US app daily for 12 weeks, lasting between 15 and 30 minutes. The app provided self-instructional materials, including illustrations and texts, to support smoking cessation efforts. Participants were encouraged to log their progress and engage with the app to enhance their chances of achieving smoking abstinence.
4	Ismail et al. [36]	2010	Group counseling: Smoking cessation package	Cognitive-behavioral strategies	School counselors received training to implement the smoking cessation package module from psychologists, health promotion experts, educators, and clinical psychiatrists, who collaborated to develop the program. Group-based counseling sessions were provided.	The intervention consisted of a structured "Stop Smoking Module" in group counseling sessions conducted weekly for 10 weeks, each lasting 1.5 to 2 hours. The intervention focuses on various aspects of smoking cessation, such as preparation to stop smoking, decision-making, managing withdrawal symptoms, and maintaining cessation status. The program employed diverse strategies like discussions, handouts, video presentations, puzzles, games, and homework to engage participants and provide them with the necessary knowledge and skills to quit smoking.

(Continued to the next page)

https://doi.org/10.24171/j.phrp.2024.0320

Table 2. Continued

No.	Study	Year	Intervention type	Theory	Intervention delivery and training provided	Description
5	Junnual et al. [37]	2019	Information-Motivation-Behavioral-Stages of Change Smoking Cessation Program	Self-efficacy theory, information-motivation-behavioral skills model, and stages of change model	Researchers from public health have yet to receive any training. Group-based counseling and face-to-face sessions were provided.	The program employed a group-based counseling program designed to help male high school students quit smoking through a structured 12-week intervention. It combines the information-motivation-behavioral skills model and the stages of change model to enhance self-esteem, attitudes toward smoking, perceived control over tobacco, and reduced smoking behavior. The program included 6 main activities: introduction and motivation, self-esteem building, alternative solutions, counseling, experience sharing, and willpower enhancement.
6	De Silva et al. [38]	2016	Brief advice	Trans-theoretical model	Medical officers did not receive any training. Individual face-to-face sessions were provided.	Participants in the intervention group received 1-time brief advice during a face-to-face individual session focused on smoking cessation. This advice aimed to enhance their motivation to quit smoking.
7	Sarayuthpitak et al. [39]	2011	Smoking cessation program	Trans-theoretical model of health behavior change, Protection motivation theory, and Theory of reasoned action	Researchers from sports science. The researchers did not receive any training—group-based and face-to-face sessions were provided.	The smoking cessation intervention was a structured program designed to assist adolescents in quitting smoking through a combination of educational, behavioral, and social support strategies over 6 weeks. Each session lasted 30 minutes, and both group and individual sessions were used. It included 10 activities to improve knowledge, attitudes, and behaviors about smoking. Key components were health assessments, personalized messages, goal-setting contracts, smoke-free parties, self-help manuals, cessation counseling, buddy support, social networks, web resources, and relaxation techniques.

FSSCP, Fit and Smart Adolescent Smoking Cessation Program.

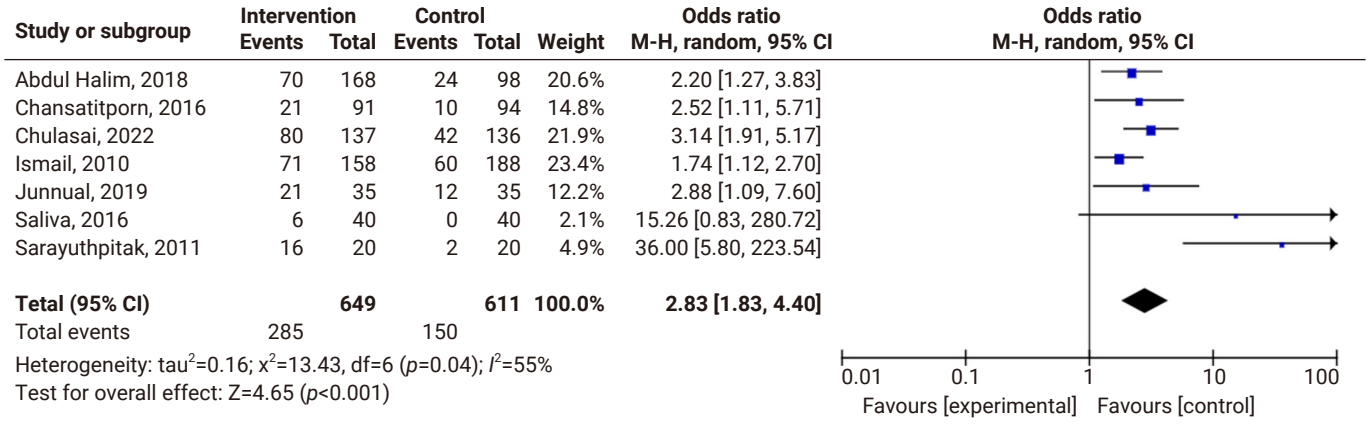


Figure 2. Non-pharmacological school-based smoking cessation interventions versus control. The plot presents the odds ratios (ORs) with 95% confidence intervals (CIs) for each study, along with the overall pooled estimate utilizing a random-effects model with the Mantel-Haenszel (M-H) method. Study-specific ORs (blue squares) are weighted by study size, with horizontal lines indicating 95% CIs. The pooled OR (black diamond) demonstrates an overall effect of 2.83 (95% CI, 1.83–4.40; Z = 4.65; $p < 0.001$) favoring the intervention. Heterogeneity analysis revealed $tau^2 = 0.16$, $x^2 = 13.43$, degrees of freedom (df) = 6, $p = 0.04$, and $I^2 = 55\%$, indicating moderate heterogeneity. The x-axis is presented on a logarithmic scale; values >1 favor the intervention, and values <1 favor the control.

and methodological heterogeneity that might influence the meta-analysis outcomes. First, smoking abstinence rates were compared based on study design. In 5 RCTs, abstinence rates ranged from 15% to 60%, while in 2 quasi-experimental studies, rates ranged from 23% to 60%. The RCT subgroup, comprising 5 studies with 1,005 participants, showed a significant positive effect (OR, 3.13; 95% CI, 1.69–5.82; $p=0.0003$) with substantial heterogeneity ($I^2=70\%$). The quasi-experimental subgroup, consisting of 2 studies with 255 participants, also demonstrated a significant effect (OR, 2.66; 95% CI, 1.42–4.97; $p=0.002$) and exhibited no heterogeneity ($I^2=0\%$). Overall, no significant subgroup difference was found in intervention effectiveness ($p=0.72$, $I^2=0\%$) (Figure 3).

Next, the subgroup with biochemical verification, which included 5 studies with 995 participants, showed statistically significant effects favoring the interventions (OR, 2.83; 95% CI, 1.68–4.77; $p<0.001$), with abstinence rates ranging from 41.8% to 60.0%, despite substantial heterogeneity ($I^2=66\%$). Conversely, the subgroup without biochemical verification, comprising 2 studies with 265 participants, reported higher quit rates; however, the differences were not statistically significant (OR, 3.62; 95% CI, 0.85–15.45; $p=0.08$), with abstinence rates ranging from 15.0% to 23.0% and lower heterogeneity ($I^2=30\%$). Overall, no subgroup differences in intervention effectiveness were identified (Figure 4).

The pooled analysis of 5 studies with follow-up assessments at 3 months or less yielded an OR of 3.15 (95% CI, 1.89–5.23; $p<0.001$), indicating a substantial increase in short-term abstinence rates among intervention participants compared to controls, with moderate heterogeneity ($I^2=55\%$). In the 4–6-month follow-up subgroup, comprising 2 studies, a positive trend was observed, although it did not reach statistical significance (OR, 3.18; 95% CI, 0.46–21.92; $p=0.24$). Overall, no subgroup differences in intervention effectiveness were identified (Figure 5).

Risk of Bias in Studies

All studies were evaluated for risk of bias using guidelines from the Cochrane Handbook for Systematic Reviews of

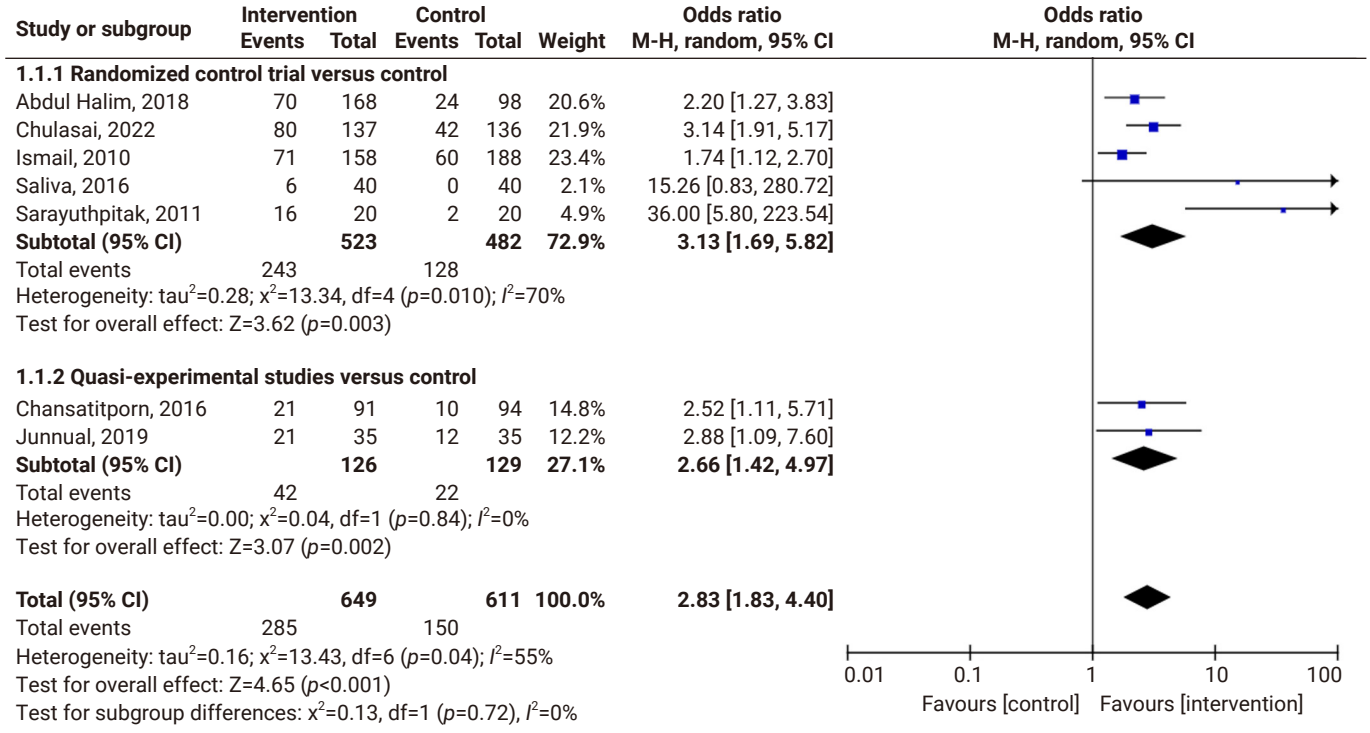


Figure 3. Efficacy of non-pharmacological smoking cessation interventions in randomized control trials vs. quasi-experimental studies. The forest plot depicts odds ratios (ORs) with 95% confidence intervals (CIs) for each study, as well as pooled estimates for each subgroup utilizing a Mantel-Haenszel (M-H) method of random-effects model. Study-specific ORs (blue squares) are weighted by study size, with horizontal lines indicating 95% CIs. The pooled OR (black diamond) illustrates the overall effect of each and the total subgroup analysis effect. The heterogeneity within randomized controlled trials is moderate ($tau^2=0.28$, $x^2=13.34$, degrees of freedom [df] = 4, $p=0.01$, $I^2=70\%$), while the heterogeneity in quasi-experimental studies is low ($tau^2=0.00$, $x^2=0.04$, df = 1, $p=0.84$, $I^2=0\%$). The overall test for effect shows statistical significance in favor of the intervention group ($Z=4.65$, $p<0.001$). The test for subgroup differences indicates no significant variation between study designs ($x^2=0.13$, df = 1, $p=0.72$, $I^2=0\%$).

https://doi.org/10.24171/j.phrp.2024.0320

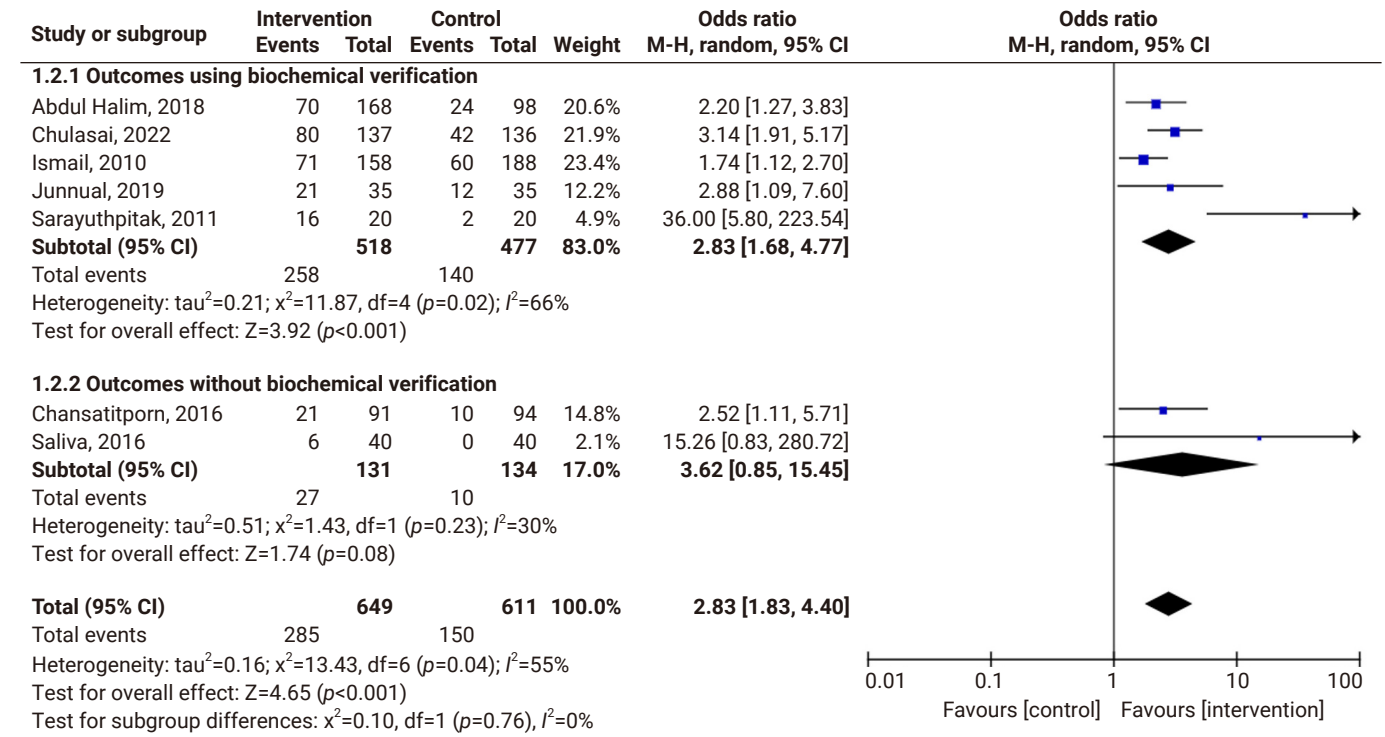


Figure 4. Comparison of interventions that included biological verification of smoking cessation versus interventions that did not. This forest plot displays odds ratios (ORs) and 95% confidence intervals (CIs) for studies using biochemical verification and those without. Blue squares represent study-specific ORs, with size proportional to study weight. Horizontal lines indicate 95% CIs, and black diamonds show pooled estimates. The x-axis is logarithmic, with values >1 favoring the intervention and <1 favoring the control. Studies using biochemical verification ($n=5$) show a significant effect favoring the intervention (OR, 2.83; 95% CI, 1.68–4.77; $p<0.001$) with moderate heterogeneity ($I^2=66\%$). Studies without biochemical verification ($n=2$) show a non-significant effect (OR, 3.62; 95% CI, 0.85–15.45; $p=0.08$) with low heterogeneity ($I^2=30\%$). The overall effect favors the intervention (OR, 2.83; 95% CI, 1.83–4.40; $p<0.001$). No significant difference was found between verification methods ($p=0.76$, $I^2=0\%$).
M-H, Mantel-Haenszel; df, degrees of freedom.

Interventions. The Cochrane Tobacco Addiction Group classified blinding-related bias as low risk, given the inherent challenges in blinding within smoking cessation intervention research [30]. A risk of bias assessment was conducted for 5 studies using the standardized RoB2 tool, which evaluates 5 key domains [33,35,36,38,39]. Two studies were categorized as having a low risk of bias (2/5, 40.0%), indicating low bias risk across all domains [33,35]. Two studies raised some concerns (2/5, 40.0%) [36,39]; specifically, 1 study had concerns in 1 domain (missing outcomes) without any high-risk ratings [36], while another study had concerns in 2 domains (randomization process and outcome measurement) but no high-risk ratings [39]. One study exhibited a higher risk of bias (1/5, 20.0%) in the randomization process domain [38].

Three quasi-experimental studies were assessed using the ROBINS-I tool across 7 domains [34,37,40]. Two studies showed moderate risk (2/3, 66.6%), indicating moderate risk in at least 2 domains [34,37], with 1 study showing concerns in 3 domains (Figures 4, 5; Supplementary Materials 1–3; Figure S1–S3) [40]. Due to the critical risk identified in the quasi-experimental study by Sidhu et al. [40] and potential publication bias, that study was excluded from the analysis. Potential bias in participant selection (domain 2) may arise from disparities in baseline characteristics or inadequate reporting of deviations from the intended intervention (domain 4), which could impact the outcomes and introduce bias due to the lack of fidelity assessment and inconsistencies in the implementation of the intervention. Low retention rates may lead to bias due to missing outcome data (domain 5) and concerns about bias in measuring the outcome (domain 6) as a result of relying solely on self-reports without biochemical verification. Bias in selecting reported results (domain 7), with significant prevention effects but not cessation effects, may indicate selective outcome reporting. Additionally, the study included a small sample size ($n=16$) of cigarette smoker participants. A sensitivity analysis comparing results with and without this study was conducted, and the main findings remained

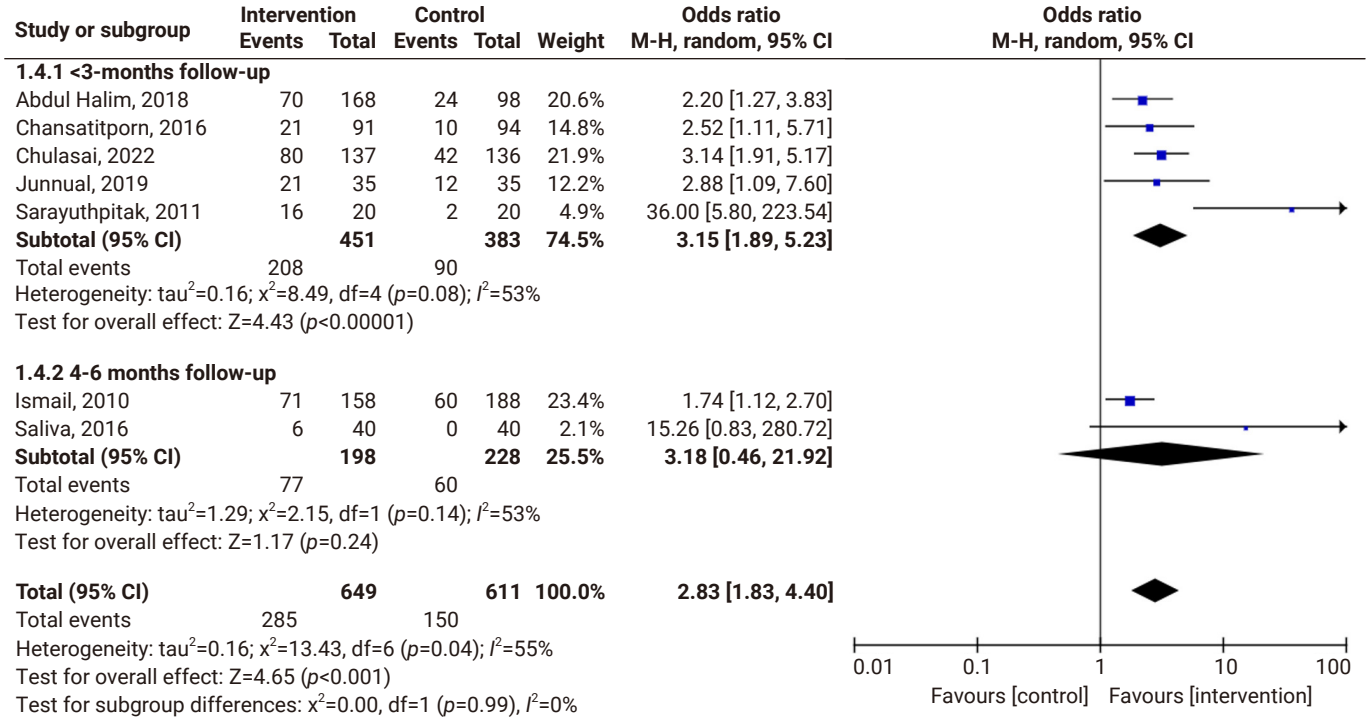


Figure 5. Comparison of studies with ≤3-month follow-up versus studies 4- to 6-month follow-up periods. The odds ratio (OR) (Mantel-Haenszel [M-H], random, 95% confidence interval [CI]) is presented for each study, with corresponding weight and CIs. The subtotal effect estimates for the <3-month and 4- to 6-month follow-up subgroups are provided, along with the overall pooled effect estimate. The results indicate that in the <3-month follow-up subgroup, the pooled OR is 3.15 (95% CI, 1.89–5.23), favoring the intervention group, with moderate heterogeneity (I^2 = 53%). In the 4- to 6-month follow-up subgroup, the pooled OR is 3.18 (95% CI, 0.46–21.92), showing a wider confidence interval and higher uncertainty. The test for subgroup differences (x^2 = 0.00, degrees of freedom [df] = 1, p = 0.99) indicates no significant variation between follow-up periods (I^2 = 0%).

statistically significant and clinically meaningful even after its exclusion [40].

GRADE Approach

The GRADE assessment indicated that the certainty of evidence for the treatment effect on quit rates at less than 6 months in the RCTs was moderate. Similarly, the 2 quasi-experimental studies demonstrated moderate certainty regarding quit rates at less than 3 months (Table S1).

Discussion

This is the first systematic review and meta-analysis examining the effectiveness of non-pharmacological, school-based therapies for smoking cessation among adolescents in South and Southeast Asian regions. The review included 7 studies conducted in the screened, eligible, approached, randomize (SEAR), providing valuable insights into the efficacy of these interventions. Notably, the review exposes a significant research gap regarding school-based smoking cessation interventions in the SAR. Despite the high prevalence of cigarette smoking among adolescents in this region [7–11], there is a notable absence of studies addressing this critical issue. This finding aligns with previous research and underscores the urgent need for more focused investigations in this geographical area [41]. The varying levels of bias across the included studies highlight the importance of employing rigorous methodological approaches. Robust testing strategies were applied to assess the risk of bias, resulting in the exclusion of 1 quasi-experimental study due to its critical risk of bias associated with a small sample size and insufficient reporting on intervention allocation and outcomes [40]. Such exclusions are essential to ensure the reliability of the synthesized evidence and maintain the integrity of the review process [31].

Results indicated that the interventions significantly increased smoking abstinence rates compared to control conditions. However, the presence of moderate heterogeneity among the studies (I^2 = 55%) suggests some variability in intervention effects, which may be attributable to differences in study populations, sample sizes, intervention types, or

https://doi.org/10.24171/j.phrp.2024.0320

methodological variations. These findings are consistent with previous systematic reviews and meta-analyzes that demonstrated the efficacy of various smoking cessation interventions despite considerable heterogeneity [18,42]. Nonetheless, the observed variability highlights the need for further research to identify the factors contributing to these differences.

The subgroup of studies that employed biochemical verification demonstrated a significant positive effect of the interventions. This robust effect size, based on objective outcome measures, provides strong evidence for the effectiveness of the smoking cessation programs evaluated in this review. In contrast, the subgroup relying on self-reported quit outcomes showed a positive effect that did not reach statistical significance, and the wide CI indicates less precision in the estimate, likely due to a smaller number of studies and participants. Previous research has documented substantial differences in smoking cessation outcomes between verified and non-verified measures, with evidence suggesting that reporting bias can inflate estimates of intervention effects [43]. This bias may occur because participants tend to overstate their progress due to social desirability [44]. Therefore, biochemical verification remains the gold standard for rigorously evaluating the effectiveness of smoking cessation interventions [45]. Once efficacy has been firmly established using objective measures, self-reported outcomes from school-based studies may still offer valuable insights into how the intervention performs in everyday life and capture social and environmental factors that are difficult to replicate in controlled trials.

Subgroup analysis of follow-up periods of 3 months or less indicated a substantial increase in short-term quit rates among intervention participants compared to controls. This finding aligns with previous meta-analyzes that demonstrated the effectiveness of behavioral interventions for short-term smoking cessation among adolescents [18,42]. The 4- to 6-month follow-up subgroup exhibited a positive trend, although it did not reach statistical significance. While less definitive, this result is consistent with the known challenges of maintaining quit rates over time, as observed in other non-pharmacological intervention studies [46].

The included studies displayed significant variability in intervention duration, ranging from short to medium-term programs. This diversity suggests that short to medium-term programs can effectively promote smoking cessation among adolescents. However, 1 study featured a 1-week intervention duration and reported lower smoking abstinence rates than interventions of longer duration [38]. These findings are consistent with previous systematic reviews that indicate the effectiveness of tobacco cessation programs across durations ranging from 4 to 12 weeks, as well as longer-term interventions lasting up to 6 months [18,42].

The individuals delivering non-pharmacological interventions were diverse, including schoolteachers, public health researchers, pharmacists, school counselors, and trained researchers. This diversity suggests a multi-faceted approach to addressing adolescent smoking, drawing on expertise from various fields to design and implement effective cessation programs. However, the reviewed studies varied in the level of detail provided regarding the training received by intervention providers. Three studies described the training interventions in detail, while the remaining studies did not report any information on the training of providers [31,32,34]. This discrepancy complicates the assessment of intervention effectiveness, as the quality of provider training can significantly influence smoking cessation outcomes. Future research would benefit from comprehensive reporting on the nature and content of training, along with ongoing supervision to ensure fidelity to the intervention protocol. Such reporting would facilitate a better understanding of the role of training in the success of smoking cessation interventions and support the development of more effective programs.

The limitations of the current research include its exclusive focus on non-pharmacological interventions within school and educational institution settings, which may not capture the full spectrum of smoking cessation programs available outside academic environments. Furthermore, the geographical concentration of the included studies in Southeast Asian countries may limit the generalizability of the findings to other regions. We also acknowledge inherent limitations in meta-analysis, such as the small number of studies and moderate heterogeneity, which may influence the overall results.

This review confirms that all studies obtained informed consent, including parental consent and adolescent assent for participants under 18. However, some studies provided insufficient details regarding procedural descriptions, raising concerns about the adequacy of ethical reporting [35,38,39]. Comprehensive documentation of consent processes is essential for transparency and ethical rigor in adolescent smoking cessation interventions [47]. Future research should prioritize standardized reporting of ethical procedures to enhance study comparability and integrity.

Moreover, future research should explore additional methodological approaches, such as the Hartung-Knapp adjustment and the use of R calculator to refine effect estimates. Studies should also address these limitations by expanding the focus beyond educational institutions to encompass a broader geographic scope, including various

South and Southeast Asian countries. These interventions show potential efficacy and warrant adaptation and evaluation in other nations in the region. Methodological considerations such as biochemical verification, detailed training protocols, fidelity evaluations, and standardized cessation assessment periods should be addressed to strengthen methodological rigor. Additional research is needed to identify the minimal intervention components required to achieve optimal outcomes, particularly in resource-constrained settings.

Future research should focus on customizing smoking cessation interventions to address the unique cultural and socioeconomic factors influencing smoking behavior in South Asian countries. Further studies should aim to increase sample sizes and standardize long-term follow-up periods to better understand the long-term efficacy of non-pharmacological, school-based smoking cessation interventions.

Conclusion

This review provides significant insights into the effectiveness of school-based non-pharmacological interventions for smoking cessation among adolescents in the Southeast Asian region. While these interventions have demonstrated short-term effectiveness, long-term follow-up studies are needed to assess sustained outcomes. Healthcare providers and policymakers should consider implementing and promoting these interventions to reduce smoking prevalence and mitigate its associated health burdens.

Supplementary Material

Table S1. GRADE approach summary of findings. **Supplementary Material 1.** Full structured search strategies for each database. **Supplementary Material 2.** Publication bias tests: Begg rank correlation. **Supplementary Material 3.** Full-text screening: excluded studies. **Figure S1.** Funnel plot of primary meta-analysis. **Figure S2.** Risk of bias assessment. (A, B) Randomized controlled trial studies risk of bias assessment (RoB2). **Figure S3.** (A, B) Risk bias ROBIN-1 assessment of quasi-experimental design. Supplementary data are available at https://doi.org/10.24171/j.phrp.2024.0320.

Notes

Ethics Approval
Not applicable.

Conflicts of Interest
The authors have no conflicts of interest to declare.

Funding
None.

Availability of Data
No additional data. The article contains all the data.

Authors' Contributions
Conceptualization: FAM, AKM; Data curation: FAM, PU, AKM; Formal analysis: FAM, SR, AKM; Methodology: PU, FAM; Supervision: PU, AKM; Writing–original draft: FAM; Writing–review & editing: all authors. All authors read and approved the final manuscript.

Acknowledgements
The author, Fahad Ali Mangrio, would like to acknowledge the scholarship program for Association of Southeast Asian Nations (ASEAN) and non-ASEAN Countries and the 90th Anniversary supported by the graduate school of Chulalongkorn University Bangkok, Thailand. This research work is part of a PhD dissertation in nursing.

References

1. World Health Organization (WHO) (CH). WHO report on the global tobacco epidemic, 2023: protect people from tobacco smoke [Internet]. WHO; 2023 [cited 2024 Aug 24]. Available from: https://www.who.int/publications/i/item/9789240077164.
2. Eriksen MP, Whitney CF. Tobacco control policies. In: Hecht SS, Hatsukami DK, editors. Tobacco and cancer: the science and the story. World Scientific Publishing; 2022. p. 343–65.
3. Jha P, Peto R. Global effects of smoking, of quitting, and of taxing tobacco. N Engl J Med 2014;370:60–8.
4. Kaur J, Rinkoo AV, Richardson S. Update on numbers of tobacco-attributable deaths by country in the South-East Asia region: implications for policy. Tob Control 2024 Jun 8 [Epub]. https://doi.org/10.1136/tc-2024-058599.
5. Institute for Health Metrics and Evaluation (US). Global Burden of Disease (GBD) 2019. University of Washington; 2021.
6. Southeast Asia Tobacco Control Alliance (SEATCA) (TH). Nicotine and tobacco addiction [Internet]. SEATCA; 2019 [cited 2024 Aug 20]. Available from: https://aseantobaccocontrolatlas.org/ch2/.
7. Mangrio FA, Uthis P, Rojnawee S, et al. Prevalence, risk factors, and intervention approaches related to adolescent smoking in Pakistan: a scoping review. J Dow Univ Health Sci 2024;18:112–25.
8. Sharma P, Pahari S, Acharya SR, et al. Tobacco consumption and its associated factors among Nepalese students. Open Public Health J 2021;14:446–54.
9. International Institute for Population Sciences (IIPS); Ministry of Health and Family Welfare (IN). Global Youth Tobacco Survey (GYTS-4), India 2019: report. IIPS; 2019.
10. Alemi Q, Stempel C, Montgomery S. Prevalence and social determinants of tobacco use in Afghanistan. Int Health 2021;13:3–12.
11. Katulanda P, Liyanage IK, Wickramasinghe K, et al. Tobacco smoking among school children in Colombo district, Sri Lanka. Asia Pac J Public Health 2015;27:NP278–87.
12. Fergusson DM, Horwood LJ. Transitions to cigarette smoking during

adolescence. Addict Behav 1995;20:627–42.
13. Goriounova NA, Mansvelder HD. Short- and long-term consequences of nicotine exposure during adolescence for prefrontal cortex neuronal network function. Cold Spring Harb Perspect Med 2012; 2:a012120.
14. Lynagh M, Schofield MJ, Sanson-Fisher RW. School health promotion programs over the past decade: a review of the smoking, alcohol and solar protection literature. Health Promot Int 1997;12:43–60.
15. Hanpatchaiyakul K, Songsuda AB, Amornrojanavaravutti MW, et al. Effectiveness of school-based smoking cessation programs. J Res Curric Dev 2022;12:67–77.
16. Fanshawe TR, Halliwell W, Lindson N, et al. Tobacco cessation interventions for young people. Cochrane Database Syst Rev 2017; 11:CD003289.
17. Hutton HE, Wilson LM, Apelberg BJ, et al. A systematic review of randomized controlled trials: web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011;13:227–38.
18. Choi Y, Lee CM, Cho B, et al. Behavioral interventions for smoking cessation among adolescents: a rapid review and meta-analysis for the Korea Preventive Services Task Force. Osong Public Health Res Perspect 2021;12:177–86.
19. Nian T, Guo K, Liu W, et al. Non-pharmacological interventions for smoking cessation: analysis of systematic reviews and meta-analyses. BMC Med 2023;21:378.
20. Selph S, Patnode C, Bailey SR, et al. Primary care-relevant interventions for tobacco and nicotine use prevention and cessation in children and adolescents: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2020;323:1599–608.
21. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane handbook for systematic reviews of interventions. Cochrane Database Syst Rev 2019; 10:ED000142.
22. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1.
23. Alexeeva NN, Sivaramamurti C, Yefremov YK, et al. South Asia [Internet]. Encyclopedia Britannica; 2024 [cited 2024 Sep 9]. Available from: https://www.britannica.com/place/South-Asia.
24. Frederick WH, Leinbach TR. Southeast Asia [Internet]. Encyclopedia Britannica; 2024 [cited 2024 Sep 9]. Available from: https://www.britannica.com/place/Southeast-Asia.
25. Munn Z, Aromataris E, Tufanaru C, et al. The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the unified management, assessment and review of information (JBI SUMARI). Int J Evid Based Healthc 2019;17:36–43.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
27. Sterne JA, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomized trials. BMJ 2019;366:l4898.
28. Eldridge S, Campbell M, Campbell M, et al. Revised Cochrane risk of bias tool for randomized trials (RoB 2.0): additional considerations for cluster-randomized trials. Cochrane Methods Cochrane Database Syst Rev 2016;10(Suppl 1).
29. Schunemann HJ, Higgins JP, Vist GE, et al. Chapter 14: completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JP, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions. 2nd ed. Cochrane Training; 2019. p. 375–402.
30. Deeks JJ, Higgins JP, Altman DG, et al. Chapter 10: analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions. 2nd ed. Cochrane Training; 2019. p. 241–84.
31. Boutron I, Page MJ, Higgins JP, et al. Chapter 7: considering bias and conflicts of interest among the included studies. In: Higgins JP, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions. 2nd ed. Cochrane Training; 2019. p. 177–204.
32. Detry MA, Lewis RJ. The intention-to-treat principle: how to assess the true effect of choosing a medical treatment. JAMA 2014;312:85–6.
33. Abdul Halim NA, Wee LH, Mohd Saat NZ, et al. Application of the logic model to the school-based fit and smart adolescent smoking cessation programme. Malays J Med Sci 2022;29:133–45.
34. Chansatitporn N, Charoenca N, Sidhu A, et al. Three-month effects of Project EX: a smoking intervention pilot program with Thai adolescents. Addict Behav 2016;61:20–4.
35. Chulasai P, Chinwong D, Vientong P, et al. Smartphone application for smoking cessation (Quit with US): a randomized controlled trial among young adult light smokers in Thailand. Int J Environ Res Public Health 2022;19:8265.
36. Ismail R, Aljunid SM, Latip KA, et al. Effectiveness of group counseling in smoking cessation program amongst adolescent smokers in Malaysia. Med J Indones 2010;19:273–9.
37. Junnual N, Sota C, Chaikoolvatana A. Effectiveness of a smoking cessation program on self-esteem, attitude, perception, and practice regarding control over smoking among male high school. J Health Res 2019;33:366–74.
38. De Silva W, Awang R, Samsudeen S, et al. A randomised single-blinded controlled trial on the effectiveness of brief advice on smoking cessation among tertiary students in Malaysia. J Health Med Inform 2016;7:217.
39. Sarayuthpitak J, Kritpet T, Nakornkhet K. Smoking cessation program for Thai male adolescents. Pacific Rim Int J Nurs Res 2011; 15:111–24.
40. Sidhu AK, Sussman S, Tewari A, et al. Project EX-India: a classroom-based tobacco use prevention and cessation intervention program. Addict Behav 2016;53:53–7.
41. Iqbal S, Barolia R, Petrucka P, et al. Smoking cessation interventions in South Asian Region: a systematic scoping review. BMC Public

Health 2022;22:1096.
42. Park E, Zhou Y, Chen C, et al. Systematic review: interventions to quit tobacco products for young adults. BMC Public Health 2023;23:1233.
43. Scheuermann TS, Richter KP, Rigotti NA, et al. Accuracy of self-reported smoking abstinence in clinical trials of hospital-initiated smoking interventions. Addiction 2017;112:2227–36.
44. Durmaz A, Dursun I, Kabadayi ET. Mitigating the effects of social desirability bias in self-report surveys: classical and new techniques. In: Baran ML, Jones JE, editors. Applied social science approaches to mixed methods research. IGI Global Scientific Publishing; 2020. p. 146–85.
45. Benowitz NL, Bernert JT, Foulds J, et al. Biochemical verification of tobacco use and abstinence: 2019 update. Nicotine Tob Res 2020; 22:1086–97.
46. Saroj SK, Bhardwaj T. Non-pharmacological interventions for tobacco cessation: a systematic review of existing practices and their effectiveness. Monaldi Arch Chest Dis 2022;92.
47. Moolchan ET, Mermelstein R. Research on tobacco use among teenagers: ethical challenges. J Adolesc Health 2002;30:409–17.

https://doi.org/10.24171/j.phrp.2024.0320

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0006
Osong Public Health Res Perspect 2025;16(3):211-222

Review Article

The role of vitamin D and the *VDR* gene polymorphism in sepsis risk and mortality: a systematic review and meta-analysis

Wiwi Jaya[1,2,3], Sumarno Reto Prawiro[4], Hani Susianti[3,5], Aswoco Andyk Asmoro[2]

[1]Doctoral Program in Medical Sciences, Faculty of Medicine, Universitas Brawijaya, Malang, Indonesia
[2]Department of Anesthesiology and Intensive Therapy, Faculty of Medicine, Universitas Brawijaya, Malang, Indonesia
[3]Dr. Saiful Anwar General Hospital, Malang, Indonesia
[4]Department of Microbiology, Faculty of Medicine, Universitas Brawijaya, Malang, Indonesia
[5]Departement of Clinical Pathology, Faculty of Medicine, Universitas Brawijaya, Malang, Indonesia

ABSTRACT

Objectives: Vitamin D regulates immune function, cell proliferation, and differentiation. Its deficiency is linked to sepsis, although the causal relationship remains unclear. Studies suggest a strong correlation between FokI polymorphism and sepsis in the context of vitamin D deficiency. This study examined the association between vitamin D levels, the *VDR* FokI polymorphism, and sepsis risk through a systematic review and meta-analysis.

Methods: Relevant articles from 2014–2024 were identified from various databases, including PubMed, Scopus, and Cochrane. A meta-analysis was conducted to assess the difference in vitamin D levels between the sepsis and control groups, as well as the relationship between *VDR* FokI genotypes (TT, CT, CC) and sepsis risk.

Results: Vitamin D levels in sepsis patients were consistently lower than in the control group, with a mean difference of −4.17 ng/mL (95% confidence interval, −7.87 to −0.47; $p=0.03$). However, the relationship between *VDR* FokI genotype and sepsis risk was not statistically significant ($p>0.05$), although several individual studies showed a positive correlation. High heterogeneity was found in the analysis of vitamin D levels ($I^2=100\%$) and FokI genotypes ($I^2=91\%$), which affected the interpretation of the results.

Conclusion: Vitamin D deficiency is a potential risk factor for sepsis, while the relationship between the *VDR* FokI polymorphism and sepsis risk requires further investigation. These findings highlight the importance of early detection of vitamin D deficiency as a preventive strategy in at-risk populations, although additional studies with more standardized designs are needed to definitively confirm this relationship.

Keywords: FokI; Polymorphism, genetic; *VDR* gene; Vitamin D; Sepsis

Received: January 14, 2025
Revised: March 16, 2025
Accepted: March 17, 2025

Corresponding author:
Wiwi Jaya
Doctoral Program in Medical Sciences, Faculty of Medicine, Universitas Brawijaya, Jl. Veteran No.10-11, Ketawanggede, Kec. Lowokwaru, Kota Malang, Jawa Timur 65145, Indonesia
E-mail: wiwi.jaya@ub.ac.id

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Sepsis is a life-threatening condition characterized by organ dysfunction resulting from a dysregulated immune response to infection [1]. Sepsis and septic shock represent major global health issues, affecting millions annually and causing death in approximately one out of every 4 individuals [2]. The global incidence of sepsis is estimated to range from 276 to 678 cases per 100,000 people annually, with mortality rates between 22.5% and 26.7%. Vulnerable populations include the elderly, pregnant women, neonates, hospitalized patients, intensive care unit (ICU) patients, individuals with compromised immune systems, and those with chronic medical conditions [3].

Vitamin D, a secosteroid, was once considered solely responsible for maintaining calcium and phosphorus balance in bones. Recent research, however, has revealed its role in immunomodulation, cell proliferation, and differentiation following vitamin D administration [4]. A retrospective cohort study by Braun, which analyzed 1,325 samples, found that ICU patients with vitamin D levels below 15 ng/mL exhibited a significant association with higher mortality rates among sepsis patients [5]. Other studies have also linked reduced serum vitamin D levels to conditions such as diabetic foot ulcers, suggesting that inflammation and increased infection susceptibility may be involved in the pathomechanism [6,7].

The vitamin D receptor (VDR) is expressed in macrophages, lymphocytes, and monocytes. When stimulated by 25-hydroxyvitamin D (25(OH)D), these cells can increase the secretion of antimicrobial peptides like cathelicidin. Cathelicidin is effective against various pathogens, including Gram-positive and Gram-negative bacteria, fungi, and microbes present in the skin, respiratory, and digestive tracts. Patients with 25(OH)D levels below 20 ng/mL are unable to produce sufficient cathelicidin, thereby increasing their susceptibility to nosocomial infections such as pneumonia and sepsis [8]. Furthermore, vitamin D binding to its receptor has been implicated in the pathogenesis of foot ulceration in individuals with diabetes [9]. There is also growing evidence that various *VDR* polymorphisms may play roles in a range of diseases, including hypertension, non-alcoholic fatty liver disease, cancer, and obesity [10].

Genetic polymorphisms, which are common variations in DNA sequences among populations, can affect biological functions. Recent studies have identified several polymorphisms in the *VDR* gene; however, their impact on protein function and VDR signaling remains incompletely understood. Among these, the FokI polymorphism, located in the translation initiation region of the *VDR* gene, produces VDR proteins with different structures. A study by Zeljic et al. [11] in Serbia involving 100 ICU sepsis patients found that the *VDR* FokI polymorphism (rs2228570) was associated with increased mortality within the first 28 days of sepsis.

HIGHLIGHTS

- Sepsis patients were found to have significantly lower vitamin D levels, suggesting that hypovitaminosis D may be a risk factor.
- This study also identifies a potential link between the *VDR* FokI gene polymorphism and increased sepsis risk.
- These findings underscore the importance of early detection of vitamin D deficiency as a preventive strategy in at-risk populations.

This study focused on the FokI polymorphism in the *VDR* gene due to its unique potential to influence vitamin D metabolism and its significant impact on VDR functionality. This polymorphism results in the production of different VDR protein forms that alter the body's response to vitamin D. Although other polymorphisms such as ApaI, BsmI, and TaqI were considered, they were not prioritized because of limited evidence supporting their clinical significance in sepsis. The ability of the FokI polymorphism to potentially affect vitamin D dosing and therapeutic responses, particularly in sepsis patients, renders it a more pertinent target for analysis.

Despite an expanding body of literature on the role of vitamin D in sepsis, several key gaps remain. Prior studies on vitamin D levels in ICU-treated sepsis patients have been constrained by small sample sizes, methodological variability, and the absence of standardized protocols for vitamin D measurement and supplementation [12,13]. Moreover, many studies have not accounted for potential genetic variations in the *VDR* gene, particularly the FokI polymorphism, which may influence vitamin D metabolism and immune responses. These limitations have contributed to inconsistent findings, making it challenging to draw definitive conclusions about the clinical implications of vitamin D deficiency in sepsis management. In contrast, our study addresses these gaps by performing a comprehensive meta-analysis that includes only studies with rigorous methodological designs, such as large sample sizes and consistent measurement protocols, while also examining the role of *VDR* gene polymorphisms in modulating vitamin D effects in sepsis patients.

This study differs from previous systematic reviews and meta-analyses in several ways. First, while earlier reviews have predominantly focused on vitamin D levels and sepsis outcomes, our review uniquely incorporates an analysis of the *VDR* FokI polymorphism, which may modulate

https://doi.org/10.24171/j.phrp.2025.0006

vitamin D effects and contribute to variability in sepsis susceptibility and outcomes. Second, previous studies often suffered from insufficient sample sizes or methodological inconsistencies, issues we addressed by including only studies that met stringent inclusion criteria in accordance with Cochrane's guidelines, even if some were of poor quality, and by addressing heterogeneity through subgroup analysis or meta-regression when necessary. Finally, this review provides a detailed risk of bias assessment using the Newcastle-Ottawa scale, which evaluates the influence of study design and methodology on the findings, allowing for a more nuanced interpretation of the evidence. Overall, our aim is to offer a comprehensive and reliable understanding of the relationships between vitamin D levels, the *VDR* FokI polymorphism, and sepsis risk and mortality.

Materials and Methods

Eligibility Criteria

The inclusion criteria for articles were (1) study type: Only studies published in English and available as full-text articles were considered. The selection included randomized controlled trials, cohort studies, case-control studies, and observational studies that focused on the relationship between vitamin D levels, *VDR* FokI gene polymorphism, and sepsis risk or mortality. (2) Patient demographics: Studies involving human subjects of any age, gender, or ethnicity were included. Priority was given to studies involving sepsis patients treated in an ICU setting. Neonates, adults, and elderly patients were considered as long as the data were relevant to the analysis. (3) Vitamin D and *VDR* FokI gene polymorphism: Studies measuring serum 25(OH)D levels and assessing the relationship between vitamin D deficiency and sepsis outcomes were included. Specifically, studies analyzing the *VDR* FokI gene polymorphism (rs2228570) were incorporated. Although other SNPs such as ApaI, BsmI, and TaqI were considered where available, the primary focus remained on FokI. (4) Variables to be collected: Eligible studies measured factors such as body mass index, age, gender, and comorbidities (e.g., diabetes or hypertension) that could influence vitamin D levels and sepsis outcomes. Data on sepsis severity (e.g., sequential organ failure assessment score) and inflammatory markers such as C-reactive protein were also considered. (5) Outcomes: Only studies that reported on the relationship between vitamin D levels, *VDR* polymorphisms, and sepsis mortality or morbidity were included. Studies providing quantitative data on the correlation between these factors and sepsis risk, hospital or ICU length of stay, or inflammatory markers were also considered.

The exclusion criteria for articles were (1) case reports and reviews: Case reports, narrative reviews, and meta-analyses were excluded because the focus was on original research providing primary data. (2) Non-human studies: Animal studies and studies that did not involve human participants were excluded. (3) Inadequate data: Studies that did not report on the relationship between vitamin D levels and *VDR* polymorphisms in sepsis patients or lacked sufficient data to perform a meta-analysis (e.g., no sample size, lack of control groups) were excluded. (4) Language and accessibility: Studies not published in English or those without full-text access were excluded. (5) Study design limitations: Articles that did not meet the inclusion criteria due to methodological issues, such as failure to control for important confounders or poor study quality (e.g., high risk of bias as per the Newcastle-Ottawa scale) were excluded from the meta-analysis.

Information Sources

A systematic literature review was conducted using electronic databases including PubMed, Embase, ProQuest, and the Cochrane Library for the period from 1996 until 2025, incorporating studies of varying quality. Keywords used in the search included "sepsis AND (Vitamin D OR 25(OH)D OR 25-Hydroxyvitamin D)" and "VDR FokI polymorphism OR Vitamin D receptor FokI gene variant."

Selection Process

The selection process began with a keyword search in various databases to identify relevant studies, followed by the removal of duplicate entries. Titles and abstracts were then screened, and full-text evaluations were conducted to determine eligibility. Studies containing relevant data were systematically analyzed and included in the meta-analysis, as depicted in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram (Figure 1).

Data Synthesis

Data from the selected studies were critically analyzed and synthesized to explore the relationship between vitamin D levels, *VDR* FokI gene polymorphism, and sepsis risk or mortality. The data were organized into 3 primary subtopics: (1) sepsis patients with *VDR* FokI gene polymorphism, (2) vitamin D levels in sepsis patients, and (3) the relationship between vitamin D levels and *VDR* FokI polymorphism in sepsis patients. The synthesis process followed these steps.

Data extraction

Two independent reviewers extracted data using a standardized form. Extracted data included study design, sample size, participant demographics (e.g., age, gender), vitamin D levels

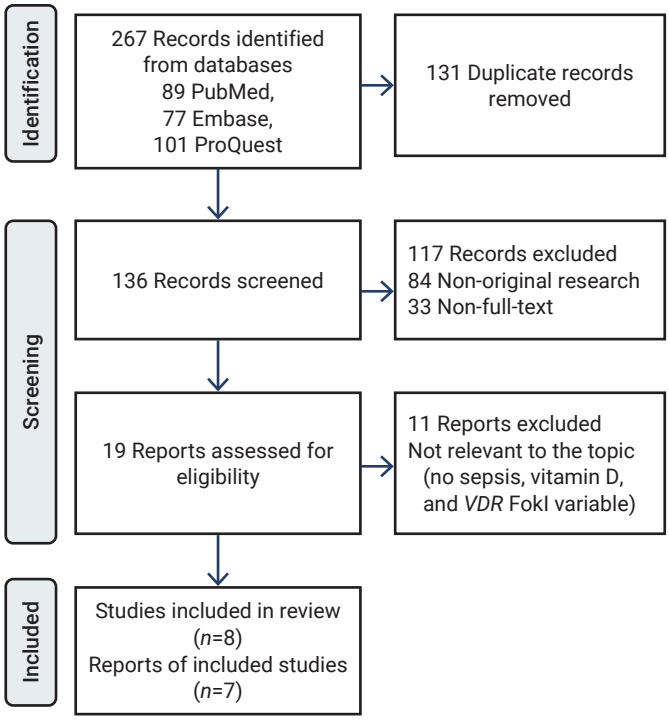


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow for systematic studies.

(25[OH]D), *VDR* FokI genotype (TT, CT, CC), and sepsis-related outcomes (e.g., sepsis risk, mortality). Discrepancies were resolved by consensus.

Statistical analysis

A meta-analysis was performed to quantitatively assess the relationship between vitamin D levels and sepsis outcomes. The pooled effect size for continuous variables (e.g., vitamin D levels) was calculated as the mean difference (MD) with 95% confidence intervals (CIs). The association between *VDR* FokI genotype (TT, CT, CC) and sepsis risk was analyzed using odds ratios (ORs) with 95% CIs, and the significance of pooled estimates was determined using the Z-statistic (p-value).

Heterogeneity assessment

Given the variability in study designs and populations, heterogeneity was evaluated using I^2 statistics. An I^2-value greater than 50% indicated substantial heterogeneity. A random-effects model was applied for studies with significant heterogeneity ($I^2 > 50\%$), while a fixed-effects model was used for studies with low heterogeneity.

Risk of bias

The risk of bias for each study was assessed using the ROB2 tool, which examines 5 domains: (1) randomization process, (2) deviations from intended intervention, (3) missing outcome data, (4) outcome measurement, and (5) selection of reported results. Studies with a high risk of bias were excluded from the analysis.

Sensitivity analysis

A sensitivity analysis was conducted by sequentially removing individual studies to determine if any single study disproportionately affected the overall results.

Publication bias

Publication bias was evaluated using funnel plots and the Egger test. A symmetrical funnel plot suggested no significant publication bias, whereas asymmetry indicated potential bias that could affect interpretation.

Statistical software

All analyses were performed using Review Manager (RevMan) software ver. 5.4 (Cochrane), with a significance threshold set at $p < 0.05$.

Additional statistical data

Where available, subgroup analyses based on patient demographics (e.g., age, gender, comorbidities) were performed to examine potential effect modifiers on the relationship between vitamin D and sepsis outcomes. Additionally, if data on vitamin D supplementation were available, its potential impact on sepsis-related outcomes in patients with vitamin D deficiency was analyzed.

Results

Study Selection

A total of 267 records were initially identified from various databases, including PubMed ($n = 89$), Embase ($n = 77$), ProQuest ($n = 101$), Cochrane Library ($n = 0$), and LILACS ($n = 0$). After screening, 136 records were deemed relevant for further evaluation. Following duplicate removal, 131 unique records remained. Further review excluded 117 records due to non-original research ($n = 84$) or lack of full-text availability ($n = 33$), leaving 19 reports for eligibility assessment. Of these, 11 reports were excluded because they did not focus on sepsis, vitamin D, or the *VDR* FokI variable. Ultimately, 8 studies were included in the review and meta-analysis. Specifically, 7 articles were excluded due to their focus on topics unrelated to sepsis—namely, Cafiero et al. [14] on periodontal disease; Gupta et al. [15] on tuberculosis disease–human immunodeficiency virus; Mokhtar et al. [16] on congenital heart disease; Ni et al. [17] on pregnancy; Rasheed et al. [18] on mycosis fungoides; Zununi Vahed et al. [19] on kidney transplant; and Zacharioudaki et al. [20] on

https://doi.org/10.24171/j.phrp.2025.0006

viral infection in infants—as well as Al-Daghri et al. [21] on cardiometabolic health in type 2 diabetes mellitus (T2DM), Kazemian et al. [22] on breast cancer, Mohseni et al. [23] on the impact of vitamin D supplementation on inflammatory markers and total antioxidant capacity in women with breast cancer, Neyestani et al. [24] in 2013 on the influence of *VDR* FokI genotypes on the effectiveness of vitamin D supplementation on glycemic, lipidemic, and inflammatory biomarkers in T2DM patients, and Sanwalka et al. [25] on the effect of *VDR* gene FokI polymorphism on bone mass accrual in Indian girls with low calcium intake. Ultimately, 7 relevant articles were identified and underwent data extraction and meta-analysis (Figure 1).

Risk of Bias in Studies

The risk of bias in the 8 reviewed articles was assessed using the Newcastle-Ottawa scale. The evaluation was conducted across 8 domains, with results presented in a table using red, yellow, and green symbols. All 7 articles demonstrated low bias in the 5 assessed domains, as indicated by green symbols (Figure 2) [11,26–32].

Vitamin D Deficiency in Sepsis Patients

Studies investigating the relationship between vitamin D levels and the FokI polymorphism in relation to sepsis risk and mortality have produced varying results. A study by Abouzeid et al. [26] in Egypt reported significantly lower vitamin D levels in sepsis cases compared to controls (17.5 ± 4.4 ng/mL vs. 36.7 ± 6.5 ng/mL). The FokI FF genotype was associated with a significantly increased risk of sepsis (OR, 3.7; 95% CI, 1.4–10.2) and mortality (OR, 3.7; 95% CI, 1.4–10.2). Tayel et al. [27] at Menoufia University Hospitals observed similar results in neonates, noting lower vitamin D levels in neonates with sepsis (8.7 ± 0.7 ng/mL vs. 19.1 ± 4.7 ng/mL) and an increased sepsis risk associated with the TT genotype (OR, 4.804; 95% CI, 1.4–16.4). In Serbia, Zeljic et al. [11] reported that the homozygous CC genotype increased sepsis risk, although no significant relationship with mortality was found. In contrast, Shaheen et al. [28] in Egypt observed no significant association between vitamin D levels, FokI polymorphism, and sepsis risk or mortality. Similarly, Das et al. [29] in India found no association between the FokI polymorphism and neonatal sepsis, despite lower vitamin D levels in cases compared to controls (12.23 ng/mL vs. 30.32 ng/mL) (Table 1) [11,26–32].

A study by Yang et al. [30] in China indicated that both the FokI FF and Ff genotypes, as well as the f allele, significantly increased sepsis risk, while Bozgul et al. [31] in Turkey found that the CC genotype was associated with a decreased risk of mortality (OR, 0.41; 95% CI, 0.17–0.92). Additionally, a study by

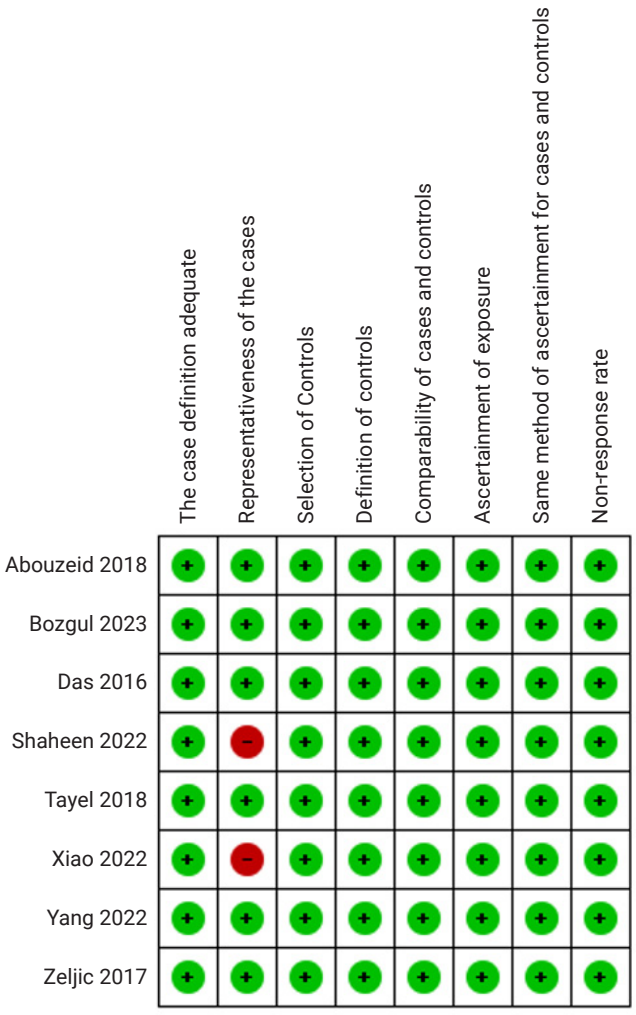


Figure 2. Risk of bias analysis using the Newcastle-Ottawa scale with 8 domains.

Xiao et al. [32] focused on these genetic variations and their potential to predict susceptibility to neonatal sepsis. This study is significant because it links genetic polymorphisms with serum biomarkers (such as 25(OH)D and sTREM-1), which could serve as diagnostic tools for neonatal sepsis.

The Relationship Between Sepsis Severity and Vitamin D Levels

Comparison of vitamin D levels between sepsis patients and healthy control groups

A meta-analysis comparing vitamin D levels between sepsis patients and healthy control groups found that vitamin D levels were lower in sepsis patients than in controls, with a MD of −4.17 (95% CI, −7.87 to −0.47) ng/mL. Although significant heterogeneity was observed among the studies (tau^2 = 17.75,

Table 1. Selected studies according to the title variables

Study	Year	Study		Sample		Exposure		Results		Bias (NOS)
		Location	Design	Age	Sample size	Vitamin D level (ng/mL)	FokI polymorphism	Sepsis risk (OR, 95% CI)	Mortality (OR, 95% CI)	
Abouzeid et al. [26]	2018	Zagazig, Ain-Shams, Cairo University, Egypt	Prospective multi-center case-control	6 mo–6 y	300 Cases, 300 controls	Case, 17.5±4.4; control, 36.7±6.5	CC: 21% (case), 6% (control), CT: 33% (case), 24% (control), TT: 46% (case), 70% (control)	CC, 3.7 (1.4–10.2); F, 1.5 (1.3–5.8)	CC, 3.7 (1.4–10.2); CT & TT not significant	8
Tayel et al. [27]	2018	Menoufia University Hospitals, Egypt	Cross-sectional study	Neonates (full-term, >37 wk)	80 Neonates with sepsis and 80 healthy neonates	Septic neonates, 8.7±0.7; healthy neonates, 19.1±4.7	TT genotype significantly more common in sepsis cases ($p=0.014$); T allele significantly more common in sepsis cases ($p=0.003$)	Sepsis risk: TT genotype, 4.804 (1.4–16.4); T allele, 2.786 (1.4–5.7)	7.5% Mortality rate (3/40 septic neonates)	8
Zeljic et al. [11]	2017	Military Medical Academy, Belgrade, Serbia	Observational	N/A	100 Patient with sepsis, 104 healthy controls	N/A	Homozygous CC (FokI rs228570) increases the risk of sepsis.	No significant difference was found between FokI polymorphism and mortality in the multivariate analysis	N/A	8
Shaheen et al. [28]	2022	Cairo University Children's Hospital	Observational	1 mo to 13 y	50 Cases, 100 controls	Mean, 29.23 (case); 20.72 (control)	TT (10%), CT (36%), CC (54%)	No significant association between vitamin D levels and the risk of sepsis	No significant association between vitamin D levels and mortality	7

(Continued on the next page)

https://doi.org/10.24171/j.phrp.2025.0006

Table 1. Continued

Study	Year	Study		Sample		Exposure		Results		Bias (NOS)
		Location	Design	Age	Sample size	Vitamin D level (ng/mL)	FokI polymorphism	Sepsis risk (OR, 95% CI)	Mortality (OR, 95% CI)	
Das et al. [29]	2016	Kalinga Institute of Medical Sciences, Bhubaneswar, India	Case-control study	Neonatal (newborns, gestational age over 37 wk)	120 Subjects (60 neonatal sepsis cases, 60 controls)	Mean serum 25(OH)D levels: case, 12.23; controls, 30.32	No association was found between the Fok1 genotype and neonatal sepsis or serum 25(OH)D levels. The distribution of Fok1, Bsm1, and Taq1 genotypes showed a significant deviation from Hardy-Weinberg equilibrium ($p < 0.01$).	N/A	N/A	8
Yang et al. [30]	2022	Yunnan Province, China	Observational, case-control	N/A	576 Sepsis patients, 421 healthy controls	Sepsis patients, 20.77±0.19; healthy controls, 24.64±0.36	CC,CT,TT	CCF, 1.77 (1.35–2.31); CT, 2.57 (1.47–4.56); T allele, 1.68 (1.36–2.08)	N/A	8
Bozgul et al. [31]	2023	İzmir, Turkey	Prospective, observational, single-center	63.3±19.0 y (sepsis), 57.4±19.8 y (non-sepsis)	96 Sepsis, 96 non-sepsis	N/A	CC,CT, TT	N/A	CC genotype, 0.41 (0.17–0.92)	8
Xiao et al. [32]	2022	Tianjin, China	Observational study analyzing genetic polymorphisms and sepsis susceptibility	Neonatal	150 Sepsis patients and 150 controls	Not directly provided in the results, but the study mentions serum 25(OH)D levels being lower in sepsis patients	The study analyzed rs739837 (*VDR* gene) and rs2234246 (*TREM-1* gene) polymorphisms.	Risk of sepsis in carriers of the G allele was 0.65 times that of the T allele (0.65 [0.50–0.83], $p < 0.001$).	The polymorphisms of VDR rs739837 and TREM-1 rs2234246 were not significantly correlated with survival in neonatal sepsis patients ($p > 0.05$)	7

OR, odds ratio; CI, confidence interval; NOS, Newcastle–Ottawa scale; N/A, not available; 25(OH)D, 25-hydroxyvitamin D.

https://doi.org/10.24171/j.phrp.2025.0006

I^2 = 100%), the overall effect was statistically significant, with a Z value of 2.21 and p = 0.03 (Figure 3) [26–30].

The distribution of *VDR* FokI genotypes (T/T, C/T, C/C) between sepsis patients and the control group

A meta-analysis of the risk associated with the CC c *VDR* FokI genotypes between sepsis patients and control groups revealed a higher risk in the sepsis group compared to controls, with a risk difference of 1.39 (95% CI, 0.70–2.73). However, the overall effect was not statistically significant, with a Z value of 0.95 and p = 0.34. The analysis also demonstrated significant heterogeneity (tau^2 = 0.81, I^2 = 91%), suggesting that these results should be interpreted with caution (Figure 4) [11,26–32].

For the CT genotype, the risk in the sepsis group was higher than in the control group, with a risk ratio of 0.75 (95% CI, 0.47–1.20). Despite a Z value of 1.19 and a reported p-value of 0.001, the overall effect was not statistically significant. Moderate heterogeneity was observed (tau^2 = 0.29, I^2 = 71%), which suggests that these results can be interpreted with somewhat greater confidence compared to previous analyses (Figure 5) [11,26–32].

The analysis of the TT genotype indicated that a lower risk in the sepsis group compared to controls, with a risk ratio of 1.04 (95% CI, 0.78–1.39). However, this difference was not statistically significant, as indicated by a Z value of 0.27 and p = 0.78. Significant heterogeneity was observed (tau^2 = 0.09, I^2 = 58%), warranting cautious interpretation of these results (Figure 6).

Discussion

We assessed the risk of bias for each study included in the meta-analysis using the Newcastle-Ottawa scale, which evaluates studies across 5 key domains. Each domain was categorized as low risk, high risk, or unclear risk. This evaluation allowed us to assess the credibility and quality of the studies in our systematic review and meta-analysis. The scale examines aspects based on predefined criteria such as the adequacy of the case definition, representativeness of the cases, selection and definition of controls, comparability of cases and controls, ascertainment of exposure, the

	Case			Control				SMD	SMD
Study or subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, random, 95% CI	IV, random, 95% CI
Yang 2022	20.77	0.19	576	24.64	0.36	421	20.0%	−14.07 [−14.70, −13.44]	
Tayel 2018	21.4	4.8	40	36.9	4.2	40	19.9%	−3.40 [−4.10, −2.71]	
Shaheen 2022	29.23	31.47	50	20.72	9.65	100	20.0%	0.43 [0.09, 0.77]	
Das 2016	12.23	36.32	60	30.32	52.76	60	20.0%	−0.40 [−0.76, −0.04]	
Abouzeid 2018	17.5	4.4	300	36.7	6.5	300	20.1%	−3.45 [−3.71, −3.20]	
Total (95% CI)			**1,026**			**921**	**100.0%**	**−4.17 [−7.87, −0.47]**	

Heterogeneity: tau^2=17.75; chi^2=1,758.49, df=4 (p<0.00001); I^2=100%
Test for overall effect: Z=2.21 (p=0.03)

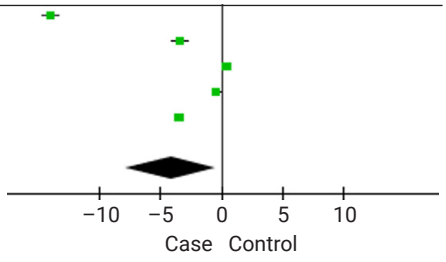


Figure 3. Forest plots comparing vitamin D levels between cases and controls.
SD, standard deviation; SMD, standard mean difference; IV, intravenous; CI, confidence interval; df, degrees of freedom.

	Case		Control			Odds ratio	Odds ratio
Study or subgroup	Events	Total	Events	Total	Weight	M-H, random, 95% CI	M-H, random, 95% CI
Abouzeid 2018	63	300	18	300	13.5%	4.16 [2.40, 7.23]	
Bozgul 2023	40	96	43	96	13.4%	0.88 [0.50, 1.56]	
Das 2016	6	10	12	20	8.3%	1.00 [0.21, 4.71]	
Shaheen 2022	27	50	44	100	12.9%	1.49 [0.76, 2.96]	
Tayel 2018	7	40	13	40	10.9%	0.44 [0.15, 1.26]	
Xiao 2022	110	150	81	150	13.8%	2.34 [1.44, 3.80]	
Yang 2022	288	576	274	421	14.5%	0.54 [0.41, 0.69]	
Zeljic 2017	34	100	13	104	12.7%	3.61 [1.77, 7.36]	
Total (95% CI)		**1,322**		**1,231**	**100.0%**	**1.39 [0.70, 2.73]**	
Total events	575		498				

Heterogeneity: tau^2=0.81; chi^2=76.12, df=7 (p<0.00001); I^2=91%
Test for overall effect: Z=0.95 (p=0.34)

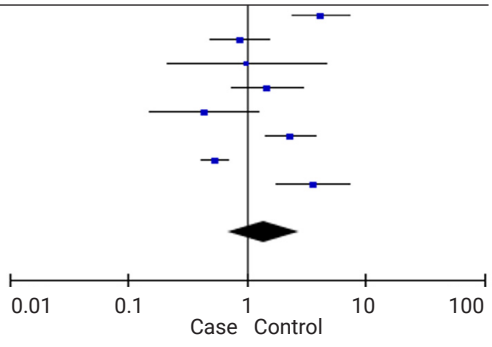


Figure 4. Forest plot comparing the distribution of the CC *VDR* FokI genotype.
M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

https://doi.org/10.24171/j.phrp.2025.0006

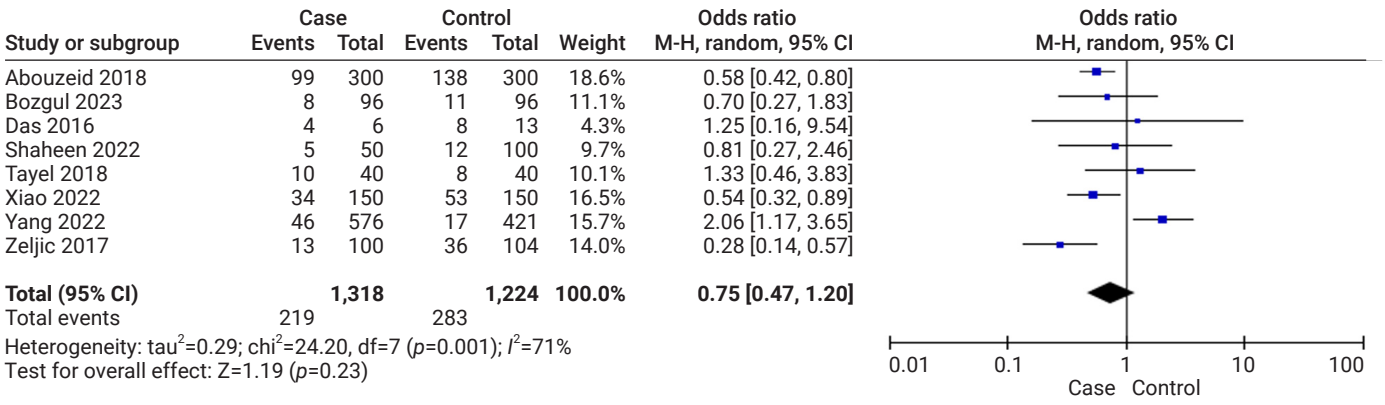


Figure 5. Forest plot comparing the distribution of the CT *VDR* FokI genotype.
M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

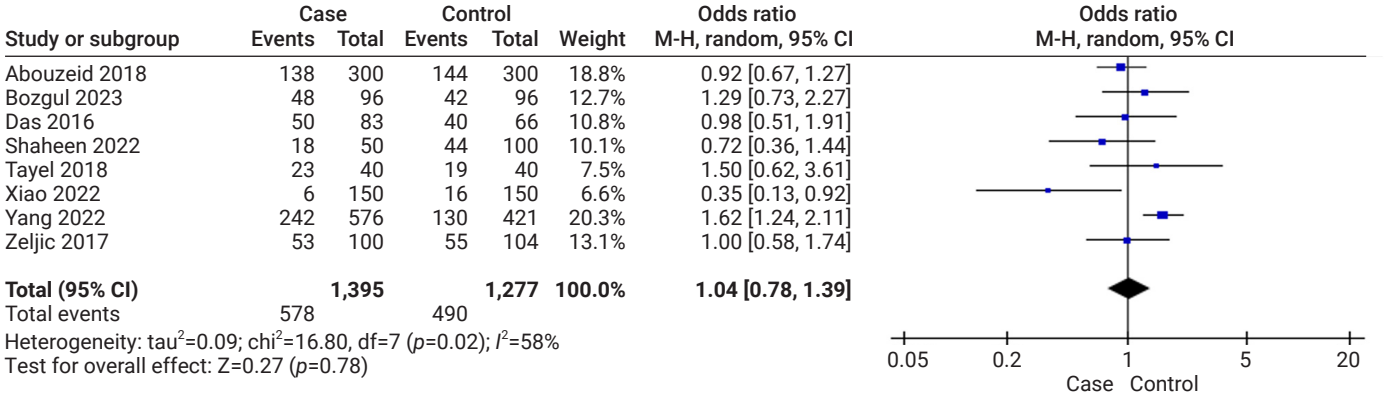


Figure 6. Forest plot comparing the distribution of the TT *VDR* FokI genotype.
M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

consistency of ascertainment methods for cases and controls, and the non-response rate. Overall, the risk of bias was primarily low, although a few studies raised concerns in specific domains—such as unclear randomization or missing outcome data—which may have impacted the results despite their inclusion based on eligibility criteria.

The results of this study indicate that vitamin D levels in sepsis patients are consistently lower than in the control group. This finding aligns with the established role of vitamin D as an immunomodulator that contributes to the body's immune response. Vitamin D enhances the activity of innate immune cells, including macrophages and neutrophils, while reducing the release of pro-inflammatory cytokines. Consequently, lower vitamin D levels may impair the body's ability to combat infections, thereby increasing the risk of sepsis.

The FokI polymorphism in the *VDR* gene is a functional variant rather than a silent one. Occurring in the translation region of the *VDR* gene, it leads to the production of 2 different forms of the VDR protein. The "F" allele produces a longer VDR protein, whereas the "f" allele results in a shorter form. This functional difference is believed to affect the receptor's ability to bind vitamin D, thereby influencing the immune system's response. Studies suggest that this polymorphism may alter VDR activity, affecting the body's ability to modulate immune responses and potentially increasing susceptibility to conditions such as sepsis. Therefore, the FokI polymorphism is considered functional because it can modify receptor activity and the immune system's responsiveness to vitamin D.

This study also indicates that *VDR* gene polymorphisms, particularly FokI, may be associated with sepsis risk, although there is variability in the results across studies. For example, some studies—such as those by Abouzeid et al. [26] and Yang et al. [30]—found that the FF genotype significantly increased the risk of sepsis. In contrast, other studies, including those by Das et al. [29] and Shaheen et al. [28], did not observe a significant association. This

variation could be attributed to differences in population characteristics, analytical methods, and sample sizes used in each study.

The meta-analysis results support the finding that vitamin D levels are lower in sepsis patients compared to the control group (standard MD, −4.17; 95% CI, −7.87 to −0.47) (Figure 3) [26–30]. The potential role of vitamin D in the pathophysiology of sepsis is further underscored by observations that patients with severe sepsis exhibit significantly lower levels of vitamin D compared to those with mild sepsis or healthy individuals [33,34]. Vitamin D plays a critical immunomodulatory role in both innate and adaptive immunity by enhancing the production of antimicrobial peptides (such as cathelicidin and defensins), suppressing pro-inflammatory cytokines (tumor necrosis factor-α and interleukin [IL]-6), and increasing anti-inflammatory cytokines (IL-10) [35]. In sepsis, dysregulation of the immune response—marked by hyper-inflammatory reactions and immunosuppression—is a key factor in morbidity and mortality. Consequently, low vitamin D levels may exacerbate this condition by reducing the body's capacity to fight infections [33,36,37]. Several mechanisms may explain the reduced vitamin D levels in sepsis patients, including pre-existing vitamin D deficiency due to comorbidities like kidney or liver disease, increased utilization of vitamin D during the immune response, and organ dysfunction that disrupts the conversion of vitamin D to its active form [37,38].

Although some studies found an association between the FokI polymorphism and sepsis risk, our meta-analysis shows that the risk associated with the genotype is not statistically significant (CC: p-value = 0.53; CT: p-value = 0.46; TT: p-value = 0.3) (Figures 4–6). In this regard, Lu et al. [39] also supported the view that the relationship between FokI and sepsis risk is not significant—particularly when FokI is considered among genetic variants that do not meet strong or moderate credibility criteria. The analysis also revealed significant heterogeneity, indicating that these results should be interpreted with caution. This variability may result from local genetic factors, differing mutation patterns in the *VDR* gene across populations, or the interaction between genetic and environmental influences.

The relationship between vitamin D levels and susceptibility to sepsis may also be influenced by genetic factors, including variations in the gene encoding VDR [11,30,40]. Among these, the FokI polymorphism is one of the most extensively studied, as it is believed to affect VDR function and, in turn, the body's immune response. However, although low vitamin D levels have been consistently identified as a risk factor that exacerbates sepsis, the evidence regarding a direct relationship between the FokI polymorphism and sepsis risk remains inconclusive [28,30,40].

Additionally, the findings of this study have important implications for the management of patients at risk for sepsis. Detecting low vitamin D levels and identifying *VDR* gene polymorphisms may serve as useful preventive or diagnostic measures to improve sepsis management. Vitamin D supplementation in high-risk populations, such as neonates or patients with immune deficiencies, could represent a potential therapeutic approach, although further research is needed to substantiate this strategy.

However, several limitations must be acknowledged. The high heterogeneity observed suggests that methodological variations, sample sizes, and population characteristics have influenced the results. Furthermore, not all studies provided detailed information on parameters that might affect outcomes—such as nutritional status, comorbidities, or medications received by the patients. Therefore, further research using prospective designs and better control of confounding factors is warranted.

Several limitations must be considered when interpreting this study's results. First, high heterogeneity was observed across the studies included in the meta-analysis, indicating variability in study designs, populations, and measurement methods. This heterogeneity suggests that the findings may not be universally applicable and that the results should be interpreted with caution. Factors such as methodological differences (e.g., varying protocols for measuring vitamin D levels, differences in the timing of these measurements, and diverse inclusion criteria for sepsis patients) likely contribute to this variability. Second, some studies had relatively small sample sizes, which may limit both the statistical power and the generalizability of the results. While larger studies could help clarify these trends, small sample sizes might lead to overestimated or underestimated effects. Another important limitation is the insufficient reporting of potential confounding variables—such as comorbid conditions (e.g., diabetes, obesity), medications, and nutritional status—that could affect both vitamin D levels and sepsis outcomes. Without controlling for these factors, drawing definitive conclusions about the causal relationship between vitamin D deficiency and sepsis risk is challenging. Lastly, variability in the assessment of *VDR* polymorphisms (i.e., differences in reporting genotypic distributions and a limited focus on the FokI polymorphism) may have influenced the consistency of the results. It is also important to note that the relationship between the FokI polymorphism and sepsis risk remains inconclusive, and further studies with more standardized genotyping methods are necessary to confirm these findings. Overall, while this study provides valuable insights into the potential role of vitamin D and *VDR* polymorphisms in

https://doi.org/10.24171/j.phrp.2025.0006

sepsis, future research should address these limitations by incorporating larger, more consistent datasets, accounting for potential confounders, and using more standardized methodologies.

Conclusion

This study demonstrates that vitamin D levels in sepsis patients are significantly lower than those in the control group, underscoring the potential role of vitamin D as a risk factor in the development of sepsis. Moreover, although several studies have identified an association between *VDR* gene polymorphisms—particularly FokI—and an increased risk of sepsis, the meta-analysis did not reveal a statistically significant relationship, with high heterogeneity observed across the studies.

Supplementary Material

Supplementary Material 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. Supplementary data are available at https://doi.org/10.24171/j.phrp.2025.0006.

Notes

Ethics Approval
Not applicable.

Conflicts of Interest
The authors have no conflicts of interest to declare.

Funding
None.

Availability of Data
All data generated or analyzed during this study are included in this published article. For other data, these may be requested through the corresponding author.

Authors' Contributions
Conceptualization: all authors; Data curation: WJ; Formal analysis: all authors; Funding acquisition: WJ; Investigation: WJ; Methodology: WJ; Project administration: all authors; Resources: all authors; Software: all authors; Supervision: SRT, HS, AAA; Validation: SRT, HS, AAA; Visualization: WJ; Writing–original draft: WJ; Writing–review & editing: all authors. All authors read and approved the final manuscript.

References

1. Shankar-Hari M, Phillips GS, Levy ML, et al. Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 2016;315:775–87.
2. Martin-Loeches I, Guia MC, Vallecoccia MS, et al. Risk factors for mortality in elderly and very elderly critically ill patients with sepsis: a prospective, observational, multicenter cohort study. Ann Intensive Care 2019;9:26.
3. Jayara A, Mascarenhas J, Gandhi B, et al. Comparison of trends of procalcitonin and neutrophil to lymphocyte ratio in patients of sepsis in intensive care unit. Indian J Crit Care Med 2024;28:942–51.
4. Lai YH, Fang TC. The pleiotropic effect of vitamin D. ISRN Nephrol 2013;2013:898125.
5. Braun AB, Gibbons FK, Litonjua AA, et al. Low serum 25-hydroxyvitamin D at critical care initiation is associated with increased mortality. Crit Care Med 2012;40:63–72.
6. Iqhrammullah M, Duta TF, Alina M, et al. Role of lowered level of serum vitamin D on diabetic foot ulcer and its possible pathomechanism: a systematic review, meta-analysis, and meta-regression. Diabet Epidemiol Manag 2024;13:100175.
7. Putra MI, Gusti N, Duta TF, et al. Vitamin D supplementation improves foot ulcers among diabetic patients: pooled analysis of randomized controlled trials. Narra X 2023;1:e104.
8. Jeng L, Yamshchikov AV, Judd SE, et al. Alterations in vitamin D status and anti-microbial peptide levels in patients in the intensive care unit with sepsis. J Transl Med 2009;7:28.
9. Saminan S, Gusti N, Amirah S, et al. Role of polymorphism in vitamin D receptor gene on diabetic foot ulcer: a systematic review and meta-analysis. J Pharm Pharmacogn Res 2024;12:348–362.
10. Voltan G, Cannito M, Ferrarese M, et al. Vitamin D: an overview of gene regulation, ranging from metabolism to genomic effects. Genes (Basel) 2023;14:1691.
11. Zeljic K, Elkilany A, Supic G, et al. Vitamin D receptor gene polymorphisms association with the risk of sepsis and mortality. Int J Immunogenet 2017;44:129–34.
12. Ghosh A, Prabhu M, Prabhu K, et al. Vitamin D as a biomarker in predicting sepsis outcome at a tertiary care hospital. Asian J Med Sci 2023;14:89–94.
13. Cecchi A, Bonizzoli M, Douar S, et al. Vitamin D deficiency in septic patients at ICU admission is not a mortality predictor. Minerva Anestesiol 2011;77:1184–9.
14. Cafiero C, Grippaudo C, Dell'Aquila M, et al. Association between Vitamin D receptor gene polymorphisms and periodontal bacteria: a clinical pilot study. Biomolecules 2022;12:833.
15. Gupta A, Montepiedra G, Gupte A, et al. Low vitamin-D levels combined with PKP3-SIGIRR-TMEM16J host variants is associated with tuberculosis and death in HIV-infected and -exposed infants. PLoS One 2016;11:e0148649.
16. Mokhtar WA, Fawzy A, Allam RM, et al. Maternal vitamin D level and vitamin D receptor gene polymorphism as a risk factor for congenital heart diseases in offspring; an Egyptian case-control study. Genes Dis 2018;6:193–200.
17. Ni M, Zhang Q, Zhao J, et al. Relationship between maternal vitamin D status in the first trimester of pregnancy and maternal and neonatal outcomes: a retrospective single center study. BMC Pediatr

2021;21:330.

18. Rasheed H, Hegazy RA, Gawdat HI, et al. Serum vitamin D and vitamin D receptor gene polymorphism in mycosis fungoides patients: a case control study. PLoS One 2016;11:e0158014.
19. Zununi Vahed S, Ahmadian E, Foroughi P, et al. Vitamin D receptor and vitamin D binding protein gene polymorphisms are associated with renal allograft outcome. Nutrients 2021;13:1101.
20. Zacharioudaki M, Messaritakis I, Galanakis E. Vitamin D receptor, vitamin D binding protein and CYP27B1 single nucleotide polymorphisms and susceptibility to viral infections in infants. Sci Rep 2021;11:13835.
21. Al-Daghri NM, Mohammed AK, Al-Attas OS, et al. Vitamin D receptor gene polymorphisms modify cardiometabolic response to vitamin D supplementation in T2DM patients. Sci Rep 2017;7:8280.
22. Kazemian E, Akbari ME, Moradi N, et al. Effect of vitamin D receptor polymorphisms on plasma oxidative stress and apoptotic biomarkers among breast cancer survivors supplemented vitamin D3. Eur J Cancer Prev 2020;29:433–44.
23. Mohseni H, Amani R, Hosseini SA, et al. Genetic variations in VDR could modulate the efficacy of vitamin D3 supplementation on inflammatory markers and total antioxidant capacity among breast cancer women: a randomized double blind controlled trial. Asian Pac J Cancer Prev 2019;20:2065–72.
24. Neyestani TR, Djazayery A, Shab-Bidar S, et al. Vitamin D receptor Fok-I polymorphism modulates diabetic host response to vitamin D intake: need for a nutrigenetic approach. Diabetes Care 2013;36:550–6.
25. Sanwalka N, Khadilkar A, Chiplonkar S, et al. Influence of vitamin D receptor gene Fok1 polymorphism on bone mass accrual post calcium and vitamin D supplementation. Indian J Pediatr 2015;82:985–90.
26. Abouzeid H, Abdelaal NM, Abdou MA, et al. Association of vitamin D receptor gene FokI polymorphism and susceptibility to CAP in Egyptian children: a multicenter study. Pediatr Res 2018;84:639–44.
27. Tayel SI, Soliman SE, Elsayed HM. Vitamin D deficiency and vitamin D receptor variants in mothers and their neonates are risk factors for neonatal sepsis. Steroids 2018;134:37–42.
28. Shaheen I, Afifi R, Abulata N, et al. Vitamin D receptor polymorphisms and vitamin D insufficiency are not associated with sepsis in critically ill children: a case-control study. Gaz Egypt Paediatr Assoc 2022;70:38.
29. Das B, Patra S, Behera C, et al. Genotyping of vitamin D receptor gene polymorphisms using mismatched amplification mutation assay in neonatal sepsis patients of Odisha, eastern India. Infect Genet Evol 2016;45:40–7.
30. Yang X, Ru J, Li Z, et al. Lower vitamin D levels and VDR FokI variants are associated with susceptibility to sepsis: a hospital-based case-control study. Biomarkers 2022;27:188–95.
31. Bozgul SM, Emecen DA, Akarca FK, et al. Association between vitamin D receptor gene FokI polymorphism and mortality in patients with sepsis. Mol Biol Rep 2023;51:44.
32. Xiao L, Que S, Mu L, et al. The relationship between vitamin D receptor gene and TREM-1 gene polymorphisms and the susceptibility and prognosis of neonatal sepsis. J Clin Lab Anal 2022;36:e24405.
33. Olejarova M, Dobisova A, Suchankova M, et al. Vitamin D deficiency: a potential risk factor for sepsis development, correlation with inflammatory markers, SOFA score and higher early mortality risk in sepsis. Bratisl Lek Listy 2019;120:284–90.
34. Parekh D, Patel JM, Scott A, et al. Vitamin D deficiency in human and murine sepsis. Crit Care Med 2017;45:282–9.
35. Watkins RR, Yamshchikov AV, Lemonovich TL, et al. The role of vitamin D deficiency in sepsis and potential therapeutic implications. J Infect 2011;63:321–6.
36. de Haan K, Groeneveld AB, de Geus HR, et al. Vitamin D deficiency as a risk factor for infection, sepsis and mortality in the critically ill: systematic review and meta-analysis. Crit Care 2014;18:660.
37. Ginde AA, Camargo CA, Shapiro NI. Vitamin D insufficiency and sepsis severity in emergency department patients with suspected infection. Acad Emerg Med 2011;18:551–4.
38. Moromizato T, Litonjua AA, Braun AB, et al. Association of low serum 25-hydroxyvitamin D levels and sepsis in the critically ill. Crit Care Med 2014;42:97–107.
39. Lu H, Wen D, Wang X, et al. Host genetic variants in sepsis risk: a field synopsis and meta-analysis. Crit Care 2019;23:26.
40. Li Q, Li W, Chen M, et al. Association of vitamin D receptor gene polymorphism with the risk of sepsis: a systematic review and meta-analysis. Medicine (Baltimore) 2023;102:e35130.

https://doi.org/10.24171/j.phrp.2025.0006

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0019
Osong Public Health Res Perspect 2025;16(3):223-235

Review Article

Effect of cognitive behavior therapy on abstinence among healthy smokers: a systematic review and meta-analysis of randomized controlled trials

Rajesh Kumar[1], Priyanka Malhotra[1], Vishwas AS[1], Kalpana Beniwal[2], Gopal Singh Charan[3]

[1]College of Nursing, All India Institute of Medical Sciences, Rishikesh, India
[2]Department of Nursing, All India Institute of Medical Sciences, Rishikesh, India
[3]Institute of Nursing, Maharishi Markandeshwar University, Ambala, India

ABSTRACT

Objectives: The objectives of this review and meta-analysis were twofold: first, to critically evaluate the effectiveness of cognitive behavioral therapy (CBT)-based interventions relative to standard care or control conditions in promoting smoking abstinence at the end of treatment, and second, to determine abstinence rates at 3 and 6 months of follow-up.

Methods: A comprehensive search of electronic databases, including PubMed, Cochrane Library, PsycINFO, Embase, and ClinicalTrials.gov, was conducted for randomized controlled trials published from 2001 to September 2024. Studies evaluating the effect of CBT on abstinence rates among healthy smokers (aged ≥12 years) were included and analyzed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

Results: Out of 1,514 study records screened, 7 studies (comprising 17 arms and 1,438 participants) met eligibility criteria for inclusion in the final analysis. The pooled analysis revealed that CBT significantly improved smoking abstinence rates, demonstrating a fourfold increase in effectiveness compared to controls (odds ratio [OR], 4.04; 95% confidence interval [CI], 2.73–5.99; $I^2 = 19\%$; $p < 0.00001$). Additionally, CBT significantly impacted smoking cessation at 3-month follow-up (OR, 1.62; 95% CI, 1.11–2.38; $I^2 = 0\%$; $p = 0.01$) and 6-month follow-up (OR, 2.19; 95% CI, 1.59–3.00; $I^2 = 0\%$; $p < 0.00001$).

Conclusion: CBT has demonstrated efficacy in facilitating smoking abstinence, particularly immediately after treatment, with sustained but diminished effects over time. However, robust conclusions on the efficacy of CBT require further studies involving larger sample sizes, diverse geographical regions, and longer follow-up periods.

Keywords: Cognitive behavior therapy; Smokers; Smoking abstinence; Smoking cessation

Received: January 22, 2025
Revised: March 1, 2025
Accepted: April 24, 2025

Corresponding author:
Rajesh Kumar
College of Nursing, All India Institute of Medical Sciences, Virbhadra Marg, Rishikesh, Uttarakhand 249203, India
E-mail: rajeshrak61@gmail.com

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Tobacco smoking remains a significant public health challenge, contributing to substantial morbidity, mortality, and economic burden worldwide [1]. The long-term use of pharmacological interventions has been associated with adverse outcomes on both physical and psychological health in tobacco users [2,3]. Tobacco smoking is linked to multiple diseases, including cardiovascular diseases (particularly stroke and coronary heart disease), cancer, chronic obstructive pulmonary disease [4], periodontal health conditions [5], and many other metabolic diseases [6]. Additionally, tobacco use is a contributory risk factor for various infections, bone-related conditions, reproductive issues such as premature ejaculation [7], delayed wound healing [8], and psychological disorders [8]. It is well-established that tobacco-related complications are associated with increased premature mortality among smokers [9].

Effective cessation strategies and interventions are urgently required to mitigate the ripple effects of tobacco use. Various interventions and strategies have been tested, yielding variable efficacy and success rates [10]. Smoking cessation interventions may involve pharmacological, non-pharmacological, or behavioral approaches [11], applied either as standalone therapies or in combination across different age groups [12].

Considering concerns regarding side effects, prolonged use [13], and the high costs associated with long-term pharmacological treatments, there has been an increasing shift toward non-pharmacological treatments [14]. Non-pharmacological interventions include behavioral skills training using cognitive behavioral therapy (CBT) [15], incentive-based methods, reinforcement behavior techniques [15–17], and motivational and brief interviewing approaches [18,19].

CBT is a structured umbrella psychological intervention that has shown promise in addressing tobacco cessation [20]. This approach focuses on identifying and modifying maladaptive thought patterns and dysfunctional behaviors [21]. Furthermore, CBT aids individuals in developing alternative thinking and behaviors to reduce psychological stress and prevent relapse episodes [22].

CBT is typically delivered through multiple sessions of 45 to 60 minutes, emphasizing problem-solving skills, recognition of social or environmental cues that trigger tobacco use, and education about the quitting process. Additionally, CBT teaches adaptive coping strategies to mitigate relapse risk [20,23]. Thus, CBT represents a more intensive intervention than brief counseling alone and is associated with higher quit rates among smokers [24]. CBT delivery formats include individual face-to-face sessions [25], group therapy [26], and digital platforms [27]. Although previous literature has demonstrated CBT's effectiveness in smoking cessation, uncertainty remains regarding its efficacy in maintaining abstinence among tobacco users, particularly as a standalone treatment over longer follow-up periods. Clarifying this issue could inform the design of more effective strategies for early smoking cessation across diverse age groups. Additionally, the comparative effectiveness of CBT versus standard care or minimal interventions for smoking cessation remains uncertain and requires pooled data analysis from randomized controlled trials (RCTs).

This systematic review and meta-analysis aimed to quantify the impact of CBT interventions on smoking cessation by synthesizing evidence from RCTs. This review focused on abstinence rates at the end of treatment and at 3-month and 6-month follow-ups.

Therefore, considering the substantial individual studies available and recent advancements in intervention methods, our objective is to provide a comprehensive pooled analysis regarding the effectiveness of CBT in achieving abstinence among healthy smokers at the specified follow-up intervals (3 and 6 months).

HIGHLIGHTS

- Cognitive behavioral therapy (CBT) is an empirically validated intervention consistently demonstrating strong efficacy in supporting sustained smoking cessation.
- This investigation offers novel insights into the longitudinal effectiveness of CBT for reinforcing adherence to smoking cessation.
- Integrating CBT as an adjunct to standard therapeutic strategies has potential as a transformative intervention for smokers attempting cessation.

Materials and Methods

Eligibility Criteria

The review included studies of sufficient quality that assessed abstinence rates at the end of treatment, as well as at 3-month and 6-month follow-up intervals.

This review specifically encompassed publications employing RCT or clinical trial designs, focusing on smoking cessation, treatment outcomes, or smoking quit rates. Included studies utilized CBT, either independently or in combination with other interventions, for smoking cessation or quitting among healthy adolescent or adult participants

https://doi.org/10.24171/j.phrp.2025.0019

aged 12 years and older. The inclusion criteria were restricted to research published in English between January 2001 and September 2024 and available electronically in full text.

We excluded studies that implemented internet-based, web-based, or mobile app interventions, as well as those that employed treatment conditions other than placebo or standard care in the control group. Additionally, reviews, meta-analyses, qualitative studies, non-RCTs, editorials, observational studies, studies focusing exclusively on pharmacotherapy, and studies in which smoking cessation was not a primary or secondary outcome were excluded. Rayyan software was used to assist in database screening according to the inclusion criteria [28]. All authors reviewed the articles, and those not meeting the inclusion criteria were excluded. Cross-referencing of both excluded and selected studies was performed with co-authors' consensus.

Information Sources

The search for this review was conducted between September 2024 and November 2024. A comprehensive and systematic literature search was carried out across major electronic databases, including PubMed, Cochrane Library, PsycINFO, Embase, and ClinicalTrials.gov.

Search Strategy

The search was performed using major online databases, refined by filters for the date range (January 2001 to September 2024) and limited to RCTs and controlled clinical trials published in English. The primary keywords and phrases used to identify relevant studies included: "cognitive behavioral therapy," OR "cognitive behavioral treatment," OR "CBT," AND "smoking cessation," OR "tobacco cessation," OR "quit smoking," AND "randomized controlled trials," OR "clinical controlled trial," AND "smoker," OR "tobacco user(s)," along with other related MeSH terms. The most recent included study was published on February 27, 2024 [29]. Additionally, 2 reviewers (P.M., R.K.) thoroughly screened references from selected studies to locate studies potentially missed during the primary search. Complete details of the search strategy are provided in Supplementary Material 1.

Extracted studies underwent a screening process based on titles, followed by an evaluation of independent and outcome variables. This meta-analysis adhered to the PICOTS framework, defined as follows: population/patient (P)—healthy smokers aged 12 years and above; intervention/indicator (I)—CBT; comparator/control (C)—standard care, consultation, health education, or brief advice; outcome (O)—smoking cessation measured at 3-month and 6-month follow-up intervals; time (T)—studies published from January 2001 to September 2024; and study design (S)—RCTs assessing the effectiveness of CBT compared with control or standard treatment conditions (Table 1).

Intervention

The interventions analyzed included various forms of CBT aimed at assessing their effects on the point-prevalence of smoking cessation among healthy individuals. Out of 7 studies employing cognitive behavioral approaches, 5 administered standard CBT [29–33], while 1 adopted a self-help CBT approach [34], and 1 used a smokeless smoking reduction program consisting of 4 weekly sessions of 2.5 hours each [35].

The interventions were delivered by researchers through group sessions with varying participant sizes: groups of 3 to 12 participants [34], 8 to 9 participants [29], 10 participants [31], and 7 (5–9) participants [35]. Sessions ranged in duration from 15 to 30 minutes [33,35] to 60 to 90 minutes [29–31], up to 2.5 hours [35], with sessions occurring weekly or twice weekly over periods ranging from 1 week to 52 weeks. Some studies provided telephone consultations to individual participants during weeks 2 and 5 of the program [35], and 1 study delivered individual counseling to the participants at baseline [33].

Study Selection Process

The outcomes assessed were smoking abstinence events at the end of treatment, as well as at 3 and 6 months among healthy smokers. Two authors (R.K. and P.M.) thoroughly reviewed all studies to gather information on the assessment measures employed. A consensus was reached to include

Table 1. The PICOTS criteria for the included studies

PICOTS	Criteria
Population (P)	Young adult (19–24), adult (19–44), older (65–79), healthy smokers
Intervention (I)	Cognitive behavioral therapy
Comparison (C)	Placebo, health education, standard treatment, hypnotherapy, brief advice, & consultation
Outcomes (O)	Abstinence at the end of treatment, 3-month abstinence rate, and 6-month abstinence rate
Time (T)	Published between January 2001 to September 2024
Study design (S)	Randomized controlled trials

only the outcomes assessed by the most frequently used instruments in this systematic review and meta-analysis. Additionally, studies reporting smoking abstinence data in frequencies or events were included. Studies reporting outcome measures in means or other formats were excluded. All outcome assessments utilized standardized tools (Table 1). Most studies conducted follow-up assessments at 1, 3, and 6 months [27,29,30], 2 studies had follow-ups at 3 and 6 months [31,34], and 3 studies had follow-ups at 1, 3, 6, and 12 months [32,33].

After exporting records from databases into Microsoft Excel, duplicate entries were removed. Subsequently, guided by eligibility criteria, 2 reviewers independently evaluated titles and abstracts of each record to identify potentially eligible studies. Following this screening, full texts were retrieved and thoroughly examined for final inclusion in the analyses. Two researchers (R.K. and K.B.) analyzed and extracted data from relevant articles. One researcher (P.M.) conducted data extraction, while another (R.K.) verified selected studies to ensure alignment with objectives and inclusion criteria. In case of discrepancies, 2 additional reviewers (V.A.S. and K.B.) were consulted to resolve disagreements. Ultimately, a senior reviewer (G.S.C.) provided a final opinion before the inclusion of records in the final analyses.

Data Collection and Processing

A Google Sheet was used to extract information, ensuring the consistency and accuracy of the collected data. Information from the studies was organized into tabular format and categorized according to the author's surname, study design, country, target participants, presence of sample size estimation (yes/no), actual sample size, age (in years), gender distribution, types of intervention and control groups, duration of the intervention, group size, follow-up intervals (days per weeks), and outcome assessment measures (Table 2) [29–35].

Assessment of Study Risk of Bias

The risk of bias in the included studies was evaluated using the Cochrane review criteria [36]. Each study was assessed across 6 evidence-based domains: allocation concealment and random sequence generation (both addressing selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), and selective reporting (reporting bias). The risk of bias for each domain was categorized as low, high, or unclear. In some instances, corresponding authors were contacted to provide missing information or clarify findings. If uncertainties remained after receiving responses, decisions were made through consensus among authors. Detailed methodological findings related to the Cochrane risk assessment domains are presented in Table 3 [29–35].

Data Analysis

Pooled data were statistically analyzed using Cochrane Review Manager software (RevMan ver. 5.3; Cochrane) [37]. In all analyses, heterogeneity was evaluated using the I^2 statistical approach, which quantifies the degree of inconsistency among study results. A fixed-effect model was employed to analyze findings in this meta-analysis when heterogeneity was less than 50%, minimizing heterogeneity risk. This model was selected based on the criterion of a p-value less than 0.05 and an I^2 value lower than 50%. The odds ratio (OR), with a 95% confidence interval (CI), was calculated for the abstinence rate outcomes. The authors included abstinence rates at 2 follow-up intervals—3 months and 6 months—in this meta-analysis, considering inconsistencies and variations across studies to mitigate potential bias in the meta-analytic outcomes.

Results

Study Selection

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [38] (Figure 1) guided the research process for study identification, selection, eligibility assessment, and final inclusion. Initially, 1,514 articles were identified, of which 1,013 were excluded due to duplication. Application of predefined inclusion and exclusion criteria to titles and abstracts reduced the number of eligible studies to 455. After comprehensive full-text evaluations, 7 articles met the criteria and were included in the final meta-analysis (Supplementary Material 2).

Study Characteristics

The study characteristics are summarized in Table 1. Seven studies involving a total of 1,438 participants were analyzed, comprising 769 participants in intervention groups and 671 in control groups, with follow-up durations ranging from 1 to 12 months at various intervals. Six studies delivered interventions through a face-to-face approach, while 1 study [35] used a digital follow-up. Comparators included a health promotion intervention [34], a brief consultation along with waitlist control [35], very brief advice [27], health education [30–32], hypnotherapy [29], and minimal interventions [33].

Two studies had an additional intervention arm [33] and control arm [35], respectively, increasing the total number of participants in the intervention group compared to the control group. Three studies were conducted in Germany

https://doi.org/10.24171/j.phrp.2025.0019

https://doi.org/10.24171/j.phrp.2025.0019

Table 2. Characteristics of the included studies

Study	Year	Design	Country	No. of participants, (intervention/control)	Sample size estimation	Age (y), mean ± SD (range)	Sex (experiment, control)	Intervention	Control	Outcome assessment measures/studied outcomes
Sykes and Marks [34]	2001	RCT	UK	229 (122/107)	No	NA	45M/86F, 49M/80F	Self-help CBT program (Quit for Life-QFL)	Health Promotion Intervention-SSME	CO reading using analyzer/abstinence at 6 mo
Abdelaal et al. [31]	2021	RCT	Egypt	146 (69/77)	Yes	21.45±1.39/20.88±1.56	100% M	CBT: 8–10 weekly group sessions lasting 90 minutes, focusing on smoking cessation and relapse prevention strategies	Basic general health education: 1 main session and 2 follow-ups	FTND/abstinence at 3, 6 mo
Webb et al. [30]	2010	Mixed factorial design RCT	USA	154 (77/77)	Yes	43±10.1/45±9.3 (18–65)	30M/47F, 37M/40F	CBT: 6 sessions focusing on barriers, previous quit attempts, risky situations	General health education	FTND/abstinence at 2 wk, 3, 6 mo
Wittchen et al. [33]	2011	RCT	Germany	256 (175/81)	No	42.4±13.9/41.5±15.1	86M/89F, 44M/37F	CBT over 9–12 weeks, focusing on motivation, smoking triggers, social support, and relapse prevention	One shot, Minimal Intervention structured quit advice by a doctor	FTND/abstinence at 3 and 12 mo
Ruther et al. [35]	2018	RCT	Germany	100 (51/49)	Yes	51.55±10.32/51.59±11.35	21M/30F, 18M/31F	Smokeless smoking reduction program: 2.5–hour weekly therapy group for 4 weeks, plus 2 telephone consultations	15-minute brief consultation	FTND/smoking reduction, abstinence
Batra et al. [29]	2024	RCT	Germany	359 (179/180)	Yes	43.28±12.18/42.86±12.88	78M/101F, 66M/114F	CBT Smoking Cessation Group Program	Hypnotherapy	FTND/abstinence at 3-time points
Selvamary et al. [32]	2016	Parallel-group RCT	India	194 (94/100)	No	35.64/34.91	185M/9F	CBT: 5 sessions over 24 weeks with follow-up reminders	Health education	Nicotine test/abstinence at 6 mo

SD, standard deviation; RCT, randomized controlled trial; UK, United Kingdom; NA, not available; M, male; F, female; CO, carbon monoxide; CBT, cognitive behavioral therapy; SSME, stopping smoking made easier; FTND, Fagerström Test of Nicotine Dependence.

Table 3. Risk of bias assessment of the included studies

Study	Year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcomes assessment	Incomplete outcome data	Selective reporting bias	Other bias
Sykes and Mark [34]	2001	Low risk	Low risk	High risk	Unclear risk	Low risk	Low risk	Low risk
Abdelaal et al. [31]	2021	Low risk	Low risk	High risk	Unclear risk	Low risk	Low risk	Low risk
Webb et al. [30]	2010	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	Low risk
Wittchen et al. [33]	2011	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk
Ruther et al. [35]	2018	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk
Batra et al. [29]	2024	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	Unclear risk
Selvamary et al. [32]	2016	Low risk	Low risk	High risk	Unclear risk	Low risk	Low risk	Low risk

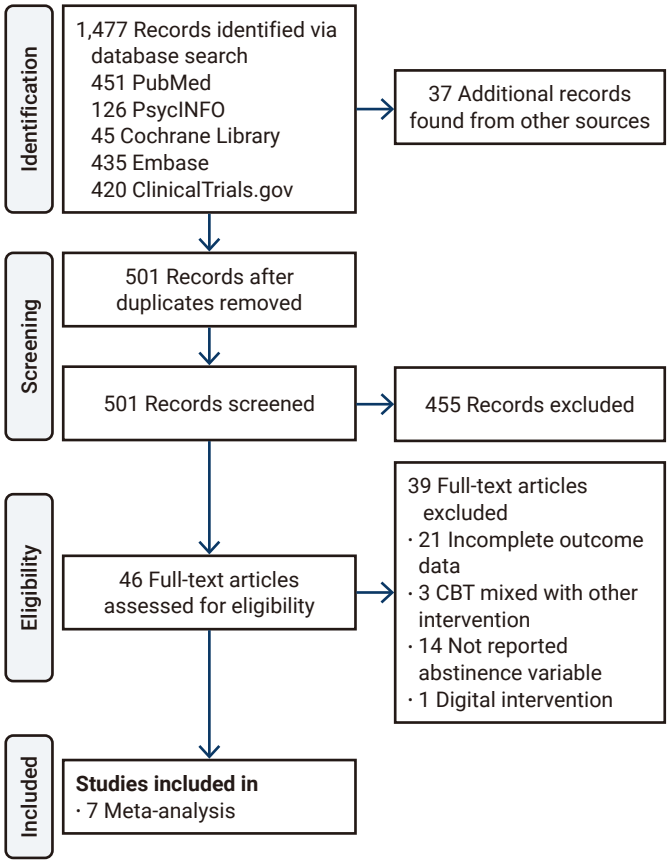


Figure 1. Prisma flow diagram. Adapted from Moher et al. PLoS Med 2009;6:e1000097, according to the Creative Commons license [38]. CBT, cognitive behavioral therapy.



Figure 2. Risk of bias graph.

[29,33,35], 1 in the United States [27,30], and 1 each in India [32], and Egypt [31].

Risk of Bias and Study Quality

The risk of bias for the included studies is presented in Figure 2 and Table 3 [29–35]. Random sequence allocation and allocation concealment were described in all 7 studies [27,29–35]. Five studies were at high risk of blinding participants and personnel [29–32,34], and 2 studies were unclear regarding this domain [33,35]. Three studies documented the blinding of outcomes assessment [29,30,33], and the remaining 4 were unclear in this regard [31,32,34,35]. Additionally, all 7 studies reported a low risk of incomplete outcome data and selective reporting bias [29–35]. Further, 5 studies [30–34]

https://doi.org/10.24171/j.phrp.2025.0019

reported a low risk of other biases, and 2 were unclear [29,35] (Figures 2, 3).

Result of Syntheses

The point prevalence at the end of treatment served as the primary outcome for this review and meta-analysis. Four studies reported the point prevalence at the end of treatment following cognitive behavior therapy [30–32,35]. The pooled analysis indicated that participants in the experimental group (n=252), who received cognitive behavioral treatment, had a fourfold greater likelihood of tobacco use abstinence compared to the control group (n=242) (OR, 4.04; 95% CI, 2.73–5.99; I^2=19%; p<0.00001) (Figure 4).

Further analyses were conducted at 3-month and 6-month follow-ups to determine the long-term impact and sustained effect of CBT on abstinence rates among tobacco users. The pooled analysis demonstrated a significant effect of CBT in maintaining higher abstinence rates at 3-month follow-up in the experimental group (n=346) compared to controls (n=283) (OR, 1.62; 95% CI, 1.11–2.38; I^2=0%; p=0.01) (Figure 5). Additionally, data from 6-month follow-ups revealed that CBT significantly improved abstinence rates in the experimental group (n=523) relative to control conditions (n=500) (OR, 2.19; 95% CI, 1.59–3.00; I^2=0%; p<0.00001) (Figure 6). Unexpectedly, the effectiveness of CBT at maintaining abstinence at 6 months was slightly higher than at the 3-month follow-up but remained lower than at the end of treatment.

Publication bias

Funnel plots were utilized to assess publication bias for the selected outcomes, displaying an approximately symmetrical distribution (Figure 7). This symmetry suggests that publication bias is unlikely to have significantly influenced the meta-analysis findings.

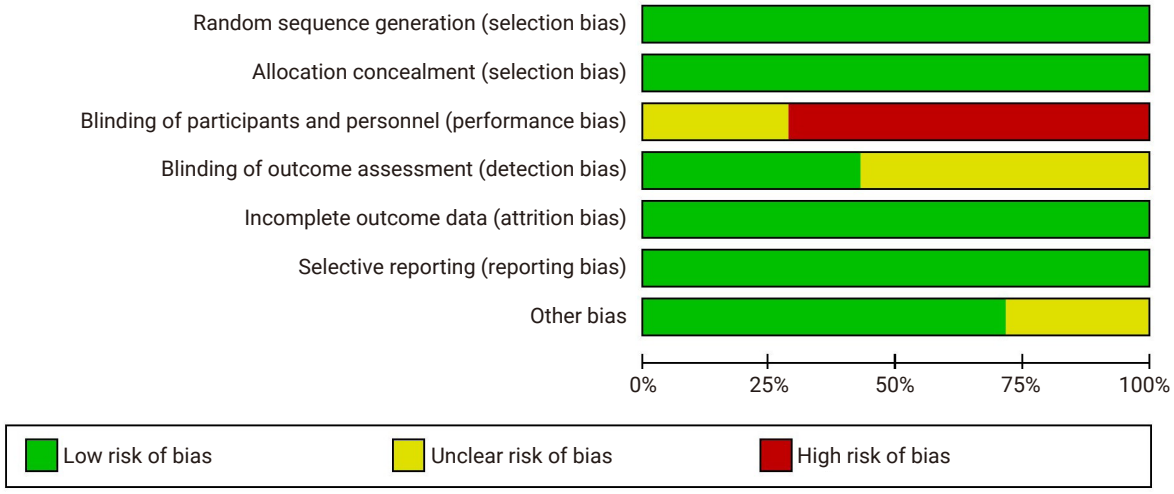


Figure 3. Risk of bias summary.

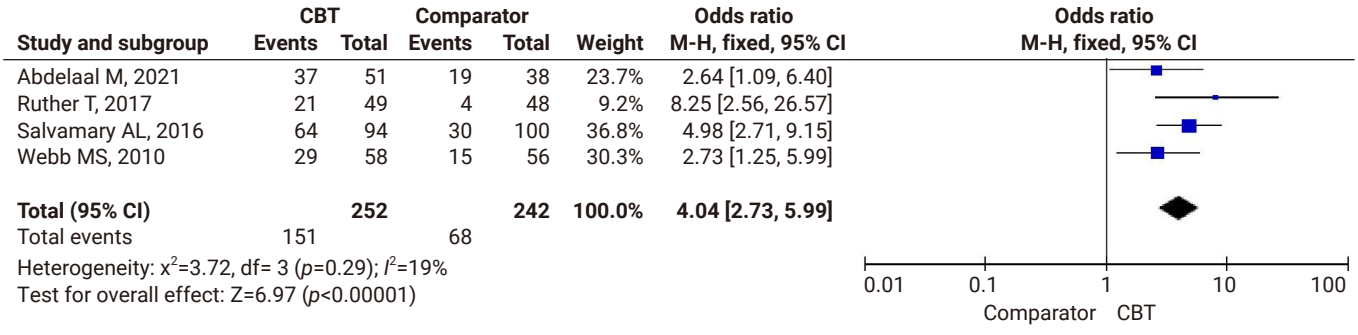


Figure 4. Forest plot: cognitive behavioral therapy (CBT) vs. control at the end of treatment; outcome: abstinence rate.
M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

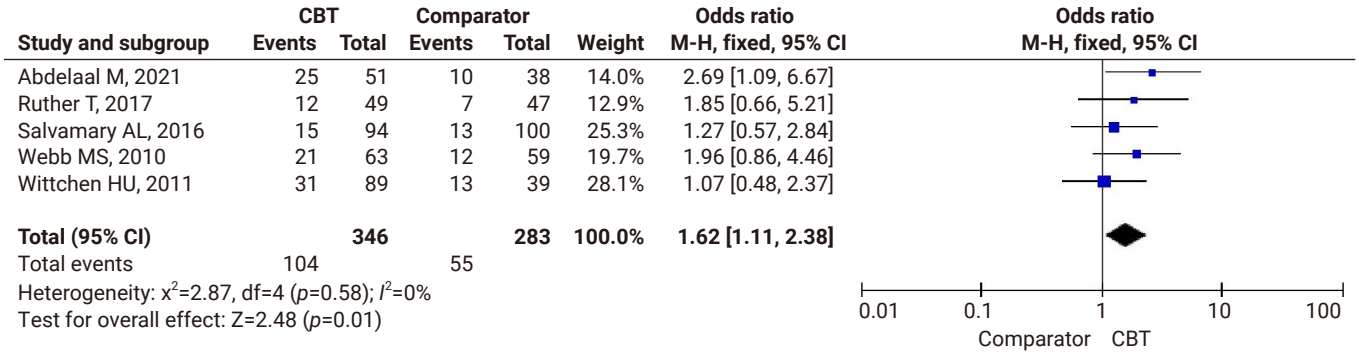


Figure 5. Forest plot: cognitive behavioral therapy (CBT) vs. control at 3 months follow-up; outcome: abstinence rate.
M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

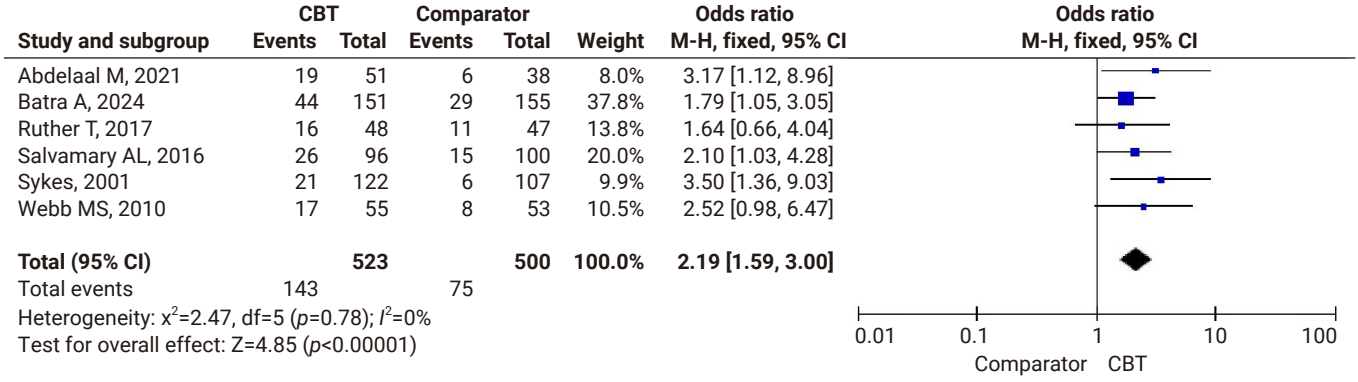


Figure 6. Forest plot: cognitive behavioral therapy (CBT) vs. control at 6 months follow-up; outcome: abstinence rate.
M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

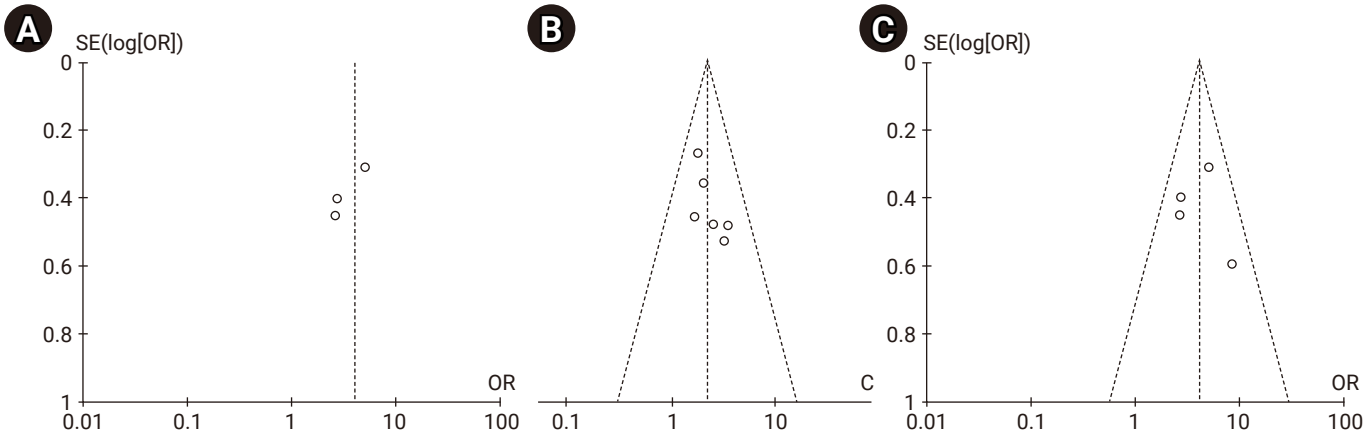


Figure 7. (A–C) Funnel plots comparing the effect of cognitive behavioral therapy (CBT) on the abstinence rate at the end of treatment and at 3- and 6-month follow-ups.

Discussion

Smoking is widely recognized for its harmful effects and is closely associated with increased morbidity and mortality. In response to these risks, numerous novel non-pharmacological interventions—including telephone counseling [39], acupuncture [40], individual [41] and group behavior therapy [42], self-help intervention [39], hypnotherapy [43], healthcare

https://doi.org/10.24171/j.phrp.2025.0019

provider-delivered interventions [44], exercise interventions [45], motivational interviewing [19], and CBT [20,30,46]—have been developed to facilitate smoking cessation, demonstrating promising efficacy. Among these, CBT has emerged as a pivotal psychological intervention specifically targeting smoking cessation [47]. CBT aims to identify and modify maladaptive thought patterns, promote healthier cognitive processes to manage stress, and subsequently reduce relapse rates among individuals with substance use disorders [21,22]. While cognitive therapy focuses primarily on helping individuals develop alternative thought processes and behaviors to alleviate psychological distress [48], CBT specifically emphasizes increasing awareness of negative interpretations and behavioral patterns that reinforce distorted thinking, empowering individuals to modify these patterns and improve mental well-being [49].

Notably, global evidence underscores the potential of CBT to facilitate cessation and improve abstinence rates among tobacco users [25], alcohol, and other illicit substances [50]. However, these studies exhibit significant methodological heterogeneity, resulting in a lack of consensus regarding CBT's effectiveness in promoting treatment adherence and sustaining abstinence, particularly among healthy smokers.

This systematic review and meta-analysis aimed to evaluate adherence to CBT and abstinence rates among tobacco users at the end of treatment and at 3-month and 6-month follow-up intervals.

Previous literature has frequently reported significant benefits of CBT in reducing smoking rates and promoting abstinence among smokers, irrespective of intervention format, whether face-to-face [25,29,51–53], online [54], or through application-based interventions [55]. To the best of our knowledge, no prior meta-analysis has systematically pooled the effects of CBT specifically in terms of treatment completion and adherence at 3-month and 6-month follow-up intervals among healthy smokers. Nonetheless, substantial individual studies have supported the effectiveness of CBT on overall smoking cessation, point prevalence, and outcomes at both short-term and long-term follow-ups [25,27,30–35,53,55,56]. Furthermore, studies indicate that CBT combined with pharmacotherapies has greater efficacy compared to usual care or pharmacotherapy alone for users of tobacco [57], alcohol, and other illicit drugs [58]. However, standalone CBT has been noted to be approximately 5 times less effective than when integrated with pharmacotherapy in smoking cessation programs [58,59]. The results of our meta-analysis explicitly confirmed CBT's effectiveness in improving abstinence among tobacco users. Additionally, analyses at 3- and 6-month follow-ups indicated sustained but variable adherence effects of CBT on smoking cessation. The inconsistency in adherence intensity over different time intervals should be explored through multiple determinants, including personalized support mechanisms, accessible complementary resources such as medication or financial assistance [60], and individual behavioral attributes, notably self-efficacy levels [61]. Similarly, numerous socio-demographic factors, such as male gender, older age, and higher educational attainment, have been identified as correlates of increased adherence rates [62].

Previous meta-analyses have supported using various non-pharmacological interventions—including hypnotherapy [43], competition and incentive-based approaches [16,18,39,40,42], telephone-based counseling [39], motivational counseling [18], group behavior therapy [42], and acupuncture and related interventions [40]—for smoking cessation. CBT approaches aimed at promoting smoking cessation outcomes incorporate cognitive strategies (e.g., learning to manage negative behaviors or urges to smoke), behavioral strategies (e.g., changing habits to avoid temptation), and motivational strategies (e.g., directly confronting adverse health consequences). The U.S. Department of Health and Human Services Guidelines Panel on tobacco treatment identified a strong positive relationship between counseling intensity and successful abstinence [47]. These findings align with our meta-analysis results, demonstrating CBT's effectiveness in smoking cessation interventions. However, the meta-analysis indicated that CBT sessions must be frequent and sustained to achieve lasting improvements in abstinence rates. Our analysis showed that the effectiveness of CBT decreased from baseline to follow-up intervals, supporting previous findings that emphasize the necessity for frequent and sustained CBT sessions [47]. Studies exploring the extended use of CBT have highlighted the importance of booster sessions to maintain CBT's positive effects on smoking cessation, noting that failure to conduct these additional sessions may diminish long-term abstinence outcomes [27,63]. Although the reviewed studies were sufficient in number, their relatively small sample sizes highlighted a high dropout rate and limited long-term adherence to CBT among tobacco users. A substantial dropout rate may lower the observed effect sizes, potentially obscuring CBT's true impact on abstinence rates over time.

Moreover, previous evidence indicates that adherence to tobacco cessation treatment is crucial for achieving successful long-term abstinence [64]. Nevertheless, maintaining consistent participation in therapy remains a significant challenge. Strategies such as financial incentives [65], the use of telehealth services [66], shared decision-making [67], and improved treatment availability and affordability [68] should be implemented to improve adherence to smoking

cessation treatment.

Conversely, research emphasizes the necessity of adopting multimodal approaches—combinations of non-pharmacological and pharmacological interventions—to optimize cessation rates and sustain abstinence [69,70]. The greatest improvements in abstinence rates occur when multiple smoking cessation interventions are combined with behavioral treatment over an extended period. However, the present meta-analysis specifically addressed the long-term effects of CBT alone on abstinence rates. Further research should investigate the comparative effectiveness of pharmacological versus non-pharmacological therapies.

Strengths and Limitations

This meta-analysis provides robust primary evidence supporting the efficacy of CBT in promoting long-term smoking cessation among healthy smokers. Firstly, given the low heterogeneity observed, a fixed-effects model was employed for data analysis, enhancing the validity and reliability of the presented evidence. Secondly, the inclusion criteria were meticulously defined, encompassing only studies involving healthy smokers who received CBT through face-to-face delivery methods. This rigorous selection approach significantly strengthens the robustness of the findings.

However, several limitations must be considered when interpreting the results. Primarily, the number of RCTs meeting the eligibility criteria was limited, and these studies typically involved relatively small sample sizes, potentially impacting the generalizability of the results. Notably, 3 out of the 7 included studies [32–34] did not report sample size calculations, introducing a possible source of bias. Moreover, the inability to assess adherence to smoking cessation facilitated by CBT beyond 12 months, due to the scarcity of eligible studies, highlights a critical gap in the existing evidence and underscores the necessity for further investigation into CBT's long-term effectiveness. Furthermore, the studies included were conducted in selected regions, potentially limiting the representativeness of findings in broader geographical contexts. To overcome this limitation and ensure generalizability, we recommend conducting larger-scale, multicentric RCTs involving diverse populations, particularly from low-socioeconomic and non-Western settings, with rigorous methodologies to offer more definitive conclusions.

Despite these limitations, this meta-analysis provides compelling initial evidence regarding the effectiveness of CBT in improving adherence and abstinence rates among tobacco users.

Implications for Clinical Practice

Cognitive behavioral interventions have been extensively and effectively utilized to support tobacco cessation efforts. Evidence indicates that behavioral improvements achieved through CBT programs can be sustained at follow-up intervals in intervention groups compared to control groups, underscoring the enduring impact of CBT on smoking cessation. Furthermore, incorporating CBT as an adjunctive therapy alongside other established interventions could help decrease smoking prevalence and reduce the global burden of tobacco-related diseases. Additionally, group CBT as a standalone treatment approach represents a potentially cost-effective strategy that clinicians may adopt in clinical practice [58].

Recommendations

Future research should prioritize designing and conducting studies with larger sample sizes to evaluate the cost-effectiveness, feasibility, and safety of CBT. Moreover, future investigations should explore the long-term applicability of CBT relative to other pharmacological and non-pharmacological therapies, thereby strengthening its evidence base and facilitating broader clinical adoption. The findings of this review emphasize the necessity of conducting large-scale, multicentric studies involving diverse populations and standardized follow-up periods to rigorously evaluate long-term adherence to CBT and its sustained efficacy as a smoking cessation intervention.

Conclusion

The pooled statistical analysis demonstrates a promising effect of CBT on abstinence rates at the end of treatment and adherence at 3-month and 6-month follow-up intervals among healthy smokers. However, the observed impact of CBT diminishes from baseline to follow-up intervals and shows variability, highlighting the need for additional research to establish its long-term efficacy conclusively.

Supplementary Material

Supplementary Material 1. Search data; **Supplementary Material 2.** Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) checklist. Supplementary data are available at https://doi.org/10.24171/j.phrp.2025.0019.

Notes

Ethics Approval
Not applicable.

https://doi.org/10.24171/j.phrp.2025.0019

Conflicts of Interest
The authors have no conflicts of interest to declare.

Funding
None.

Availability of Data
The datasets are not publicly available but are available from the corresponding author upon reasonable request.

Authors' Contributions
Conceptualization: RK, PM, KB; Data curation: PM, RK, GSC; Formal analysis: RK, PM, GSC; Methodology: RK, PM, KB, GSC; Project administration: PM, RK, GSC; Resources: RK, PM, KB, GSC; Software: RK, PM; Supervision: RK, PM, KB; Validation: RK, VAS; Visualization: RK, PM, KB, GSC; Writing–original draft: RK, VAS, KB; Writing–review & editing: all authors. All authors read and approved the final manuscript.

Acknowledgements
While preparing this work, the author(s) used Grammarly to improve vocabulary and correct grammar. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

1. Jones P. Nicotine addiction. Aust J Clin Hypnother Hypn 2006;27:1–4.
2. Hilt RJ, Chaudhari M, Bell JF, et al. Side effects from use of one or more psychiatric medications in a population-based sample of children and adolescents. J Child Adolesc Psychopharmacol 2014;24:83–9.
3. Aubin HJ, Luquiens A, Berlin I. Pharmacotherapy for smoking cessation: pharmacological principles and clinical practice. Br J Clin Pharmacol 2014;77:324–36.
4. Saha SP, Bhalla DK, Whayne TF Jr, et al. Cigarette smoke and adverse health effects: an overview of research trends and future needs. Int J Angiol 2007;16:77–83.
5. Bergstrom J. Tobacco smoking and chronic destructive periodontal disease. Odontology 2004;92:1–8.
6. Office of the Surgeon General (US); Office on Smoking and Health (US). The health consequences of smoking: a report of the surgeon general. Centers for Disease Control and Prevention (US); 2004.
7. Kovac JR, Khanna A, Lipshultz LI. The effects of cigarette smoking on male fertility. Postgrad Med 2015;127:338–41.
8. Kock L, Brown J, Cox S, et al. Association of psychological distress with smoking cessation, duration of abstinence from smoking, and use of non-combustible nicotine-containing products: a cross-sectional population survey in Great Britain. Addict Behav 2023;138:107570.
9. Doll R, Peto R, Boreham J, et al. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004;328:1519.
10. US Preventive Services Task Force, Krist AH, Davidson KW, et al. Interventions for tobacco smoking cessation in adults, including pregnant persons: US Preventive Services Task Force recommendation statement. JAMA 2021;325:265–79.
11. Stitzer ML. Combined behavioral and pharmacological treatments for smoking cessation. Nicotine Tob Res 1999;1 Suppl 2:S181–7.
12. Onwuzo CN, Olukorode J, Sange W, et al. A review of smoking cessation interventions: efficacy, strategies for implementation, and future directions. Cureus 2024;16:e52102.
13. Laland KN, Rendell L. Social learning: theory. In: Choe JC: editor. Encyclopedia of animal behavior. 2nd ed. Academic Press; 2010. p. 380–6.
14. Cornuz J, Gilbert A, Pinget C, et al. Cost-effectiveness of pharmacotherapies for nicotine dependence in primary care settings: a multinational comparison. Tob Control 2006;15:152–9.
15. Stead LF, Koilpillai P, Fanshawe TR, et al. Combined pharmacotherapy and behavioural interventions for smoking cessation. Cochrane Database Syst Rev 2016;3:CD008286.
16. Cahill K, Perera R. Competitions and incentives for smoking cessation. Cochrane Database Syst Rev 2008;(3):CD004307.
17. Halpern SD, French B, Small DS, et al. Randomized trial of four financial-incentive programs for smoking cessation. N Engl J Med 2015;372:2108–17.
18. Lindson N, Thompson TP, Ferrey A, et al. Motivational interviewing for smoking cessation. Cochrane Database Syst Rev 2019;7:CD006936.
19. Kumar R, Sahu M, Rodney T. Efficacy of Motivational Interviewing and Brief Interventions on tobacco use among healthy adults: a systematic review of randomized controlled trials. Invest Educ Enferm 2022;40:e03.
20. Nwosu NC, Ede MO, Onah NG, et al. Cognitive behavioral therapy for challenges to quitting tobacco smoking among social science and religion students. Medicine (Baltimore) 2022;101:e31913.
21. Nathan PE, Gorman JM. A guide to treatments that work. 4th ed. Oxford University Press; 2015.
22. Carroll RS. Rational psychotherapy. South Med J 1914;7:451–8.
23. Marlatt GA, Gordon JR. Relapse prevention: maintenance strategies in the treatment of addictive behaviors. Guilford Press; 1986.
24. Fiore MC, Jaen CR, Baker TB, et al. Treating tobacco use and dependence: 2008 update. In: Tobacco Use and Dependence Guideline Panel, editor. Clinical practice guideline. US Department of Health and Human Services, Public Health Service; 2008. p. 1196–9.
25. Farooq MU, Puranik MP, Uma SR. Effectiveness of cognitive-behavioral therapy compared with basic health education for tobacco cessation among smokers: a randomized controlled trial. J Indian Assoc Public Health Dent 2020;18:25–30.
26. Bjornsson AS, Bidwell LC, Brosse AL, et al. Cognitive-behavioral group therapy versus group psychotherapy for social anxiety disorder among college students: a randomized controlled trial. Depress Anxiety 2011;28:1034–42.
27. Webb J, Peerbux S, Ang A, et al. Long-term effectiveness of a clinician-assisted digital cognitive behavioral therapy intervention for smoking cessation: secondary outcomes from a randomized controlled trial. Nicotine Tob Res 2022;24:1763–72.
28. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan: a web and mobile app for systematic reviews. Syst Rev 2016;5:210.
29. Batra A, Eck S, Riegel B, et al. Hypnotherapy compared to cognitive-

behavioral therapy for smoking cessation in a randomized controlled trial. Front Psychol 2024;15:1330362.

30. Webb MS, de Ybarra DR, Baker EA, et al. Cognitive-behavioral therapy to promote smoking cessation among African American smokers: a randomized clinical trial. J Consult Clin Psychol 2010;78:24–33.
31. Abdelaal M, El Kholy O, Rady A, et al. The efficacy of cognitive behavioral therapy for smoking cessation among Egyptian male students in Alexandria university: a randomized clinical study. Senses Sci 2021;8.
32. Selvamary AL, Narayanan MA, Doss J, et al. Effectiveness of cognitive behavior therapy in tobacco cessation at a dental setting: a hospital-based randomized controlled trial. J Indian Assoc Public Health Dent 2016;14:370–6.
33. Wittchen HU, Hoch E, Klotsche J, et al. Smoking cessation in primary care: a randomized controlled trial of bupropione, nicotine replacements, CBT and a minimal intervention. Int J Methods Psychiatr Res 2011;20:28–39.
34. Sykes CM, Marks DF. Effectiveness of a cognitive behaviour therapy self-help programme for smokers in London, UK. Health Promot Int 2001;16:255–60.
35. Ruther T, Kiss A, Eberhardt K, et al. Evaluation of the cognitive behavioral smoking reduction program "Smoke_less": a randomized controlled trial. Eur Arch Psychiatry Clin Neurosci 2018;268:269–77.
36. The Cochrane Collaboration (GB). Review manager web [Internet]. The Cochrane Collaboration; 2024 [cited 2024 Dec 12]. Available from: https://revman.cochrane.org.
37. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343: d5928.
38. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
39. Stead LF, Lancaster T, Perera R. Telephone counselling for smoking cessation. Cochrane Database Syst Rev 2003;(1):CD002850.
40. White AR, Rampes H, Campbell JL. Acupuncture and related interventions for smoking cessation. Cochrane Database Syst Rev 2006;(1):CD000009.
41. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2017;(3):CD001292.
42. Stead LF, Lancaster T. Group behaviour therapy programmes for smoking cessation. Cochrane Database Syst Rev 2005;(2):CD001007.
43. Abbot NC, Stead LF, White AR, et al. Hypnotherapy for smoking cessation. Cochrane Database Syst Rev 2000;(2):CD001008.
44. Gorin SS, Heck JE. Meta-analysis of the efficacy of tobacco counseling by health care providers. Cancer Epidemiol Biomarkers Prev 2004; 13:2012–22.
45. Ussher MH, Faulkner GE, Angus K, et al. Exercise interventions for smoking cessation. Cochrane Database Syst Rev 2019;2019: CD002295.
46. Lee M, Miller SM, Wen KY, et al. Cognitive-behavioral intervention to promote smoking cessation for pregnant and postpartum inner city women. J Behav Med 2015;38:932–43.
47. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update: a U.S. Public Health Service report. Am J Prev Med 2008;35:158–76.
48. Beck AT. Cognitive therapy: nature and relation to behavior therapy: republished article. Behav Ther 2016;47:776–84.
49. Ellis A. Rational psychotherapy. J Gen Psychol 1958;59:35–49.
50. Magill M, Kiluk BD, Ray LA. Efficacy of cognitive behavioral therapy for alcohol and other drug use disorders: is a one-size-fits-all approach appropriate? Subst Abuse Rehabil 2023;14:1–11.
51. Araujo MS, Silva LG, Pereira GM, et al. Mindfulness-based treatment for smoking cessation: a randomized controlled trial. J Bras Pneumol 2022;47:e20210254.
52. Garcia-Fernandez G, Krotter A, Garcia-Perez A, et al. Pilot randomized trial of cognitive-behavioral treatment plus contingency management for quitting smoking and weight gain prevention among smokers with overweight or obesity. Drug Alcohol Depend 2022;236:109477.
53. Kumar R, Goel MK, Talukdar T. Effect of Cognitive Behavioral Therapy on tobacco usage and dependency among adult male tobacco users in Palam area of Delhi. Indian J Community Health 2024;36:26–32.
54. Tang J, Yang J, Liu Y, et al. Efficacy of WeChat-based online smoking cessation intervention ('WeChat WeQuit') in China: a randomized controlled trial. EClinicalMedicine 2023;60:102009.
55. Chen S, Tang J, Wu C, et al. Preliminary efficacy of a cognitive behavioral therapy-based smartphone app for smoking cessation in China: randomized controlled pilot trial. JMIR Form Res 2024; 8:e48050.
56. Krotter A, Garcia-Fernandez G, Garcia-Perez A, et al. Contingency management for smoking cessation for individuals with overweight or obesity: a randomized controlled trial. Drug Alcohol Depend 2024; 258:111269.
57. Ray LA, Meredith LR, Kiluk BD, et al. Combined pharmacotherapy and cognitive behavioral therapy for adults with alcohol or substance use disorders: a systematic review and meta-analysis. JAMA Netw Open 2020;3:e208279.
58. Magill M, Ray LA. Cognitive-behavioral treatment with adult alcohol and illicit drug users: a meta-analysis of randomized controlled trials. J Stud Alcohol Drugs 2009;70:516–27.
59. Irvin JE, Bowers CA, Dunn ME, et al. Efficacy of relapse prevention: a meta-analytic review. J Consult Clin Psychol 1999;67:563–70.
60. Aksel O, Kucuktepe N, Yaslica Z, et al. Providing free access to smoking cessation medications: does it have an impact on the treatment adherence and success of smoking cessation? Turk Thorac J 2021; 22:224–30.
61. Catz SL, Jack LM, McClure JB, et al. Adherence to varenicline in the COMPASS smoking cessation intervention trial. Nicotine Tob Res 2011;13:361–8.
62. Figueiro LR, Barros HM, Ferigolo M, et al. Assessment of factors

https://doi.org/10.24171/j.phrp.2025.0019

related to smokers' adherence to a short-term support group for smoking cessation: a longitudinal study in a developing country. Trends Psychiatry Psychother 2017;39:19–28.

63. Killen JD, Fortmann SP, Schatzberg AF, et al. Extended cognitive behavior therapy for cigarette smoking cessation. Addiction 2008; 103:1381–90.
64. Cheng A, Liu Z, Zhou X, et al. Real-world treatment adherence to smoking cessation medications and association with treatment outcome: 27 centers, China, 2017-2022. China CDC Wkly 2024;6:522–9.
65. Kendzor DE, Businelle MS, Frank-Pearce SG, et al. Financial incentives for smoking cessation among socioeconomically disadvantaged adults: a randomized clinical trial. JAMA Netw Open 2024;7:e2418821.
66. Gong J, Baker CL, Zou KH, et al. A pragmatic randomized trial comparing telephone-based enhanced pharmacy care and usual care to support smoking cessation. J Manag Care Spec Pharm 2016;22:1417–25.
67. Baskerville WA, Friedman TC, Hurley B, et al. Embedding comprehensive smoking cessation programs into community clinics: study protocol for a cluster-randomized controlled trial. Trials 2022;23:109.
68. Hollands GJ, Naughton F, Farley A, et al. Interventions to increase adherence to medications for tobacco dependence. Cochrane Database Syst Rev 2019;8:CD009164.
69. Shiffman S, Paty JA, Rohay JM, et al. The efficacy of computer-tailored smoking cessation material as a supplement to nicotine patch therapy. Drug Alcohol Depend 2001;64:35–46.
70. Hall SM, Reus VI, Munoz RF, et al. Nortriptyline and cognitive-behavioral therapy in the treatment of cigarette smoking. Arch Gen Psychiatry 1998;55:683–90.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2024.0349
Osong Public Health Res Perspect 2025;16(3):236-251

Original Article

Design and evaluation of a multi-epitope subunit vaccine against human norovirus using an immunoinformatics approach

Zxcy L. Nonog[1], Edward C. Banico[2], Ma. Easter V. Sajo[1], Pablo V. Serrano[1], Fredmoore L. Orosco[2,3,4]

[1]Department of Biology, College of Science, University of the Philippines Baguio, Baguio, Philippines
[2]Virology and Vaccine Research Program, Industrial Technology Development Institute, Department of Science and Technology, Taguig City, Philippines
[3]S&T Fellows Program, Department of Science and Technology, Taguig City, Philippines
[4]Department of Biology, College of Arts and Sciences, University of the Philippines Manila, Manila City, Philippines

ABSTRACT

Objectives: This study aimed to identify safe, conserved, and highly immunogenic epitopes from all proteins of human-infecting norovirus (NoV) and to design a multi-epitope subunit vaccine construct from these epitopes using an immunoinformatics approach. Additionally, the vaccine construct was evaluated using both sequence- and structure-based assessments.

Methods: Conserved fragments were identified from all proteins of human-infecting NoV, and B and T lymphocyte epitopes were subsequently predicted using multiple epitope prediction tools. The selected epitopes were linked to form a multi-epitope construct, incorporating various adjuvants in the design. Vaccine constructs with different adjuvants were analyzed for their physicochemical properties and immune simulation profiles, and the optimal combination was selected as the final vaccine candidate for further study. Finally, molecular docking and dynamics simulations were performed to visualize the interaction between the construct and a host immune receptor.

Results: Twenty-two safe, conserved, and highly immunogenic epitopes were identified from all human-infecting NoV proteins. The construct adjuvanted with 50S ribosomal protein L7/L12 (50SrpL7/L12) was chosen as the final vaccine candidate due to its optimal physicochemical properties and favorable immune simulation profile. Furthermore, the construct exhibited high binding affinity and a stable interaction with toll-like receptor 4).

Conclusion: The multi-epitope subunit vaccine designed in this study shows promise as a potential NoV vaccine candidate for human immunization. Further *in vitro* and *in vivo* experiments are warranted to validate these findings.

Keywords: Epitopes; Immunoinformatics; Protein subunit vaccines; Norovirus; Vaccine adjuvants

Received: December 16, 2024
Revised: February 8, 2025
Accepted: March 18, 2025

Corresponding author:
Fredmoore L. Orosco
Virology and Vaccine Research Program, Industrial Technology Development Institute, Department of Science and Technology, General Santos Avenue, Bicutan, Taguig City 1634, Philippines
E-mail: florosco@up.edu.ph

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Norovirus (NoV), a member of the *Caliciviridae* family, is the leading cause of viral acute gastroenteritis (AGE) worldwide [1]. It is associated with approximately 685 million AGE cases [2] and over 200,000 deaths annually [3]. NoV-induced AGE imposes a significant economic and clinical burden, with an estimated global loss of $60 billion in healthcare and societal costs [4].

The NoV virion comprises 3 major proteins: the major capsid protein (VP1), the minor capsid protein (VP2), and a nonstructural polyprotein (NSP), which includes nonstructural proteins such as p48, NTPase, p22, VPg, 3CLpro, and RNA-dependent RNA polymerase (RdRp) [5,6]. NoVs exhibit extensive genetic diversity and are currently classified into 10 genogroups (GI–GX) based on the VP1 sequence [7]. Strains from each genogroup infect specific host species [5,7,8]. Notably, GI, GII, GIV, GVIII, and GIX are associated with human infections [5], with GII being the most prevalent, accounting for approximately 90% of all reported cases and outbreaks [5,8].

To date, no specific therapeutics for NoV-induced AGE exist. Current treatments primarily focus on hydration and supportive care [5]. Severe cases may require hospitalization for intravenous fluid administration [9]. However, in low-income countries with limited access to rehydration therapies, preventing NoV-AGE is even more critical [5]. Preventing NoV infection largely depends on proper hygiene practices, such as thorough hand washing, correct food handling, and minimizing contact with infected individuals [10]. Although improved hygiene and sanitation are crucial, they may not fully control NoV transmission, particularly in high-density or communal environments [11]. Vaccination remains the most effective measure to prevent viral spread [12]. By priming the immune system to recognize and neutralize pathogens, vaccination directly prevents infection and offers protection regardless of environmental conditions [13]. Currently, no licensed NoV vaccine is available, largely due to the virus's extensive genetic diversity, limited understanding of immune responses, and the lack of an effective cell culture system for viral propagation, which renders conventional live attenuated or inactivated vaccine approaches unfeasible [14,15]. Consequently, non-replicating recombinant protein-based vaccines have emerged as a preferred alternative [15].

Multi-epitope subunit vaccines are recombinant protein-based formulations that incorporate immunogenic regions, or epitopes, derived from 1 or multiple viral proteins into a single construct. This design preserves immunogenicity because only the specific epitopes interact directly with immune cells, including B and T lymphocytes, to stimulate antibody production and activate cell-mediated immune responses [16]. By eliminating non-essential (non-immunogenic) or potentially immunosuppressive regions, the vaccine minimizes unwanted reactions and optimizes immune targeting [17].

HIGHLIGHTS

- Norovirus (NoV) is the leading cause of viral gastroenteritis worldwide, and no licensed vaccine is currently available.
- Immunoinformatics streamlines the vaccine development process.
- The multi-epitope subunit vaccine against NoV, designed in this study using an immunoinformatics approach, exhibited optimal physicochemical properties and a favorable immune simulation profile.
- This is the first multi-epitope subunit vaccine against NoV that maps epitopes from all proteins of the human-infecting NoV genogroups.

Traditional development of multi-epitope subunit vaccines has relied on *in vivo* and *in vitro* methods [18]. Initially, target proteins are isolated from the pathogen and enzymatically or chemically fragmented to produce smaller peptide segments. These segments are then tested for immunogenicity through animal inoculations or immunoassays assessing antibody production, cytotoxicity, T lymphocyte proliferation, and cytokine release [18]. However, the advent of immunoinformatics has revolutionized vaccine development by streamlining epitope identification. Modern prediction tools can accurately identify epitopes from viral proteins that are capable of activating B and T lymphocytes based on their sequence and structural properties [17]. The computational tools employed in this study leverage established correlations between these properties and biological outcomes, as demonstrated in several studies [19–23]. This approach enables the prediction of biological functions without extensive wet lab experiments, thereby accelerating vaccine development. Immunoinformatics has already been successfully applied to vaccine design for various pathogens [24,25].

This study aimed to design and evaluate a multi-epitope subunit vaccine against NoV using immunoinformatics. The approach involved (1) mapping epitopes across all proteins of human-infecting NoV genogroups, (2) designing the vaccine using the predicted epitopes, and (3) conducting *in silico* evaluations through a series of sequence- and structure-based tests. This work is vital for advancing NoV vaccine candidates, reducing NoV-AGE-related morbidity and mortality, and achieving a significant breakthrough in global public health.

Materials and Methods

The design and evaluation of the NoV vaccine construct were performed entirely using *in silico* approaches. Figure 1 presents a summary of the methodology employed in this study.

Identification of Conserved Regions

The full-length reference sequences for VP1, VP2, and NSP proteins of human-infecting NoV genogroups (GI, GII, GIV, GVIII, and GIX) were retrieved from the National Center for Biotechnology Information (NCBI) database [26] (https://www.ncbi.nlm.nih.gov/) in November 2023. Sequences for each protein were aligned within their respective genogroups using the Multiple Sequence Alignment (MSA) tool in Clustal Omega [27] (https://www.ebi.ac.uk/jdispatcher/msa/clustalo). The resulting alignment files were then uploaded to the Protein Variability Server (PVS, http://imed.med.ucm.es/PVS/) to extract conserved fragments. A Shannon Diversity Index (H) threshold of 1.0 was applied to identify variable residues. This index is calculated as

$$H = -\sum_{i=1}^{M} p_i log_2 p_i$$ [28],

where p_i denotes the proportion of residues of amino type i, and M refers to the total number of amino acid types present at a specific site. The H value ranges from 0 (indicating a single amino acid type at that position) to 4.322 (indicating equal representation of all 20 amino acids at that position), with an H value=1.0 considered low variability or conservation [29]. Conserved fragments with ≥9-mer, ≥15-mer, and ≥16-mer lengths were collected.

Epitope Mapping

The 9-mer fragments were uploaded to NetMHCCons-1.1 (https://services.healthtech.dtu.dk/services/NetMHCcons-1.1/) to identify potential cytotoxic T lymphocyte (CTL) epitopes. NetMHCCons-1.1 predicts CTL epitopes by assessing their binding affinity (BA) to known major histocompatibility complex (MHC) class I molecules using 3 methods: NetMHC (an artificial neural network-based allele-specific method),

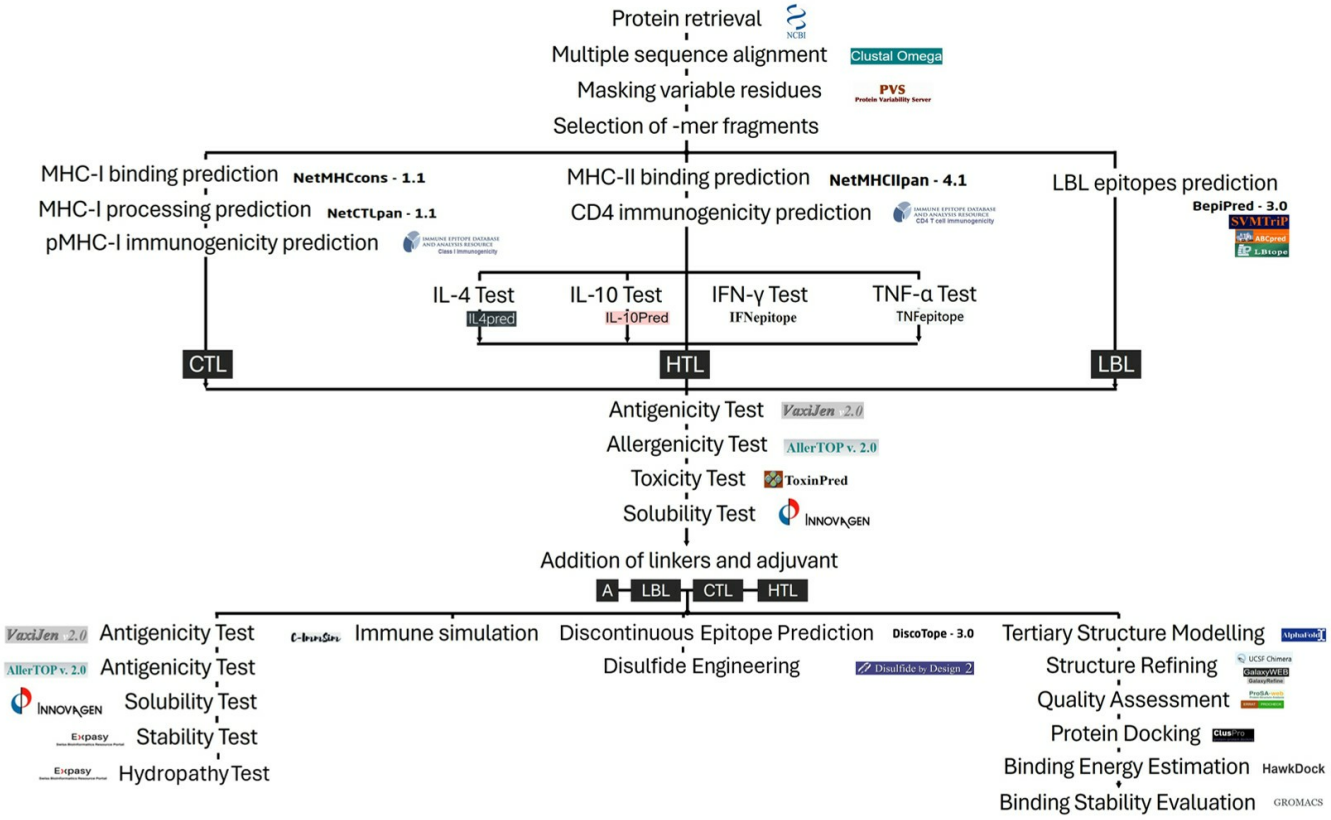


Figure 1. Workflow of the study, showing the immunoinformatics tools used.
MHC, major histocompatibility complex; pMHC, peptide-bound major histocompatibility complex; CD, cluster of differentiation; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte; LBL, linear B lymphocyte; A, adjuvant; IL, interleukin; IFN, interferon; TNF, tumor necrosis factor.

https://doi.org/10.24171/j.phrp.2024.0349

NetMHCpan (a pan-specific ANN method trained on extensive binding data), and PickPocket (a matrix-based method analyzing receptor-pocket similarities) [30]. The analysis used 9-mer peptide lengths and reference human leukocyte antigen (HLA) class I molecules [31], with a screening threshold set at a percentile rank (PR) of ≤0.5 or a half-maximal inhibitory constant (IC50) of ≤50 nM. Selected CTL epitopes were further evaluated using NetCTLpan 1.1 [32] (https://services.healthtech.dtu.dk/services/NetCTLpan-1.1) with the same HLA parameters and a PR threshold of ≤1.0. NetCTLpan 1.1 integrates predictions for MHC class I binding (via an ANN method similar to NetMHCpan), proteasomal C-terminal cleavage (using NetChop neural networks) [33], and TAP transport efficiency (using a weight matrix-based method described by Peters et al. [34]). The Class I Immunogenicity tool (http://tools.iedb.org/immunogenicity) was then used with a score threshold of >0 to predict immunogenicity based on amino acid enrichment, positional importance, and MHC binding and processing affinities [35]. Finally, the CTL epitopes were subjected to additional filtering using a score threshold of ≥0.4 in VaxiJen 2.0 (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html), confirmation as non-allergenic by AllerTOP v2.1 (http://www.ddg-pharmfac.net/AllerTOP), non-toxicity by ToxinPred (https://webs.iiitd.edu.in/raghava/toxinpred/multi_submit.php), and favorable water solubility in Innovagen's PepCalc (https://pepcalc.com/). VaxiJen employs an alignment-free method based on auto-cross covariance (ACC) to transform protein sequences into uniform vectors, with descriptors for hydrophobicity, molecular size, and polarity [19]. Similarly, AllerTOP uses ACC to represent sequences with descriptors for hydrophobicity, molecular size, helix-forming propensity, abundance, and β-strand forming propensity [20]. ToxinPred identifies toxic peptides using a combination of similarity-based, motif-based, and machine/deep learning techniques [23]. Innovagen's PepCalc estimates water solubility by evaluating molecular weight, hydrophobicity, and isoelectric point. Epitopes meeting these criteria were selected as the final CTL epitopes.

The 15-mer fragments were uploaded to NetMHCIIpan 4.3 (https://services.healthtech.dtu.dk/services/NetMHCIIpan-4.3/) to identify potential helper T lymphocyte (HTL) epitopes. NetMHCIIpan 4.3 predicts HTL epitopes by assessing their binding to known MHC class II molecules using an ANN method trained on extensive BA and eluted ligand (EL) data [36]. The analysis employed 15-mer peptide lengths and reference HLA class II molecules [37] with a PR threshold of ≤1.0. Selected HTL epitopes were further evaluated for immunogenicity using the CD4 T cell immunogenicity prediction tool (http://tools.iedb.org/CD4episcore) with a score threshold of ≥40. This tool combines neural networks trained on experimentally validated epitopes and negative peptides to assess amino acid enrichment and positional importance, identifying peptides with the potential to bind T cell receptors and induce a T cell response [38]. The HTL epitopes were also filtered using VaxiJen 2.0, AllerTOP v2.1, ToxinPred, and Innovagen's PepCalc as described above. Furthermore, only epitopes predicted to induce at least 1 of the following cytokines were selected: tumor necrosis factor (TNF)-α (using TNFepitope at https://webs.iiitd.edu.in/raghava/tnfepitope), interferon (IFN)-γ (using IFNepitope at https://webs.iiitd.edu.in/raghava/ifnepitope), interleukin (IL)-4 (using IL4Pred at http://crdd.osdd.net/raghava/il4pred/), and IL-10 (using IL10Pred at https://webs.iiitd.edu.in/raghava/il10pred/). TNFepitope employs alignment-free machine learning models based on composition-based features and alignment-based methods using BLAST to compare peptide sequences [39]. These methods, trained on datasets of TNF-α-inducing and non-inducing epitopes, score peptides based on their ability to activate TNF-α production. Similarly, IFNepitope, IL4Pred, and IL10Pred use a hybrid approach combining motif-based prediction with Support Vector Machine (SVM) classification [40–42] to distinguish cytokine-inducing peptides.

The 16-mer fragments were uploaded to BepiPred 3.0 (https://services.healthtech.dtu.dk/services/BepiPred-3.0) to predict potential B lymphocyte epitopes. BepiPred 3.0 identifies B lymphocyte epitopes by analyzing protein sequences using ESM-2 protein language model representations and a neural network that accounts for hydrophilicity, flexibility, surface accessibility, and amino acid composition [43]. The predicted epitopes were cross-validated using SVMTrip (http://sysbio.unl.edu/SVMTriP/prediction.php) and the ABCPred server (https://webs.iiitd.edu.in/raghava/abcpred); SVMTrip utilizes SVM classifiers [44] and ABCPred is based on an ANN [45]. A threshold of 0.85 was used for ABCPred while SVMTrip was run with default parameters. Epitopes validated by either tool were then submitted to the LBtope server (https://webs.iiitd.edu.in/raghava/lbtope) to assess prediction confidence. LBtope distinguishes potential B lymphocyte epitopes using datasets of verified epitopes and non-epitopes from the Immune Epitope Database [46]. Only epitopes with a probability of ≥60% were considered for further analysis. Finally, the linear B lymphocyte (LBL) epitopes were further filtered using VaxiJen 2.0, AllerTOP v2.1, ToxinPred, and Innovagen's PepCalc as previously described, and those meeting all criteria were selected as the final LBL epitopes.

Vaccine Designing

The AAY linker was employed to connect CTL epitopes. This linker facilitates cleavage and selective binding of the attached epitopes to chaperones or TAP transporters, thereby increasing the number of epitopes available for MHC binding and presentation [47]. The KK linker was used to join LBL epitopes; pairs of basic residues such as KK serve as cleavage sites for cathepsin B, aiding in the release of LBL epitopes for efficient presentation [48]. The GPGPG linker connected the HTL epitopes, with regions rich in glycine and proline being associated with β-turns and hypothesized to enhance secondary structure formation and antigen processing [49]. The HEYGAEALERAG linker was used to connect the 3 epitope groups; it contains 5 specific proteasome recognition sites that promote efficient cleavage during antigen processing [50]. The combined epitopes were then joined with 3 distinct adjuvants—truncated human β-defensin 2 (thβd2), truncated human β-defensin 3 (thβd3), and 50S ribosomal protein L7/L12 (50SrpL7/L12)—using the EAAAK linker. Both β-defensins and 50SrpL7/L12 are recognized adjuvants [51–53] and are frequently employed in multi-epitope vaccine design studies [54–61]. The EAAAK linker acts as a structural spacer, controlling the distance between the epitope and adjuvant regions and minimizing potential interference [62].

Assessment of Physicochemical Properties

The physicochemical properties of the adjuvanted vaccine constructs were evaluated using VaxiJen 2.0, AllerTOP v2.1, and Innovagen's PepCalc under the conditions described above. Expasy ProtParam [22] (https://web.expasy.org/protparam) was used to estimate construct stability and the Grand Average of Hydropathy (GRAVY). ProtParam calculates protein stability using the instability index—which predicts whether a protein is stable *in vitro* based on its dipeptide composition as described by Guruprasad et al. [63]—and the GRAVY score, which indicates overall hydrophobicity or hydrophilicity, computed as:

$$GRAVY = \frac{\sum_{i=1}^{N} H_i}{N}$$ [22],

where H_i is the hydropathy value of each amino acid (based on Kyte and Doolittle's hydropathy index [64]), and N is the number of residues in the sequence.

Immune Response Simulation

Host immune response profiles were predicted using C-Immsim (https://kraken.iac.rm.cnr.it/C-IMMSIM). Simulations involved 3 injections of 1,000 particles administered at intervals of 1, 84, and 168 time steps, with the response analyzed over 300 time steps. C-ImmSim integrates several established immunological theories—including clonal selection, clonal deletion, hypermutation, T lymphocyte senescence, anergy, antigen-dose tolerance in B lymphocytes, and the danger theory [65,66]—to simulate immune responses. Immune parameters such as antibody titers, IFN-γ levels, and the populations of B lymphocytes, CTLs, and HTLs were compared across constructs, and the vaccine candidate demonstrating the most favorable combination of physicochemical properties and immune simulation outcomes was selected as the final NoV vaccine candidate.

Tertiary Structure Prediction

The tertiary structure of the final vaccine construct was predicted using AlphaFold v2 (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb). AlphaFold v2 employs a deep learning model that uses MSAs, co-evolutionary data, and a neural network to generate highly accurate 3-dimensional structures from the amino acid sequence, iteratively refining predictions through an attention-based mechanism [67]. The predicted model was submitted to UCSF Chimera 1.17.1 for energy minimization using 10,000 steps of steepest descent; Chimera 1.17.1 applies algorithms such as steepest descent or conjugate gradient to achieve a more stable, lower-energy conformation [68]. Further refinement was performed with the GalaxyWEB Refine service (http://galaxy.seoklab.org/refine), which optimizes protein structures by rebuilding side chains, resolving steric clashes, and applying both mild and aggressive relaxation via short molecular dynamics simulations [69]. Structural quality was evaluated using ERRAT and Procheck in SAVES v6.0 (https://saves.mbi.ucla.edu); ERRAT assesses the distribution of non-bonded atomic interactions [70], while Procheck examines stereochemical parameters such as bond angles, torsion angles, and backbone conformation [71]. The highest-scoring model was selected as the representative structure of the vaccine construct. Additionally, Protein Structure Analysis (ProSA)-web (https://prosa.services.came.sbg.ac.at/prosa.php) was used to further assess structural quality by comparing the energy profile of the protein structure to a database of experimentally determined structures, providing a Z-score where lower values indicate a more stable conformation [72]. Finally, ChimeraX 1.9 [73] was employed to visualize the final vaccine construct.

Discontinuous B lymphocyte Epitope Mapping and Disulfide Engineering

Discontinuous B lymphocyte epitopes within the vaccine construct were predicted using the DiscoTope-3.0 webserver (https://services.healthtech.dtu.dk/services/DiscoTope-3.0/)

https://doi.org/10.24171/j.phrp.2024.0349

with default parameters. DiscoTope-3.0 uses inverse folding structure representations and positive-unlabeled learning to predict B lymphocyte epitopes [74].

Additionally, Disulfide by Design 2.0 (DbD2) (http://cptweb.cpt.wayne.edu/DbD2/) was used to identify residue pairs within the vaccine construct that could be mutated to cysteine to enhance structural stability. DbD2 incorporates B-factor values to assess regional flexibility and ranks potential disulfide bonds based on mobility [75]. The analysis was performed with default parameters, applying a bond energy threshold of ≤ 2.0 kcal/mol. The instability index of the mutated construct was then determined using Expasy ProtParam and compared to that of the original construct.

Molecular Docking and Dynamics Simulation

The crystal structure of human toll-like receptor 4 (TLR4) was downloaded from the RCSB Protein Data Bank [76] (https://www.rcsb.org) (PDB ID: 3FXI). Peptides, water molecules, and extraneous atoms were removed from the structure using PyMOL v2 [77], and missing residues were modeled using Modeller [78] (https://salilab.org/modeller). The final vaccine construct was forcibly docked to the TLR4 structure using ClusPro 2.0 (https://cluspro.bu.edu/login.php), which performs (1) rigid body docking, (2) root-mean-square deviation (RMSD)-based clustering of the 1,000 lowest-energy structures to identify the largest clusters, and (3) refinement via energy minimization of selected structures [79]. Forced docking was achieved by applying an attraction parameter to constrain binding at the agonist-binding site of TLR4. Binding residues were identified from the LPS-bound TLR4 crystal structure (PDB ID: 3FXI) using ChimeraX's Contacts function, selecting residues with a van der Waals overlap of −0.40 or lower as attraction residues to ensure biologically relevant binding interactions.

The BA and per-residue energy decomposition of the complex were determined using HawkDock 2.0 (http://huanglab.phys.hust.edu.cn/hpepdock). This server calculates BA using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method, which estimates the free energy difference between bound and unbound states [80]. It computes per-residue energy decomposition by breaking down the total binding energy into contributions from individual residues, identifying key stabilizing or destabilizing interactions [80].

The stability of the complex was analyzed through molecular dynamicssimulations using the Southampton Interface and Reduction Algorithm for Hydrodynamics force field [81] in GROMACS 2023.2 [82] (https://manual.gromacs.org/2023.2/download.html). The system was prepared by placing the complex in a cubic simulation box with periodic boundary conditions, ensuring a minimum distance of 1 Å between the complex and the box boundaries to avoid interactions between periodic images. The box was solvated with the TIP3P water model, and sodium and chloride ions were added to achieve a 0.15 M concentration. Energy minimization was performed with up to 50,000 steepest descent steps, followed by 2 equilibration phases: an NVT equilibration to gradually stabilize the temperature at 312 K using a modified Berendsen thermostat, and an NPT equilibration to stabilize the pressure at 1 bar using a Berendsen barostat. A 100 ns production run was then executed, generating 50 million time steps of 2 fs each, with snapshots recorded every 100 ps (totaling 1,000 frames). The trajectory's stability was evaluated using RMSD analysis.

Codon Optimization and *In Silico* Cloning

The JAVA Codon Adaptation Tool (JCat) (https://www.jcat.de/) was employed to generate an optimized codon sequence of the vaccine construct for efficient expression in an *Escherichia coli* K-12 system. JCat evaluates codon adaptation based on the Codon Adaptation Index (CAI) and adjusts rare or less frequently used codons to improve translation efficiency [83]. BamHI and XhoI restriction sites were incorporated at the N-terminal and C-terminal ends, respectively. The optimized sequence was subsequently inserted into the pET-28a(+) vector via restriction cloning using the SnapGene tool (http://www.snapgene.com).

Results

Identification of Conserved Regions

Nineteen NSP, 15 VP1, and 18 VP2 reference proteins were retrieved from the NCBI database. Table 1 summarizes the distribution of these proteins across the genogroups. Sequence alignment was performed for each genogroup, and the conserved fragments identified from these alignments are detailed in Tables S1–S3. These conserved fragments served as inputs for epitope prediction.

Epitope Mapping

Through epitope screening, 22 epitopes were identified

Table 1. Distribution of the retrieved proteins across genogroups

Protein	No. of proteins					
	GI	GII	GIV	GVIII	GIX	Total
NSP	5	5	3	1	5	19
VP1	5	5	3	1	1	15
VP2	5	5	3	2	3	18

G, genogroup; NSP, nonstructural polyprotein; VP1, major viral capsid protein; VP2, minor viral capsid protein.

from the VP1, VP2, and NSP proteins of the human-infecting NoV genogroups. Of these, 4 were classified as potential CTL inducers, 11 as potential HTL inducers, and 7 as potential B lymphocyte inducers (Table 2). Most epitopes were derived from NSPs (73%), while the 2 capsid proteins contributed an equal number of predicted epitopes (13% each). All genogroups were represented, with the epitopes covering GI (9%), GII (9%), GIV (32%), GVIII (32%), and GIX (36%).

All predicted CTL epitopes exhibited high affinity for HLA class I molecules and strong $CD8^+$ immunogenicity. Similarly, all HTL epitopes showed high affinity for HLA class II molecules and robust $CD4^+$ immunogenicity. A summary of the physicochemical properties of the T lymphocyte epitopes is provided in Table 3. Additionally, HTL epitopes induced at least 1 key cytokine associated with NoV infection, such as TNF-α, IFN-γ, IL-4, or IL-10 (Table 4).

All 7 predicted B lymphocyte epitopes demonstrated high predictive accuracy across at least 3 prediction tools. Table 5 summarizes the physicochemical properties of the B lymphocyte epitopes. Overall, the epitopes were predicted to be highly conserved, strongly antigenic, non-allergenic, non-toxic, and soluble.

Vaccine Designing

Three vaccine constructs were designed using the identified epitopes combined with 3 different adjuvants: thβd2, thβd3, and 50SrpL7/L12. Figure 2 illustrates the general schematic

Table 2. Final epitope components of the vaccine construct with corresponding protein sources and covered genogroup(s)

Epitopes	Protein sources	Genogroup(s) covered
Cytotoxic T lymphocyte		
YLGGRDPRV	NSP	I, II, IV, VIII
MPLLDDFEL	NSP	IV
SMLDVGDYV	NSP	IV
RPTGECLPL	NSP	IV
Helper T lymphocyte		
GPIIFEKHSRYKYHY	NSP	II
VGFTAEKAGRLLSSA	NSP	IV
PHVIVDVRQLEPVRL	VP1	IX
ARIAAVRSLAFKAKE	NSP	IV
QPAFAHLRKRGISEA	VP2	VIII
TTAISASTASSRTSD	VP2	VIII
KASILSNMAVTFKKA	NSP	VIII
KIKKVANAVLCALGS	NSP	IX
PIIFDKHAKYKYHYD	NSP	IX
RIDFLVYAEAPDIEK	NSP	IX
RDHFKADSSHIKLSL	NSP	IX
Linear B lymphocyte		
DKHAKYKCHYDADYSR	NSP	IX
EIRKRNPDDFQPKGNL	NSP	VIII, IX
PSSPKATGRFIIESKD	NSP	VIII
PPNGEDATHFKKETKT	NSP	IX
EFTISPNNTPGDILFD	VP1	I
EAYVNTTDSDFAPATGNTKI	VP1	IV
GFSPTDAARGAVNAPMTK	VP2	VIII

NSP, nonstructural polyprotein; VP1, major viral capsid protein; VP2, minor viral capsid protein.

Table 3. Physicochemical properties of predicted T lymphocyte epitopes

Epitopes	Best binding HLA[a)]	Antigenicity score[b)]	Allergenicity[c)]	Toxicity[d)]	Solubility[e)]
CTL					
YLGGRDPRV	A*02:01	1.22	Non-allergenic	Non-toxin	Soluble
MPLLDDFEL	B*35:01	0.62	Non-allergenic	Non-toxin	Soluble
SMLDVGDYV	A*02:01	0.97	Non-allergenic	Non-toxin	Soluble
RPTGECLPL	B*07:02	1.42	Non-allergenic	Non-toxin	Soluble
HTL					
GPIIFEKHSRYKYHY	DRB1*12:01	0.86	Non-allergenic	Non-toxin	Soluble
VGFTAEKAGRLLSSA	DRB1*01:01	0.85	Non-allergenic	Non-toxin	Soluble
PHVIVDVRQLEPVRL	DRB1*03:01	0.97	Non-allergenic	Non-toxin	Soluble
ARIAAVRSLAFKAKE	DRB1*12:01	0.81	Non-allergenic	Non-toxin	Soluble
QPAFAHLRKRGISEA	DRB1*11:01	0.92	Non-allergenic	Non-toxin	Soluble
TTAISASTASSRTSD	DRB5*01:01	0.85	Non-allergenic	Non-toxin	Soluble
KASILSNMAVTFKKA	DRB3*02:02	0.76	Non-allergenic	Non-toxin	Soluble
KIKKVANAVLCALGS	DPA1*02:01-B1*14:01	0.50	Non-allergenic	Non-toxin	Soluble
PIIFDKHAKYKYHYD	DRB1*03:01	0.64	Non-allergenic	Non-toxin	Soluble
RIDFLVYAEAPDIEK	DRB1*09:01	0.54	Non-allergenic	Non-toxin	Soluble
RDHFKADSSHIKLSL	DRB5*01:01	0.65	Non-allergenic	Non-toxin	Soluble

HLA, human leukocyte antigen; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte.
[a)]Identified using NetMHCCons 1.1 (https://services.healthtech.dtu.dk/services/NetMHCcons-1.1/) for CTL and NetMHCIIpan 4.3 (https://services.healthtech.dtu.dk/services/NetMHCIIpan-4.3/) for HTL epitope prediction. [b)]Identified using VaxiJen 2.0 (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html); antigenic ≥0.40. [c)]Identified using AllerTOP v2.1 (http://www.ddg-pharmfac.net/AllerTOP). [d)]Identified using ToxinPred (https://webs.iiitd.edu.in/raghava/toxinpred/multi_submit.php). [e)]Solubility in water; identified using Innovagen's PepCalc (https://pepcalc.com/).

https://doi.org/10.24171/j.phrp.2024.0349

Table 4. Cytokine-inducing potential of the predicted helper T lymphocyte epitopes

Epitopes	TNF-α[a)]	IFN-γ[b)]	IL-4[c)]	IL-10[d)]
GPIIFEKHSRYKYHY	(-)	(+)	(+)	(-)
VGFTAEKAGRLLSSA	(-)	(+)	(-)	(+)
PHVIVDVRQLEPVRL	(-)	(+)	(+)	(-)
ARIAAVRSLAFKAKE	(-)	(+)	(+)	(-)
QPAFAHLRKRGISEA	(-)	(+)	(+)	(+)
TTAISASTASSRTSD	(-)	(+)	(+)	(-)
KASILSNMAVTFKKA	(-)	(+)	(+)	(-)
KIKKVANAVLCALGS	(-)	(+)	(+)	(-)
PIIFDKHAKYKYHYD	(-)	(+)	(+)	(-)
RIDFLVYAEAPDIEK	(-)	(+)	(+)	(+)
RDHFKADSSHIKLSL	(+)	(-)	(+)	(-)

TNF, tumor necrosis factor; IFN, interferon; IL, interleukin.
[a)]Identified using TNFepitope (https://webs.iiitd.edu.in/raghava/tnfepitope). [b)]Identified using IFNepitope (https://webs.iiitd.edu.in/raghava/ifnepitope). [c)]Identified using IL4Pred (http://crdd.osdd.net/raghava/il4pred/). [d)]Identified using IL10Pred (https://webs.iiitd.edu.in/raghava/il10pred/).

Table 5. Physicochemical properties of predicted B lymphocyte epitopes

LBL epitopes	Correct prediction (%)[a)]	Antigenicity score[b)]	Allergenicity[c)]	Toxicity[d)]	Solubility[e)]
DKHAKYKCHYDADYSR	66.4	0.52	Non-allergenic	Non-toxin	Soluble
EIRKRNPDDFQPKGNL	85.8	1.24	Non-allergenic	Non-toxin	Soluble
PSSPKATGRFIIESKD	74.3	0.73	Non-allergenic	Non-toxin	Soluble
PPNGEDATHFKKETKT	70.1	0.52	Non-allergenic	Non-toxin	Soluble
EFTISPNNTPGDILFD	66.3	0.78	Non-allergenic	Non-toxin	Soluble
EAYVNTTDSDFAPATGNTKI	65.1	0.88	Non-allergenic	Non-toxin	Soluble
GFSPTDAARGAVNAPMTK	72.0	0.65	Non-allergenic	Non-toxin	Soluble

LBL, linear B lymphocyte.
[a)]Identified using LBtope server (https://webs.iiitd.edu.in/raghava/lbtope). [b)]Identified using VaxiJen 2.0 (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html); antigenic ≥0.40. [c)]Identified using AllerTOP v2.1 (http://www.ddg-pharmfac.net/AllerTOP). [d)]Identified using ToxinPred (https://webs.iiitd.edu.in/raghava/toxinpred/multi_submit.php). [e)]Identified using Innovagen's PepCalc (https://pepcalc.com/).

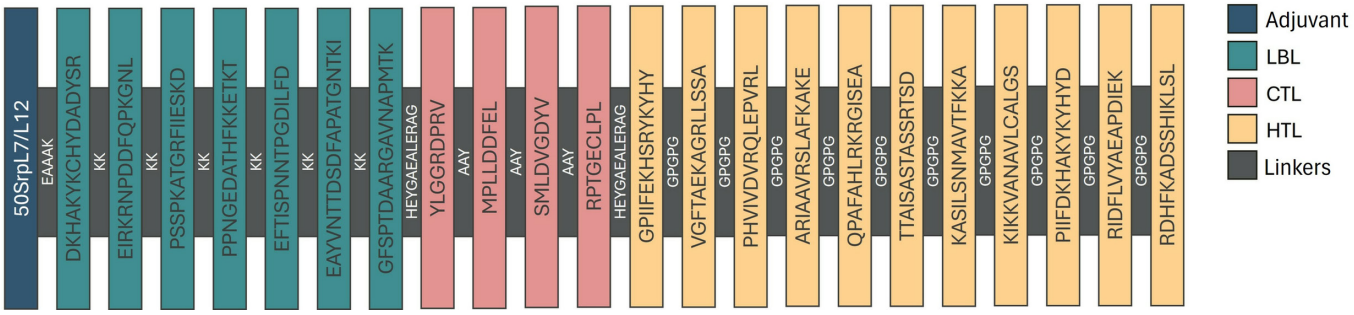


Figure 2. Configuration of the multi-epitope subunit vaccine construct, showing the position of epitopes, linkers, and the adjuvant.
LBL, linear B lymphocyte; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte; 50SrpL7/L12, 50S ribosomal protein L7/L12.

configuration of the constructs, depicting the arrangement of epitopes, linkers, and the adjuvant.

Physicochemical Properties Assessment and Immune Response Simulation

All 3 constructs met the thresholds for antigenicity, non-allergenicity, solubility, and stability (Table 6). The construct adjuvanted with 50SrpL7/L12 demonstrated the most favorable immune simulation profile (Figure 3). As shown in Figure 3, all constructs displayed relatively similar CTL populations, IFN-γ levels, and IL-10 levels. In contrast, the thβd3-adjuvanted construct exhibited the lowest antibody titers, HTL and B

Table 6. Physicochemical properties of the vaccine construct adjuvanted with thβd2, thβd3, and 50SrpL7/L12

Adjuvant	Antigenicity score[a)]	Allergenicity[b)]	Solubility[c)]	Instability index[d)]
thβd2	0.66	Non-allergenic	Soluble	31.4 (stable)
thβd3	0.62	Non-allergenic	Soluble	32.2 (stable)
50SrpL7/L12	0.57	Non-allergenic	Soluble	27.4 (stable)

thβd2, truncated human β-defensin 2; thβd3, truncated human β-defensin 3; 50SrpL7/L12, 50S ribosomal protein L7/L12.
[a)]Identified using VaxiJen 2.0 (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html); antigenic ≥0.40. [b)]Identified using AllerTOP v2.1 (http://www.ddg-pharmfac.net/AllerTOP). [c)]Solubility in water; identified using Innovagen's PepCalc (https://pepcalc.com/). [d)]Identified using Expasy ProtParam (https://web.expasy.org/protparam).

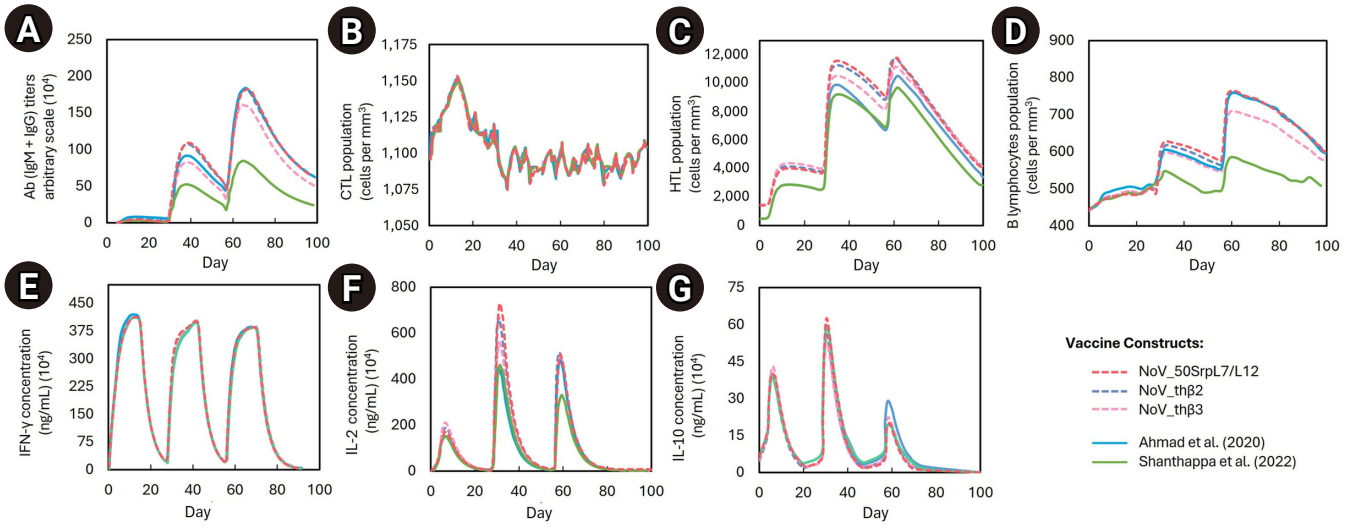


Figure 3. Immune simulation profiles of vaccine constructs differentiated by color. (A) Antibody (Ab) titers; (B) cytotoxic T lymphocyte (CTL) populations; (C) helper T lymphocyte (HTL) populations; (D) B lymphocyte populations; (E) interferon (IFN)-γ populations; (F) interleukin (IL)-2 concentrations; and (G) IL-10 concentrations.
Ig, immunoglobulin; NoV, norovirus; 50SrpL7/L12, 50S ribosomal protein L7/L12; thβd2, truncated human β-defensin 2; thβd3, truncated human β-defensin 3.

lymphocyte populations, and IL-2 levels, rendering it the least favored design. Although the thβd2 and 50SrpL7/L12-adjuvanted constructs showed generally similar immune responses, the latter exhibited relatively higher HTL and B lymphocyte populations following the second immunization, leading to its selection as the final vaccine construct. When compared with the immune simulation profiles of multi-epitope subunit vaccines previously designed by Shanthappa et al. [84] and Ahmad et al. [85], the 50SrpL7/L12-adjuvanted vaccine construct demonstrated remarkably higher antibody titers, HTL and B lymphocyte populations, and IL-2 concentrations. The final vaccine construct has an antigenicity score of 0.54, an instability index of 27.4 (classified as stable), and a GRAVY score of −0.457.

Tertiary Structure Prediction

Figure 4 displays the tertiary structural model of the 50SrpL7/L12-adjuvanted construct. This predicted model achieved high-quality scores in structural quality assessments (Table S4; Figure S1) and was selected for subsequent molecular docking analysis as the best representation of the construct.

Discontinuous B Lymphocyte Epitope Mapping and Disulfide Engineering

A total of 41 residues were identified as discontinuous epitopes using the DiscoTope-3.0 server (Figure 5A), with these epitopes being uniformly distributed across the construct. Additionally, 4 residue pairs suitable for cysteine mutation were identified using DbD2 (Figure 5B). These pairs were substituted with cysteine, and the instability index of the modified construct was analyzed. Unexpectedly, rather than decreasing, the instability index increased from 27.4 to 30.1, possibly due to unfavorable structural constraints or disruption of local interactions. Therefore, the original non-mutated construct was retained for further analysis.

Molecular Docking and Dynamics Simulation

The 50SrpL7/L12-adjuvanted construct exhibited strong

https://doi.org/10.24171/j.phrp.2024.0349

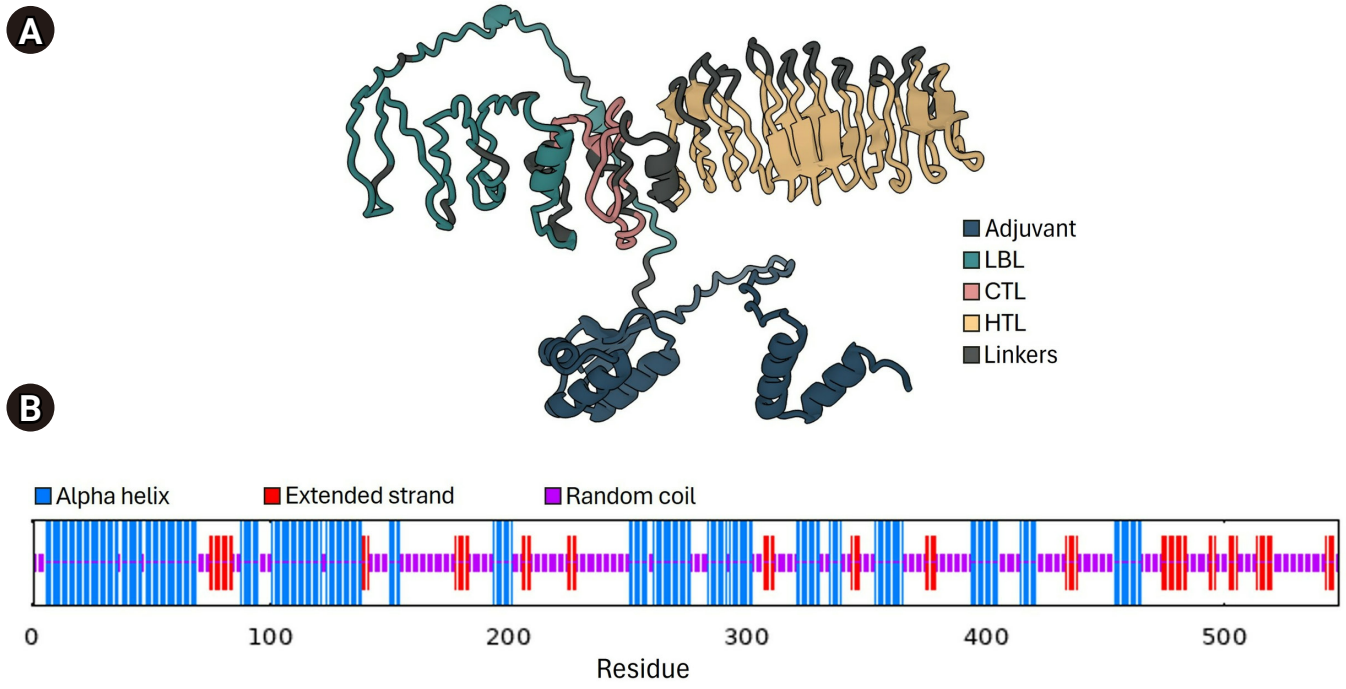


Figure 4. Tertiary structure of the 50S ribosomal protein L7/L12-adjuvanted vaccine construct. (A) Model of the construct. (B) Per-residue secondary structure annotation.
LBL, linear B lymphocyte; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte.

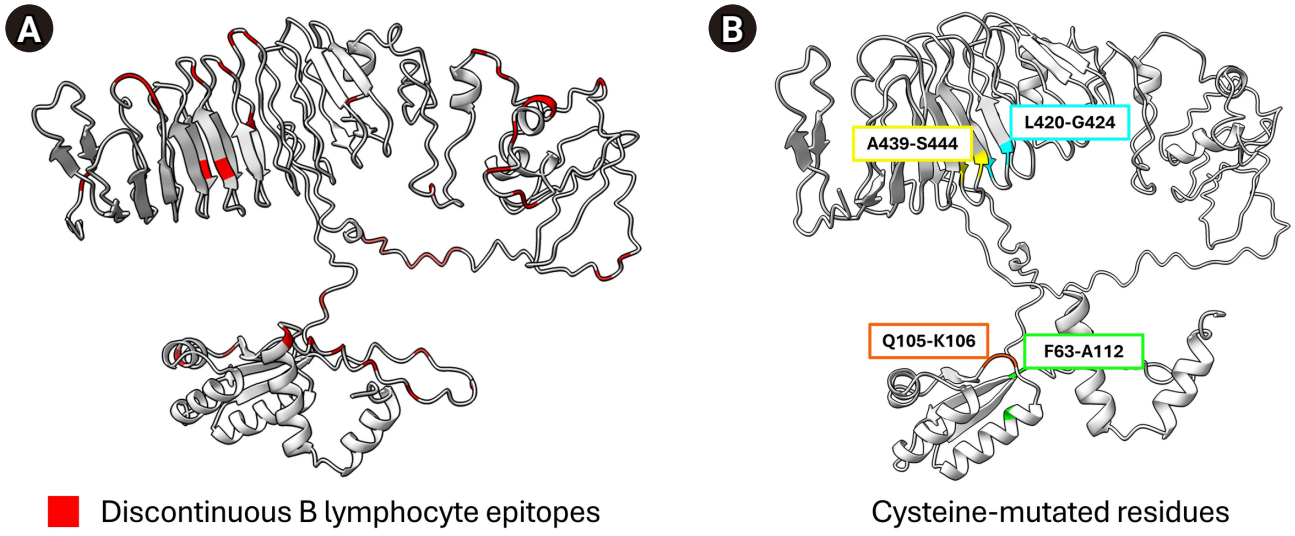


Figure 5. (A) Distribution of discontinuous epitopes in 50SrpL7/L12-adjuvanted vaccine construct identified by DiscoTope-3.0. (B) Residues within the construct identified as suitable candidates for cysteine mutations identified by Disulfide by Design 2.0. 50SrpL7/L12, 50S ribosomal protein L7/L12.

affinity and stable interactions with the immune receptor TLR4 and its co-receptor, myeloid differentiation 2 (MD2), as demonstrated by molecular docking and molecular dynamics simulation analyses. The complex achieved a binding free energy of –145.46 kcal/mol, a value more negative than that observed for previously docked multi-epitope vaccine constructs [86,87]. Moreover, significant residues involved in canonical binding to lipopolysaccharide—such as R264, Y296, K341, and K362 from TLR4, and F121 from MD2—displayed notably high negative ΔG values (Figure 6B). Additionally,

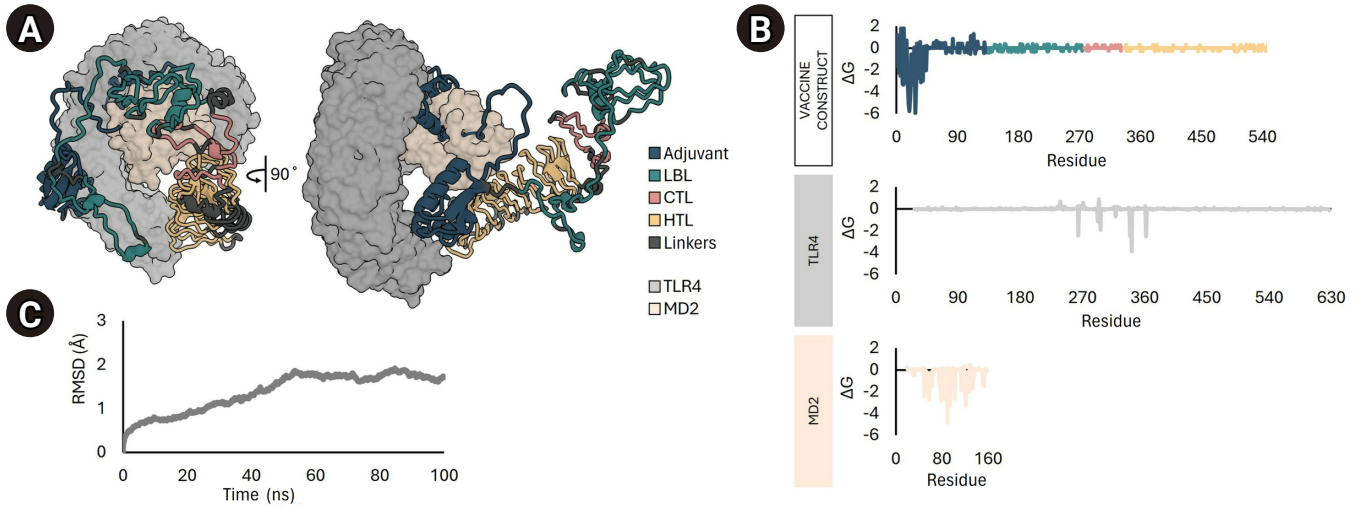


Figure 6. Molecular docking of the 50SrpL7/L12-adjuvanted vaccine construct to the TLR4-MD2 and molecular dynamics simulation results. (A) Model of the complex. (B) Per-residue binding free energy of the complexes. (C) Root-mean-square deviation (RMSD) graph of the complex following 100 ns coarse-grained dynamics simulations.
LBL, linear B lymphocyte; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte; 50SrpL7/L12, 50S ribosomal protein L7/L12; TLR, toll-like receptor; MD, myeloid differentiation.

residues not typically involved in canonical binding, including Y292 in TLR4 and Y58, L87, R90, and P151 from MD2, also interacted with the vaccine construct, exhibiting highly negative ΔG values.

During the molecular dynamics simulation, the complex demonstrated stable interactions, as evidenced by a plateau in RMSD around 40 ns before the simulation's end (Figure 6C). This stability was further confirmed through visual inspection of the simulation trajectory.

Codon Optimization and *In Silico* Cloning

The optimized codon sequence of the construct was generated using JCat for potential expression in the *E. coli* K-12 strain. This sequence achieved a CAI value of 1.0 and a GC content of 51.8%. Figure 7 presents a sample of the cloned construct within the pET28(a)+ vector, with the sequence inserted at the XhoI (158) and BamHI (198) restriction sites.

Discussion

Using an immunoinformatics approach, this study successfully designed a multi-epitope subunit vaccine construct against NoV. Three previously designed multi-epitope subunit vaccines against NoV have been reported by Ahmad et al. [85], Azim et al. [88], and Shanthappa et al. [84]. Compared with these designs, the current study introduces critical advancements that significantly enhance the potential for broad immunological protection. Notably, this study included epitopes from all known human-infecting NoV genogroups, whereas Azim et al. [88] focused solely on GI and GII and Shanthappa et al. [84] restricted their design to the prevalent GII genogroup. The inclusion of GV by Ahmad et al. [85]—a murine genogroup—is not relevant for human vaccine development because murine genogroup V does not infect humans [89]. By targeting all human-infecting genogroups, the study ensures comprehensive coverage of both prevalent and less common genogroups, which is essential for mitigating outbreaks caused by emerging or underrepresented strains.

Another key feature of this study was sourcing epitopes from both capsid and nonstructural proteins. Although Azim et al. [88] and Shanthappa et al. [84] focused exclusively on capsid proteins, this study recognized the potential of nonstructural proteins to contribute to immune activation. Although nonstructural proteins are not initially accessible to receptors during early infection, their release following synthesis and cell lysis allows them to interact with lymphocyte receptors and trigger immune responses [61]. This approach is consistent with previous multi-epitope vaccine designs [61,90–92] that incorporated epitopes from nonstructural proteins.

This study further distinguished itself by prioritizing epitope conservancy. Unlike the studies by Ahmad et al. [85], Azim et al. [88], and Shanthappa et al. [84], which relied on a single protein variant for epitope mapping, this study used conserved fragments obtained from multiple protein variants. This strategy ensures that the designed vaccine effectively targets diverse viral strains and offers a significant

https://doi.org/10.24171/j.phrp.2024.0349

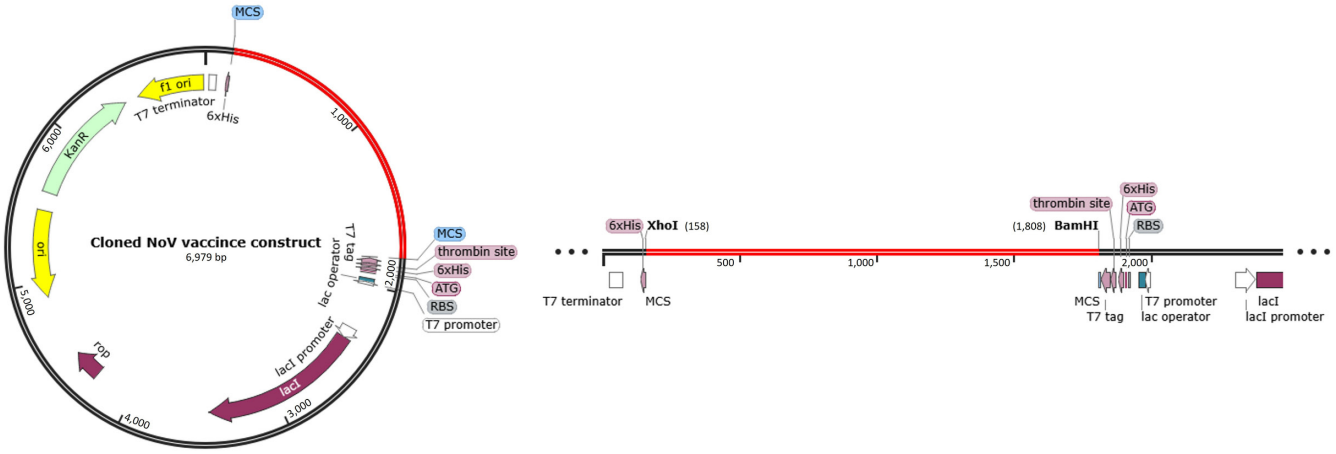


Figure 7. A sample cloned construct for expression in the *Escherichia coli* K-12 strain.

advantage in vaccine coverage. Several multi-epitope vaccine designs have also employed conserved regions for epitope mapping to potentially increase vaccine coverage [93,94].

The study aimed to activate all 3 arms of the immune response—CTL, HTL, and B lymphocyte responses—which each play a critical role during NoV infection [89]. CTL responses were ensured by selecting epitopes that are effectively processed for proteasomal cleavage, transported by the endoplasmic reticulum, bind to MHC I molecules, and are recognized by $CD8^+$ receptors. These predictions were facilitated by NetMHCcons, NetCTLpan, and $CD8^+$ immunogenicity tools. HTL responses were secured by choosing epitopes that bind effectively to MHC II molecules and are recognized by $CD4^+$ receptors, as predicted by NetMHCIIpan and $CD4^+$ immunogenicity tools. In addition, these HTL epitopes induce at least 1 key cytokine associated with NoV infection, such as TNF-α, IFN-γ, IL-4, or IL-10. B lymphocyte responses were confirmed using multiple prediction methods based on amino acid properties and validated epitope databases, ensuring high accuracy. Furthermore, all predicted T and B lymphocyte epitopes met thresholds for antigenicity, non-allergenicity, and solubility. Antigenicity is essential for an epitope to be recognized as foreign, thereby triggering an immune response [95]. Non-allergenicity and non-toxicity assessments are crucial to ensure host safety when the components are processed [91], and sufficient solubility is necessary to prevent aggregation—a common bottleneck in vaccine development [96]. All prediction tools used in this study confirmed that the epitope components were safe, immunogenic, and suitable for vaccine development.

The final vaccine construct incorporated a 50SrpL7/L12 adjuvant to activate innate immunity. This adjuvant is recognized for its role in dendritic cell (DC)-based tumor immunotherapy [53]. It activates TLR signaling by binding to TLR4, which induces both the MyD88 and TRIF signaling pathways, leading to DC activation [53]. Activated DCs, in turn, stimulate naive T lymphocytes, polarize both $CD4^+$ and $CD8^+$ T lymphocytes, secrete IFN-γ, and induce T lymphocyte-mediated cytotoxicity [53]. Molecular docking results demonstrated a strong BA and stable interaction between the 50SrpL7/L12 component and TLR4. BA was estimated using the binding free energy (ΔG) calculated via MM/GBSA analysis, and the highly negative ΔG value indicates strong binding [97].

In contrast to BA analysis, which estimates the strength of the complex based on a single binding pose, the molecular dynamics simulation provided insights into the stability of the complex over time under physiological conditions [98]. The 50SrpL7/L12-adjuvanted vaccine construct docked to TLR4 exhibited a plateauing RMSD, signifying that the structure reached equilibrium during the simulation [99]. Visualization of the simulation further confirmed that the vaccine construct remained bound to TLR4 throughout the simulation period. The combined results from molecular docking and molecular dynamics simulations consistently demonstrated that the 50SrpL7/L12-adjuvanted vaccine exhibited favorable binding characteristics with TLR4. In addition to effective TLR4 binding, the vaccine constructs also displayed favorable physicochemical properties, including high antigenicity, non-allergenicity, solubility, and stability. The presence of widely distributed discontinuous epitopes may further contribute to robust immune recognition. A sample cloned construct was designed for expression in the *E. coli* K-12 strain; however, alternative expression systems may be considered based on production requirements.

The primary limitation of this study is the inherent accuracy

constraint of the predictive models used by the servers. Despite selecting tools with high accuracy, these models can still produce inaccuracies, especially when applied to novel constructs. Future work should focus on experimentally validating these predictions through protein expression in various systems and assessing the construct's stability, solubility, and immunogenicity through *in vitro* and *in vivo* studies. Additionally, exploring multi-epitope designs for other pathogens would broaden the application of this approach. As computational tools improve, refining prediction models and incorporating new data will further increase accuracy.

The computationally designed vaccine construct holds significant potential for clinical applications, provided it is validated through further experimental studies. The construct developed in this study could serve as a strong candidate for future vaccine development, pending successful preclinical and clinical evaluations.

Conclusion

This study successfully designed a multi-epitope subunit vaccine against NoV using multiple immunoinformatics tools. The vaccine design incorporated 22 conserved, non-toxic, non-allergenic, and highly immunogenic T and B lymphocyte epitopes derived from all proteins of the human-infecting genogroups of the virus. The multi-epitope construct adjuvanted with 50SrpL7/L12 exhibited optimal physicochemical properties and a favorable immune simulation profile following a series of sequence- and structure-based assessments. These promising evaluation results suggest that this construct is a potential NoV vaccine candidate. Further *in vitro* and *in vivo* experiments are required to validate these findings.

Supplementary Material

Table S1. Conserved fragments identified from the major viral capsid protein (VP1) of human-infecting genogroups of norovirus; **Table S2.** Conserved fragments identified from the minor viral capsid protein (VP2) of human-infecting genogroups of norovirus; **Table S3.** Conserved fragments identified from the nonstructural polyprotein of human-infecting genogroups of norovirus; **Table S4.** Tertiary structure quality assessment scores of the 50S ribosomal protein L7/L12-adjuvanted vaccine construct structures; **Figure S1.** Tertiary structure quality assessment scores of the 50S ribosomal protein L7/L12-adjuvanted vaccine construct structures in (A) Procheck and (B) ProSA-web. Supplementary data are available at Supplementary data are available at https://doi.org/10.24171/j.phrp.2024.0349.

Notes

Ethics Approval
Not applicable.

Conflicts of Interest
The authors have no conflicts of interest to declare.

Funding
This study was financially supported by the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (PCAARRD).

Availability of Data
The datasets are not publicly available but are available from the corresponding author (F.L.O.) upon reasonable request.

Authors' Contributions
Conceptualization: ZLN, ECB, FLO; Data curation: ZLN, ECB; Formal analysis: ZLN, ECB; Funding acquisition: FLO; Investigation: ZLN, ECB; Methodology: ZLN, ECB, FLO; Project administration: PVS, MEVS, FLO; Supervision: PVS, MEVS, FLO; Visualization: ZLN, ECB; Writing–original draft: ZLN, ECB; Writing–review & editing: all authors. All authors read and approved the final manuscript.

Additional Contributions
Advanced Science and Technology Institute of the Department of Science and Technology (DOST-ASTI) Philippines provided high-performance computing units to run molecular dynamics analyses.

References

1. Burke RM, Hall AJ. Global burden of norovirus. In: Melhem NM, editor. Norovirus. Springer; 2019. p. 1–29.
2. Centers for Disease Control and Prevention (CDC) (US). Norovirus facts and stats [Internet]. CDC; 2024 [cited 2024 Nov 12]. Available from: https://www.cdc.gov/norovirus/data-research/index.html.
3. PLOS (US). The global burden of norovirus & prospects for vaccine development [Internet]. PLOS; 2017 [cited 2024 Nov 12]. Available from: https://collections.plos.org/collection/norovirus/.
4. Bartsch SM, Lopman BA, Ozawa S, et al. Global economic burden of norovirus gastroenteritis. PLoS One 2016;11:e0151219.
5. Carlson KB, Dilley A, O'Grady T, et al. A narrative review of norovirus epidemiology, biology, and challenges to vaccine development. NPJ Vaccines 2024;9:94.
6. de Graaf M, van Beek J, Koopmans MP. Human norovirus transmission and evolution in a changing world. Nat Rev Microbiol 2016;14:421–33.
7. Chhabra P, de Graaf M, Parra GI, et al. Updated classification of norovirus genogroups and genotypes. J Gen Virol 2019;100:1393–406.
8. Cortes-Penfield NW, Ramani S, Estes MK, et al. Prospects and challenges in the development of a norovirus vaccine. Clin Ther 2017;39:1537–49.
9. Kaufman SS, Green KY, Korba BE. Treatment of norovirus infections: moving antivirals from the bench to the bedside. Antiviral Res 2014; 105:80–91.

10. Barclay L, Park GW, Vega E, et al. Infection control for norovirus. Clin Microbiol Infect 2014;20:731–40.
11. Armah G, Lopman BA, Vinje J, et al. Vaccine value profile for norovirus. Vaccine 2023;41(Suppl 2):S134–52.
12. Tripathi T. Advances in vaccines: revolutionizing disease prevention. Sci Rep 2023;13:11748.
13. Laupeze B, Del Giudice G, Doherty MT, et al. Vaccination as a preventative measure contributing to immune fitness. NPJ Vaccines 2021;6:93.
14. Ford-Siltz LA, Tohma K, Parra GI. Understanding the relationship between norovirus diversity and immunity. Gut Microbes 2021;13:1–13.
15. Tan M. Norovirus vaccines: current clinical development and challenges. Pathogens 2021;10:1641.
16. Sanchez-Trincado JL, Gomez-Perosanz M, Reche PA. Fundamentals and methods for T- and B-cell epitope prediction. J Immunol Res 2017;2017:2680160.
17. Sira EM, Banico EC, Fajardo LE, et al. Current strategies, advances, and challenges in multi-epitope subunit vaccine development for African swine fever virus. Vet Integr Sci 2025;23:1–48.
18. Janeway CA, Travers P, Walport M, et al. Immunobiology: the immune system in health and disease. 5th ed. Garland Science; 2001.
19. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007;8:4.
20. Dimitrov I, Bangov I, Flower DR, et al. AllerTOP v.2: a server for in silico prediction of allergens. J Mol Model 2014;20:2278.
21. Magnan CN, Randall A, Baldi P. SOLpro: accurate sequence-based prediction of protein solubility. Bioinformatics 2009;25:2200–7.
22. Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Springer protocols handbooks. Humana Press; 2005. p. 571–607.
23. Gupta S, Kapoor P, Chaudhary K, et al. In silico approach for predicting toxicity of peptides and proteins. PLoS One 2013;8:e73957.
24. Tan LK, Carlone GM, Borrow R. Advances in the development of vaccines against Neisseria meningitidis. N Engl J Med 2010;362:1511–20.
25. Rappuoli R. Reverse vaccinology. Curr Opin Microbiol 2000;3:445–50.
26. Sayers EW, Bolton EE, Brister JR, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res 2022;50:D20–6.
27. Madeira F, Madhusoodanan N, Lee J, et al. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. Nucleic Acids Res 2024;52(W1):W521–5.
28. Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, et al. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res 2008;36(Web Server issue):W35–41.
29. Stewart JJ, Lee CY, Ibrahim S, et al. A Shannon entropy analysis of immunoglobulin and T cell receptor. Mol Immunol 1997;34:1067–82.
30. Karosiene E, Lundegaard C, Lund O, et al. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. Immunogenetics 2012;64:177–86.
31. Weiskopf D, Angelo MA, de Azeredo EL, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci U S A 2013;110:E2046–53.
32. Stranzl T, Larsen MV, Lundegaard C, et al. NetCTLpan: pan-specific MHC class I pathway epitope predictions. Immunogenetics 2010;62:357–68.
33. Kesmir C, Nussbaum AK, Schild H, et al. Prediction of proteasome cleavage motifs by neural networks. Protein Eng 2002;15:287–96.
34. Peters B, Bulik S, Tampe R, et al. Identifying MHC class I epitopes by predicting the TAP transport efficiency of epitope precursors. J Immunol 2003;171:1741–9.
35. Calis JJ, Maybeno M, Greenbaum JA, et al. Properties of MHC class I presented peptides that enhance immunogenicity. PLoS Comput Biol 2013;9:e1003266.
36. Nilsson JB, Kaabinejadian S, Yari H, et al. Accurate prediction of HLA class II antigen presentation across all loci using tailored data acquisition and refined machine learning. Sci Adv 2023;9:eadj6367.
37. Greenbaum J, Sidney J, Chung J, et al. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics 2011;63:325–35.
38. Dhanda SK, Karosiene E, Edwards L, et al. Predicting HLA CD4 immunogenicity in human populations. Front Immunol 2018;9:1369.
39. Dhall A, Patiyal S, Choudhury S, et al. TNFepitope: a webserver for the prediction of TNF-α inducing epitopes. Comput Biol Med 2023;160:106929.
40. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. Biol Direct 2013;8:30.
41. Dhanda SK, Gupta S, Vir P, et al. Prediction of IL4 inducing peptides. Clin Dev Immunol 2013;2013:263952.
42. Nagpal G, Usmani SS, Dhanda SK, et al. Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. Sci Rep 2017;7:42851.
43. Clifford JN, Hoie MH, Deleuran S, et al. BepiPred-3.0: improved B-cell epitope prediction using protein language models. Protein Sci 2022;31:e4497.
44. Yao B, Zhang L, Liang S, et al. SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. PLoS One 2012;7:e45152.
45. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 2006;65:40–8.
46. Singh H, Gupta S, Gautam A, et al. Designing B-cell epitopes for immunotherapy and subunit vaccines. Methods Mol Biol 2015;1348:327–40.
47. Wang QM, Sun SH, Hu ZL, et al. Epitope DNA vaccines against tuberculosis: spacers and ubiquitin modulates cellular immune responses elicited by epitope DNA vaccine. Scand J Immunol 2004;60:219–25.

48. Sarobe P, Lasarte JJ, Larrea E, et al. Enhancement of peptide immunogenicity by insertion of a cathepsin B cleavage site between determinants recognized by B and T cells. Res Immunol 1993;144:257–62.
49. Livingston B, Crimi C, Newman M, et al. A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes. J Immunol 2002;168:5499–506.
50. Dolenc I, Seemuller E, Baumeister W. Decelerated degradation of short peptides by the 20S proteasome. FEBS Lett 1998;434:357–61.
51. Kim J, Yang YL, Jang SH, et al. Human β-defensin 2 plays a regulatory role in innate antiviral immunity and is capable of potentiating the induction of antigen-specific immunity. Virol J 2018;15:124.
52. Funderburg N, Lederman MM, Feng Z, et al. Human -defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2. Proc Natl Acad Sci U S A 2007;104:18631–5.
53. Lee SJ, Shin SJ, Lee MH, et al. A potential protein adjuvant derived from Mycobacterium tuberculosis Rv0652 enhances dendritic cells-based tumor immunotherapy. PLoS One 2014;9:e104351.
54. Behmard E, Abdulabbas HT, Abdalkareem Jasim S, et al. Design of a novel multi-epitope vaccine candidate against hepatitis C virus using structural and nonstructural proteins: an immunoinformatics approach. PLoS One 2022;17:e0272582.
55. Rouzbahani AK, Kheirandish F, Hosseini SZ. Design of a multi-epitope-based peptide vaccine against the S and N proteins of SARS-COV-2 using immunoinformatics approach. Egypt J Med Hum Genet 2022;23:16.
56. Kumar J, Qureshi R, Sagurthi SR, et al. Designing of nucleocapsid protein based novel multi-epitope vaccine against SARS-COV-2 using immunoinformatics approach. Int J Pept Res Ther 2021;27:941–56.
57. Khan A, Khan S, Saleem S, et al. Immunogenomics guided design of immunomodulatory multi-epitope subunit vaccine against the SARS-CoV-2 new variants, and its validation through in silico cloning and immune simulation. Comput Biol Med 2021;133:104420.
58. Tahir Ul Qamar M, Ahmad S, et al. Designing multi-epitope vaccine against Staphylococcus aureus by employing subtractive proteomics, reverse vaccinology and immuno-informatics approaches. Comput Biol Med 2021;132:104389.
59. Shetty S, Dash S, Kumar A, et al. Immunoinformatics design of a multi-epitope vaccine for Chlamydia trachomatis major outer membrane proteins. Sci Rep 2024;14:29919.
60. Banico EC, Sira EM, Fajardo LE, et al. Bridging One Health: computational design of a multi-epitope messenger RNA vaccine for cross-species immunization against Nipah virus. Int J One Health 2024;10:216–29.
61. Banico EC, Sira EM, Fajardo LE, et al. Advancing One Health vaccination: in silico design and evaluation of a multi-epitope subunit vaccine against Nipah virus for cross-species immunization using immunoinformatics and molecular modeling. PLoS One 2024;19:e0310703.
62. Arai R, Ueda H, Kitayama A, et al. Design of the linkers which effectively separate domains of a bifunctional fusion protein. Protein Eng 2001;14:529–32.
63. Guruprasad K, Reddy BV, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng 1990;4:155–61.
64. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–32.
65. Rapin N, Lund O, Bernaschi M, et al. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PLoS One 2010;5:e9862.
66. Castiglione F, Deb D, Srivastava AP, et al. From infection to immunity: understanding the response to SARS-CoV2 through in-silico modeling. Front Immunol 2021;12:646972.
67. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596:583–9.
68. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera: a visualization system for exploratory research and analysis. J Comput Chem 2004; 25:1605–12.
69. Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res 2013;41(Web Server issue):W384–8.
70. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 1993;2:1511–9.
71. Laskowski RA, MacArthur MW, Moss DS, et al. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–91.
72. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35(Web Server issue):W407–10.
73. Pettersen EF, Goddard TD, Huang CC, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci 2021;30:70–82.
74. Hoie MH, Gade FS, Johansen JM, et al. DiscoTope-3.0: improved B-cell epitope prediction using inverse folding latent representations. Front Immunol 2024;15:1322712.
75. Craig DB, Dombkowski AA. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics 2013; 14:346.
76. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. Nucleic Acids Res 2000;28:235–42.
77. Schrodinger LL. The PyMOL Molecular Graphics System, version 1.8. Scientific Research Publishing; 2015.
78. Webb B, Sali A. Comparative protein structure modeling using MODELLER. Curr Protoc Bioinformatics 2016;54:5.6.1–5.6.37.
79. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein-protein docking. Nat Protoc 2017;12:255–78.
80. Zhou P, Jin B, Li H, et al. HPEPDOCK: a web server for blind peptide-protein docking based on a hierarchical algorithm. Nucleic Acids Res 2018;46(W1):W443–50.
81. Klein F, Sonora M, Helene Santos L, et al. The SIRAH force field: a

https://doi.org/10.24171/j.phrp.2024.0349

suite for simulations of complex biological systems at the coarse-grained and multiscale levels. J Struct Biol 2023;215:107985.
82. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1:19–25.
83. Grote A, Hiller K, Scheer M, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 2005;33(Web Server issue):W526–31.
84. Shanthappa PM, Suravajhala R, Suravajhala P, et al. In silico based multi-epitope vaccine design against norovirus. J Biomol Struct Dyn 2023;41:5696–706.
85. Ahmad I, Ali SS, Zafar B, et al. Development of multi-epitope subunit vaccine for protection against the norovirus' infections based on computational vaccinology. J Biomol Struct Dyn 2022;40:3098–3109.
86. Kar PP, Araveti PB, Kuriakose A, et al. Design of a multi-epitope protein as a subunit vaccine against lumpy skin disease using an immunoinformatics approach. Sci Rep 2022;12:19411.
87. Pang F, Long Q, Liang S. Designing a multi-epitope subunit vaccine against Orf virus using molecular docking and molecular dynamics. Virulence 2024;15:2398171.
88. Azim KF, Hasan M, Hossain MN, et al. Immunoinformatics approaches for designing a novel multi epitope peptide vaccine against human norovirus (Norwalk virus). Infect Genet Evol 2019;74:103936.
89. van Loben Sels JM, Green KY. The antigenic topology of norovirus as defined by B and T cell epitope mapping: implications for universal vaccines and therapeutics. Viruses 2019;11:432.
90. Fajardo LE, Banico EC, Sira EM, et al. Computational multi-epitope based design of a multivalent subunit vaccine against co-infecting African swine fever virus and porcine circovirus type 2. Vet Integr Sci 2025;23:1–30.
91. Simbulan AM, Banico EC, Sira EM, et al. Immunoinformatics-guided approach for designing a pan-proteome multi-epitope subunit vaccine against African swine fever virus. Sci Rep 2024;14:1354.
92. Ullah A, Rehman B, Khan S, et al. An in silico multi-epitopes vaccine ensemble and characterization against nosocomial Proteus penneri. Mol Biotechnol 2024;66:3498–513.
93. Ahmad S, Demneh FM, Rehman B, et al. In silico design of a novel multi-epitope vaccine against HCV infection through immunoinformatics approaches. Int J Biol Macromol 2024;267(Pt 2):131517.
94. Farhani I, Yamchi A, Madanchi H, et al. Designing a multi-epitope vaccine against the SARS-CoV-2 variant based on an immunoinformatics approach. Curr Comput Aided Drug Des 2024;20:274–90.
95. Fishman J, Wiles K, Wood K. The acquired immune system response to biomaterials, including both naturally occurring and synthetic biomaterials. In: Badylak SF, editor. Host response to biomaterials. Academic Press; 2015. p. 151–87.
96. Rosace A, Bennett A, Oeller M, et al. Automated optimisation of solubility and conformational stability of antibodies and proteins. Nat Commun 2023;14:1937.
97. Schon A, Madani N, Smith AB, et al. Some binding-related drug properties are dependent on thermodynamic signature. Chem Biol Drug Des 2011;77:161–5.
98. Sinha S, Tam B, Wang SM. Applications of molecular dynamics simulation in protein study. Membranes (Basel) 2022;12:844.
99. Macalalad MA, Orosco FL. In silico identification of multi-target inhibitors from medicinal fungal metabolites against the base excision repair pathway proteins of African swine fever virus. RSC Adv 2024; 14:10039–55.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0036
Osong Public Health Res Perspect 2025;16(3):252-260

Original Article

Metabolic and systemic inflammation status in rheumatoid arthritis—fasting blood glucose as a primary predictor of rheumatoid arthritis risk: a cross-sectional study in Iran

Sorayya Kheirouri[1], **Mohammad Alizadeh**[1,2], **Arash Tandorost**[1]

[1]Department of Nutrition, Faculty of Nutrition and Food Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
[2]Nutrition Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Received: February 9, 2025
Revised: April 12, 2025
Accepted: April 21, 2025

Corresponding author:
Sorayya Kheirouri
Department of Nutrition, Faculty of Nutrition and Food Sciences, Tabriz University of Medical Sciences, Attar Nishabouri St., Tabriz, I. R. Postal code: 5166614711, POBOX: 14711, Iran
E-mail: Kheirouris@tbzmed.ac.ir

Co-Corresponding author:
Mohammad Alizadeh
Department of Nutrition, Faculty of Nutrition and Food Sciences, Tabriz University of Medical Sciences, Attar Nishabouri St., Tabriz, I. R. Postal code: 5166614711, POBOX: 14711, Iran
E-mail: mdalizadeh@tbzmed.ac.ir

ABSTRACT

Objectives: This study investigated the relationship between metabolic factors (blood lipids and glucose) and inflammatory indicators (tumor necrosis factor-alpha [TNF-α] and high-sensitivity C-reactive protein [hs-CRP]), disease activity, and the rheumatoid arthritis (RA) risk.
Methods: Serum fasting blood glucose (FBG) and lipid profiles—including total cholesterol (Chol), triglycerides (TG), high-density lipoprotein (HDL), and low-density lipoprotein—were measured in 100 RA patients and 100 healthy individuals. Disease severity was assessed using the disease activity score 28. Inflammatory indicators (TNF-α and hs-CRP) were measured using the enzyme-linked immunosorbent assay method.
Results: In RA patients, serum FBG, TG, Chol/HDL, and TG/HDL were significantly elevated, whereas HDL levels reduced compared to healthy individuals. Multivariate analysis indicated that each unit increase in serum FBG, HDL, Chol/HDL, and TG/HDL was associated with a 64% increase ($p < 0.001$), a 7% reduction ($p = 0.001$), a 52% increase ($p = 0.007$), and a 54% increase ($p = 0.001$) in the odds of RA, respectively. Disease activity showed no correlation with metabolic factors ($p > 0.05$). Among all metabolic factors studied, FBG had the largest area under the curve (0.981) ($p < 0.0001$) for predicting RA. Across the total participant group, FBG, TG, and TG/HDL were positively associated with hs-CRP and TNF-α ($p < 0.05$). HDL showed an inverse association with hs-CRP ($p = 0.008$). Among RA patients specifically, TNF-α positively correlated with TG and TG/HDL, while hs-CRP correlated only with TG/HDL.
Conclusion: These findings indicate that increased FBG and Chol/HDL and decreased HDL may elevate RA risk by promoting systemic inflammation. Among these, elevated FBG may serve as the strongest predictor of RA risk.

Keywords: Arthritis, rheumatoid; Blood glucose; Inflammation; Lipids

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune disorder characterized by inflammation, affecting approximately 0.5% to 1% of the global population, with increased prevalence among individuals aged over 60 years [1]. The primary cause of RA is inflammation of the synovial membrane, leading to cartilage and bone degradation, joint damage, and impaired joint function [2–4]. RA typically presents with debilitating joint pain, morning stiffness, and swelling [5]. Additionally, extra-articular manifestations may involve other organ systems, including renal, cardiovascular, ocular, and respiratory systems [6]. Besides the burden of the disease itself, RA patients frequently develop metabolic abnormalities such as type 2 diabetes, hypertension, arteriosclerosis, and osteoporosis, significantly increasing cardiovascular risk [7], mortality, and disability rates [8]. Although RA's exact etiology remains unclear, genetic predisposition, autoimmune responses, environmental influences, and gut microbiota are believed to play critical roles in disease development [9,10].

Previous studies have reported disturbances in metabolic factors, including fasting blood glucose (FBG) and lipid profiles, among RA patients [11–13]. Dyslipidemia observed in RA differs from the general population, typically involving reduced total cholesterol (Chol) and low high-density lipoprotein (HDL) [14,15]. An imbalance in HDL and Chol leads to increased atherogenicity (elevated Chol/HDL ratio), an important cardiovascular prognostic marker. Previously, Chol, HDL, and low-density lipoprotein (LDL) levels were thought to decrease during active, untreated RA [16–18]. Recent studies have supported these findings, reporting decreased HDL and LDL [17–21] and increased Chol/HDL ratios [17,19] in RA patients.

Changes in blood lipids have also been linked with inflammatory markers in RA [22,23]. Recent investigations have further proposed that glucose metabolism disruption in RA might influence inflammatory properties of rheumatoid synovial fibroblasts. This metabolic alteration could significantly contribute to autoimmune disease pathogenesis by affecting immune responses and altering autoantigen expression [24].

Existing studies have frequently attributed elevated FBG levels to RA itself. However, studies addressing the predictive roles of blood glucose and lipid profiles regarding RA risk remain scarce. Examining lipid and glucose metabolism and their metabolic pathways may enhance our understanding of their roles in RA onset and persistence, as well as other autoimmune disorders. Thus, this study was designed to explore relationships between metabolic factors (blood lipids and glucose), inflammatory markers (including tumor necrosis factor-alpha [TNF-α] and high-sensitivity C-reactive protein [hs-CRP]), disease activity, and RA risk.

HIGHLIGHTS

- Metabolic disturbances are common in patients with rheumatoid arthritis (RA) and are often attributed to the disease itself.
- Few studies have explored whether metabolic disturbances might contribute to RA development.
- Elevated fasting blood glucose (FBG), cholesterol-to-high-density lipoprotein ratio, triglyceride-to-high-density lipoprotein ratio, and reduced high-density lipoprotein ratio may increase RA risk by promoting inflammation.
- Increased FBG might be the strongest predictor of RA.

Materials and Methods

Study Design and Setting

In this cross-sectional study, participants were recruited consecutively from patients referred to the rheumatology clinic of Imam Khomeini Hospital in Urmia City, from January 2019 to February 2020.

Participants

One hundred newly diagnosed RA patients aged 20 to 50 years, comprising both males and females, were enrolled in the study. A control group consisting of 100 healthy participants was also included. These individuals were selected from people visiting the clinic for routine health check-ups. The sample size determination was based on previously published studies [25]. Convenience sampling was employed for participant selection. Participants in both groups were balanced for age, sex, and body mass index (BMI). Exclusion criteria included pregnant and lactating women, individuals with other chronic diseases such as diabetes, and those following special diets.

Variables

The primary outcomes of the study were the risk and severity of RA. Exposure variables were metabolic factors (including blood lipids and glucose) and inflammatory markers (including TNF-α and hs-CRP). Potential confounders included age, sex, BMI, and energy intake.

Data Sources/Measurement

RA diagnosis was confirmed by a rheumatologist according to the American College of Rheumatology/European League Against Rheumatism criteria [26]. The severity of RA was evaluated using the disease activity score of 28 joints (DAS-28) [27]. DAS-28 incorporates tender and swollen joint counts (TJC and SJC) of 28 joints, erythrocyte sedimentation rate (mm/h) or C-reactive protein (CRP, mg/L) levels, and general health assessed via visual analogue scale (VAS). DAS-28 (CRP) scores were calculated as follows:

$$DAS{-}28(CRP) = 0.56 \times \sqrt{TJC28} + 0.28 \times SJC28 + 0.70 \times Ln(CRP) + 0.014 \times VAS$$

DAS-28 scores were categorized into disease activity levels as follows: remission (≤2.6), low activity (>2.6 to ≤3.2), moderate activity (>3.2 to ≤5.1), and high activity (>5.1).

Participants' body weight was measured using a SECA scale (Germany, model 769) to the nearest 0.1 kg, with participants wearing minimal clothing and no footwear. Height was measured to the nearest 0.5 cm. BMI was calculated as weight (kg) divided by height squared (m^2).

After 12 hours of overnight fasting, 5 mL blood samples were collected. Serum was separated by centrifugation at 3,500 rpm for 10 minutes at 4 °C and stored at −70 °C for subsequent analysis. Levels of total Chol, HDL, and triglycerides (TG) were quantified using a biochemical auto-analyzer and enzymatic kits (Pars Azmoon). LDL levels were calculated using the Friedewald equation [28]. FBG concentrations were measured using the glucose oxidase technique with a commercial kit (Pars Azmoon).

Serum levels of TNF-α were measured using an enzyme-linked immunosorbent assay kit from Glory Science Co., Ltd. Additionally, hs-CRP levels were quantified using an immunoturbidimetric assay (Pars Azmoon).

Bias

To minimize potential bias, participants in both groups were matched according to age, sex, and BMI.

Statistical Methods

Data analysis was conducted using IBM SPSS ver. 21 software (IBM Corp.). The normality of data was evaluated using skewness and kurtosis tests. Qualitative results were presented as frequency (%), and quantitative results as mean ± standard deviation. For comparisons between the 2 groups, the independent-samples t-test (for continuous variables) or the chi-square test (for categorical variables) was employed. Binary logistic regression analysis was conducted to estimate odds ratios (ORs) for RA disease risk, in both univariate and multivariate models. Linear regression analysis was used to evaluate associations between serum metabolic factors and disease activity, as well as inflammatory markers. Age, sex, BMI, and energy intake were considered as covariates. Receiver operating characteristic (ROC) curves were used to assess and compare the predictive ability of serum metabolic factors for RA disease. A p-value <0.05 was considered statistically significant in all analyses.

Ethics Statement

The study protocol was approved by the Institutional Review Board (IRB) of Tabriz University of Medical Sciences (No. IR.TBZMED.REC.1403.045, IRB No: 73560). The requirement for informed consent was waived by the IRB.

Results

Participants

The study included 100 subjects in the RA group and 99 healthy individuals in the control group. One healthy participant was excluded from the study due to abnormal serum results.

Descriptive Data

As presented in Table 1, no significant differences were observed between the RA patients and healthy individuals

Table 1. Demographic characteristics of participants

Characteristic	RA group (n = 100)	Healthy group (n = 99)	p
Age (y)	43.00 (33.00–60.00)	44.00 (32.00–60.00)	0.37
Sex			0.75
Male	35 (35.0)	45 (45.0)	
Female	65 (65.0)	55 (55.0)	
Smoking	10 (10.0)	15 (15.0)	0.43
Body mass index (kg/m²)	27.04 (19.93–40.79)	25.22 (19.47–34.29)	0.19
Energy intake	2,222.219 ± 714.337	2,101.012 ± 529.661	0.17

Data are presented as median (minimum–maximum), n (%), or mean ± standard deviation.
RA, rheumatoid arthritis.

https://doi.org/10.24171/j.phrp.2025.0036

regarding age, sex, smoking status, BMI, and total energy intake ($p > 0.05$). According to Table 2, serum levels of FBG ($p < 0.001$), TG ($p = 0.013$), Chol/HDL ($p = 0.003$), and TG/HDL ($p < 0.001$) were significantly higher in RA patients compared to healthy participants, whereas HDL levels were significantly lower ($p < 0.001$). Levels of total Chol and LDL did not differ significantly between the 2 groups ($p > 0.05$). Serum TNF-α and hs-CRP levels were significantly elevated in RA patients.

Main Results

As presented in Table 3, increased serum FBG and Chol/HDL significantly elevated the odds of RA disease in the multivariate model. Specifically, each unit increase in serum FBG was associated with a 64% increase in the odds of RA (OR, 1.639; 95% confidence interval [CI], 1.381–1.946; $p < 0.001$). Each 1-unit increase in Chol/HDL corresponded to a 52% increase in the odds of RA (OR, 1.519; 95% CI, 1.119–2.061; $p = 0.007$), and each 1-unit increase in TG/HDL corresponded to a 54% increase (OR, 1.539; 95% CI, 1.203–1.968; $p = 0.001$). Conversely, increased HDL was protective, with each 1-unit increase in serum HDL associated with approximately a 7% reduction in RA odds (OR, 0.926; 95% CI, 0.886–0.968; $p = 0.001$), both in univariate and multivariate analyses. No significant associations were found between RA odds and SBP, DBP, total Chol, TG, or LDL serum levels ($p > 0.05$).

As shown in Table 4, none of the metabolic factors studied were significantly correlated with disease activity ($p > 0.05$). Among all participants, serum levels of FBG (β = 0.399, $p < 0.001$), TG (β = 0.164, $p = 0.024$), and TG/HDL (β = 0.256, $p < 0.001$) were positively associated with hs-CRP. Additionally, serum levels of FBG (β = 0.253, $p = 0.001$), TG (β = 0.314, $p < 0.001$), and TG/HDL (β = 0.295, $p < 0.001$) were positively associated with TNF-α. Serum HDL levels showed an inverse association with hs-CRP (β = −0.199, $p = 0.008$).

In RA patients specifically, TNF-α correlated positively with TG (β = 0.312, $p = 0.001$) and TG/HDL (β = 0.248, $p = 0.011$), but not with other metabolic factors ($p > 0.05$). In this group, hs-CRP correlated only with TG/HDL (β = 0.203, $p = 0.047$). In healthy individuals, neither TNF-α nor hs-CRP significantly correlated with metabolic factors ($p > 0.05$) (data not shown).

ROC curve analysis showed a considerable predictive role for serum FBG regarding RA disease (Figure 1). FBG exhibited the highest area under the curve (AUC = 0.981, $p < 0.001$), indicating strong predictive capacity, whereas HDL had the lowest predictive capacity (AUC = 0.337). The AUC values for

Table 2. Comparison of metabolic factors between groups

	Healthy ($n = 99$)	RA ($n = 100$)	p
FBG	90.12 ± 7.175	110.80 ± 7.665	< 0.001
TG	158.828 ± 37.114	175.980 ± 56.986	0.013
Chol	171.840 ± 28.893	175.80 ± 25.476	0.306
HDL	46.252 ± 7.419	42.190 ± 7.658	< 0.001
LDL	127.545 ± 24.975	140.070 ± 85.370	0.162
Chol/HDL	3.848 ± 1.102	4.308 ± 1.047	0.003
TG/HDL	3.535 ± 1.051	4.380 ± 1.880	< 0.001

Data are presented as mean ± standard deviation.
RA, rheumatoid arthritis; FBG, fasting blood glucose; TG, triglyceride; Chol, cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Table 3. Correlations between metabolic factors and odds of rheumatoid arthritis

	Univariate model		Multivariate model	
	OR (95% CI)	p	OR (95% CI)	p
FBG	1.610 (1.364–1.900)	< 0.001	1.639 (1.381–1.946)	< 0.001
TG	1.008 (1.001–1.014)	0.015	1.008 (1.001–1.014)	0.020
Chol	1.005 (0.995–1.016)	0.305	1.006 (0.995–1.016)	0.281
HDL	0.929 (0.892–0.968)	< 0.001	0.926 (0.886–0.968)	0.001
LDL	1.004 (0.998–1.011)	0.205	1.005 (0.998–1.012)	0.178
Chol/HDL	1.531 (1.144–2.050)	0.004	1.519 (1.119–2.061)	0.007
TG/HDL	1.505 (1.201–1.887)	< 0.001	1.539 (1.203–1.968)	0.001

Age, sex, body mass index, and energy intake were considered as confounders in the multivariate analysis.
OR, odds ratio; CI, confidence interval; FBG, fasting blood glucose; TG, triglyceride; Chol, cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Table 4. Correlations between metabolic factors and disease activity (in patients with rheumatoid arthritis) and inflammatory markers (total participants)

	Disease activity		hs-CRP		TNF-α	
	β	*p*	β	*p*	β	*p*
FBG	0.102	0.292	0.399	<0.001	0.253	0.001
TG	0.167	0.082	0.164	0.024	0.314	<0.001
Chol	−0.050	0.600	−0.028	0.699	0.010	0.892
HDL	−0.046	0.634	−0.199	0.008	−0.102	0.176
LDL	−0.020	0.832	0.003	0.963	0.086	0.234
Chol/HDL	0.001	0.994	0.113	0.122	0.071	0.338
TG/HDL	0.130	0.177	0.256	<0.001	0.295	<0.001

hs-CRP, high-sensitivity C-reactive protein; TNF-α, tumor necrosis factor-alpha; FBG, fasting blood glucose; TG, triglyceride; Chol, cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

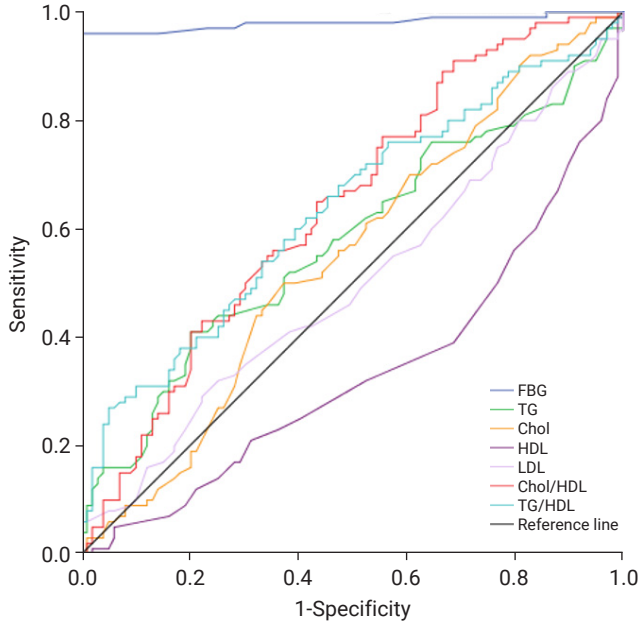


Figure 1. Receiver operating characteristic curve showing sensitivity and specificity to compare the predictive power of various metabolic factors for rheumatoid arthritis.
FBG, fasting blood glucose; TG, triglyceride; Chol, cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

TG, total Chol, LDL, Chol/HDL, and TG/HDL were 0.577, 0.548, 0.498, 0.641, and 0.635, respectively (Table 5).

Table 5. AUC of metabolic factors for predicting rheumatoid arthritis

	AUC (95% CI)	*p*
FBG	0.981 (0.959–1.000)	<0.0001
TG	0.577 (0.497–0.657)	0.060
Chol	0.548 (0.467–0.628)	0.246
HDL	0.337 (0.261–0.412)	<0.0001
LDL	0.498 (0.417–0.579)	0.960
Chol/HDL	0.641 (0.565–0.718)	0.001
TG/HDL	0.635 (0.558–0.712)	0.001

AUC, area under the curve; CI, confidence interval; FBG, fasting blood glucose; TG, triglyceride; Chol, cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Discussion

The present study found that patients with RA had higher serum levels of FBG, TG, Chol/HDL, and TG/HDL, and lower HDL levels compared with healthy individuals. Elevated serum levels of FBG, Chol/HDL, and TG/HDL were associated with increased odds of having RA, whereas higher HDL levels reduced the odds. Among all participants, FBG, TG, and TG/HDL showed positive associations with hs-CRP and TNF-α, while HDL was inversely correlated with hs-CRP. Within the RA group, TNF-α correlated positively with TG and TG/HDL, whereas hs-CRP correlated only with TG/HDL. Among metabolic factors studied, FBG showed the strongest predictive power for RA.

The current study highlights the potential role of elevated FBG levels in RA development. Extensive evidence suggests that RA is associated with increased diabetes mellitus risk. Ruscitti et al. [29], in a study involving 500 RA patients and 500 matched controls, observed a higher prevalence of type 2 diabetes and abnormal fasting glucose levels among RA patients. Similarly, Ali et al. [30], in research with 70 females with early RA and 35 healthy controls, reported significantly elevated insulin resistance and FBG concentrations in the RA group. Another single-center study with 90 RA patients and 37 matched controls also identified significantly greater insulin resistance in RA individuals [31]. Furthermore, a high prevalence of diabetes has been reported in patients with RA [32]. Current studies have often compared serum FBG between RA patients and healthy

https://doi.org/10.24171/j.phrp.2025.0036

controls [29,30,33], attributing elevated FBG levels to RA and suggesting inflammation-induced insulin resistance as a causal factor for increased blood glucose [34]. However, research examining whether elevated blood glucose itself independently increases RA risk remains limited.

Elevated FBG levels may increase RA risk through several mechanisms. High glucose concentrations could promote immune dysregulation, contributing to autoantibody production and synovial inflammation [35]. Hyperglycemia also leads to chronic inflammation, a central aspect of RA pathogenesis. Specifically, high glucose stimulates the release of inflammatory mediators such as TNF-α and IL-6, involved in joint inflammation and damage [36,37]. The positive relationship between FBG and inflammatory markers observed in this study further corroborates these mechanisms. Additionally, hyperglycemia-induced oxidative stress may worsen inflammatory responses, potentially triggering autoimmune reactions against joint tissues [38]. Collectively, these findings underscore the potential importance of controlling blood glucose levels to mitigate RA risk.

In our study, lower HDL levels were associated with increased RA risk. Dyslipidemia has frequently been observed in RA patients. Erum et al. [39], investigating 200 RA patients, found dyslipidemia in 53.5% and low HDL in 41.5%. Similarly, Lakatos and Harsagyi [40], in a study including 129 RA patients and 1,374 controls, reported significantly lower HDL levels in the RA group. Likewise, Dursunoglu et al. [41], studying 87 women with RA and 50 healthy women, observed lower HDL levels among the RA participants. Associations between RA severity and HDL functionality indicators have also been noted [42]. HDL possesses anti-inflammatory and antioxidant properties, potentially mitigating the chronic inflammation characteristic of RA [43]. HDL particles inhibit adhesion molecule expression on endothelial cells, reducing inflammatory cell recruitment into joint tissues [44]. Furthermore, HDL neutralizes oxidized lipids and decreases inflammatory cytokine release, such as TNF-α and IL-6 [45]. RA patients often exhibit reduced HDL levels, impairing HDL's anti-inflammatory functions [46]. The inflammatory score in RA patients depends not only on plasma HDL levels but also on HDL particle structure, composition, and functionality [47]. The inverse correlation between HDL and inflammatory markers noted in this study aligns with these earlier reports. Certain biological therapies, such as tofacitinib, may also improve disease activity and functional status in RA patients experiencing lipid alterations [48]. These mechanisms collectively suggest a protective role of HDL in RA prevention and management, highlighting the importance of maintaining healthy HDL levels.

In this study, higher Chol/HDL ratios correlated with an increased RA risk. Elevated Chol/HDL ratios, indicative of higher cardiovascular risk, are frequently observed in RA patients [19,41]. This scenario, termed the “lipid paradox,” involves lower total Chol and LDL-Chol levels in RA patients alongside paradoxically elevated cardiovascular disease risk [49]. Chronic inflammation in RA, reflected by increased CRP and TNF-α levels, disrupts lipid metabolism, reducing total Chol, LDL, and HDL levels; however, this reduction does not equate to decreased cardiovascular risk [50,51]. Despite lower Chol concentrations, the atherogenic index (Chol/HDL ratio) may remain elevated, indicating greater atherosclerosis risk [15]. Thus, the higher Chol/HDL ratio observed in RA patients highlights the importance of careful cardiovascular risk management in this group.

The present study identified higher TG levels as associated with increased RA risk. Elevated TG concentrations have been implicated in RA pathogenesis [19,52]. Hypertriglyceridemia can aggravate systemic inflammation characteristic of RA [53]. Increased TG levels facilitate the production of pro-inflammatory cytokines [53,54], exacerbating joint inflammation and damage. The direct correlation between TG and inflammatory markers observed in this research further supports these previous findings. Thus, monitoring lipid profiles, particularly TG levels, is essential as part of comprehensive RA management.

The integration of antioxidants, complementary therapies, and dietary modifications plays a crucial role in managing RA, particularly through modulating lipid profiles, glycemic factors, and inflammatory processes. Pycnogenol, as an antioxidant, has demonstrated effectiveness in reducing oxidative stress and stabilizing inflammatory biomarkers [55], potentially aiding lipid metabolism regulation. In addition, propolis exhibits significant anti-inflammatory properties, offering therapeutic benefits for individuals with RA [56]. Royal jelly has also been reported to enhance glycemic control among individuals with compromised health [57]. Moreover, conjugated linoleic acids have shown the ability to reduce inflammation and oxidative stress while simultaneously improving physical performance and metabolic parameters [58]. Taurine, flaxseed oil, and *Gundelia tournefortii* supplementation might further improve glycemic control and lipid metabolism [59–61]. Collectively, these findings emphasize the multifaceted benefits of integrating antioxidants, complementary therapies, and dietary interventions as adjunctive approaches to conventional RA treatment strategies.

Regarding clinical and therapeutic implications, the findings from this study suggest that routine monitoring and management of FBG, TG, and HDL levels should become integral

components of care for individuals at risk of or diagnosed with RA. Clinicians should consider strategies such as dietary modifications, increased physical activity, and pharmacological interventions to control blood glucose and lipid levels, thereby reducing systemic inflammation. Addressing these metabolic factors could reduce both the incidence and severity of RA, ultimately enhancing patient outcomes and quality of life. This highlights the importance of adopting a comprehensive, multidisciplinary approach in RA management.

A major limitation of the current study is the inherent potential for selection bias associated with case-control studies. The selected cases and controls may not precisely represent the broader population, limiting the ability to establish causal relationships. Further investigations are necessary to clarify exact mechanisms and determine optimal therapeutic strategies. Large-scale prospective cohort studies would be beneficial in confirming cause-and-effect relationships between metabolic factors and RA.

Conclusion

The findings highlight that elevated FBG, Chol/HDL, TG/HDL, and reduced HDL may increase RA risk by promoting systemic inflammation. Among these metabolic factors, elevated FBG might be the strongest predictor of RA risk.

Notes

Ethics Approval

This study was approved by the ethical committee of Tabriz University of Medical Sciences (No. IR.TBZMED.REC.1403.045, IRB No: 73560) and performed in accordance with the principles of the Declaration of Helsinki. The requirement for informed consent was waived because of the retrospective nature of this study.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Funding

This study was supported by Tabriz University of Medical Sciences.

Availability of Data

All data generated or analyzed during this study are included in this published article. For other data, these may be requested through the corresponding author.

Authors' Contributions

Conceptualization: SK, MA; Data curation: AT; Formal analysis: SK, MA; Funding acquisition: SK, MA; Investigation: SK; Methodology: SK; Project administration: SK, MA; Software: SK; Supervision: SK, MA; Validation: SK; Visualization: SK; Writing–original draft: SK; Writing–review & editing: all authors. All authors read and approved the final manuscript.

References

1. Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: a review. JAMA 2018;320:1360–72.
2. Gibofsky A. Epidemiology, pathophysiology, and diagnosis of rheumatoid arthritis: a synopsis. Am J Manag Care 2014;20(7 Suppl): S128–35.
3. Guo Q, Wang Y, Xu D, et al. Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. Bone Res 2018; 6:15.
4. Symmons D, Turner G, Webb R, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford) 2002;41:793–800.
5. Sharif K, Sharif A, Jumah F, et al. Rheumatoid arthritis in review: clinical, anatomical, cellular and molecular points of view. Clin Anat 2018;31:216–23.
6. Sparks JA. Rheumatoid arthritis. Ann Intern Med 2019;170:ITC1–6.
7. Kerekes G, Nurmohamed MT, Gonzalez-Gay MA, et al. Rheumatoid arthritis and metabolic syndrome. Nat Rev Rheumatol 2014;10:691–6.
8. Gheorghe A, Griffiths U, Murphy A, et al. The economic burden of cardiovascular disease and hypertension in low- and middle-income countries: a systematic review. BMC Public Health 2018;18:975.
9. Deane KD, Demoruelle MK, Kelmenson LB, et al. Genetic and environmental risk factors for rheumatoid arthritis. Best Pract Res Clin Rheumatol 2017;31:3–18.
10. Kang Y, Cai Y, Zhang X, et al. Altered gut microbiota in RA: implications for treatment. Z Rheumatol 2017;76:451–7.
11. Gabriel SE, Crowson CS. Risk factors for cardiovascular disease in rheumatoid arthritis. Curr Opin Rheumatol 2012;24:171–6.
12. Boyer JF, Gourraud PA, Cantagrel A, et al. Traditional cardiovascular risk factors in rheumatoid arthritis: a meta-analysis. Joint Bone Spine 2011;78:179–83.
13. Choy E, Sattar N. Interpreting lipid levels in the context of high-grade inflammatory states with a focus on rheumatoid arthritis: a challenge to conventional cardiovascular risk actions. Ann Rheum Dis 2009;68:460–9.
14. Steiner G, Urowitz MB. Lipid profiles in patients with rheumatoid arthritis: mechanisms and the impact of treatment. Semin Arthritis Rheum 2009;38:372–81.
15. Amezaga Urruela M, Suarez-Almazor ME. Lipid paradox in rheumatoid arthritis: changes with rheumatoid arthritis therapies. Curr Rheumatol Rep 2012;14:428–37.
16. Myasoedova E, Crowson CS, Kremers HM, et al. Lipid paradox in rheumatoid arthritis: the impact of serum lipid measures and systemic inflammation on the risk of cardiovascular disease. Ann Rheum Dis 2011;70:482–7.
17. Park YB, Lee SK, Lee WK, et al. Lipid profiles in untreated patients with rheumatoid arthritis. J Rheumatol 1999;26:1701–4.
18. Myasoedova E, Crowson CS, Kremers HM, et al. Total cholesterol and LDL levels decrease before rheumatoid arthritis. Ann Rheum Dis 2010;69:1310–4.
19. Kim JY, Lee EY, Park JK, et al. Patients with rheumatoid arthritis show altered lipoprotein profiles with dysfunctional high-density

https://doi.org/10.24171/j.phrp.2025.0036

lipoproteins that can exacerbate inflammatory and atherogenic process. PLoS One 2016;11:e0164564.
20. Govindan KP, Basha S, Ramesh V, et al. A comparative study on serum lipoprotein (a) and lipid profile between rheumatoid arthritis patients and normal subjects. J Pharm Bioallied Sci 2015;7(Suppl 1):S22–5.
21. Dessie G, Tadesse Y, Demelash B, et al. Assessment of serum lipid profiles and high-sensitivity C-reactive protein among patients suffering from rheumatoid arthritis at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia: a cross-sectional study. Open Access Rheumatol 2020;12:223–32.
22. Gan L, He Y, Liu L, et al. Association of serum lipids with autoantibodies and inflammatory markers in rheumatoid arthritis patients. Clin Chim Acta 2018;486:282–90.
23. Weber B, He Z, Yang N, et al. Divergence of cardiovascular biomarkers of lipids and subclinical myocardial injury among rheumatoid arthritis patients with increased inflammation. Arthritis Rheumatol 2021;73:970–9.
24. Masuko K. Glucose as a potential key to fuel inflammation in rheumatoid arthritis. Nutrients 2022;14:2349.
25. Vallve JC, Paredes S, Girona J, et al. Tumor necrosis factor-alpha-1031 T/C polymorphism is associated with smaller and more proatherogenic low density lipoprotein particles in patients with rheumatoid arthritis. J Rheumatol 2008;35:1697–703.
26. Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.
27. Wells G, Becker JC, Teng J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954–60.
28. Friedwald WT, Leve RI, Fredrickson DS. Estimation of concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499–502.
29. Ruscitti P, Ursini F, Cipriani P, et al. Prevalence of type 2 diabetes and impaired fasting glucose in patients affected by rheumatoid arthritis: results from a cross-sectional study. Medicine (Baltimore) 2017;96:e7896.
30. Ali MH, Hassan AJ, Hasan EJ. Risk factors and early detection of diabetes mellitus in early rheumatoid arthritis women. J Fac Med Baghdad 2018;60:74–6.
31. Ristic GG, Subota V, Stanisavljevic D, et al. Impact of disease activity on impaired glucose metabolism in patients with rheumatoid arthritis. Arthritis Res Ther 2021;23:95.
32. Albrecht K, Luque Ramos A, Hoffmann F, et al. High prevalence of diabetes in patients with rheumatoid arthritis: results from a questionnaire survey linked to claims data. Rheumatology (Oxford) 2018;57:329–36.
33. Zhang P, Liu J, Tan B, et al. Hypercoagulable state is associated with NF-kappa B activation and increased inflammatory factors in patients with rheumatoid arthritis. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi 2016;32:364–8. Chinese.
34. Li J, Chen Y, Liu Q, et al. Mechanistic and therapeutic links between rheumatoid arthritis and diabetes mellitus. Clin Exp Med 2023;23:287–99.
35. Masoumi M, Mehrabzadeh M, Mahmoudzehi S, et al. Role of glucose metabolism in aggressive phenotype of fibroblast-like synoviocytes: latest evidence and therapeutic approaches in rheumatoid arthritis. Int Immunopharmacol 2020;89(Pt A):107064.
36. Matuschik L, Riabov V, Schmuttermaier C, et al. Hyperglycemia induces inflammatory response of human macrophages to CD163-mediated scavenging of hemoglobin-haptoglobin complexes. Int J Mol Sci 2022;23:1385.
37. Taile J, Patche J, Veeren B, et al. Hyperglycemic condition causes pro-inflammatory and permeability alterations associated with monocyte recruitment and deregulated NFκB/PPARγ pathways on cerebral endothelial cells: evidence for polyphenols uptake and protective effect. Int J Mol Sci 2021;22:1385.
38. Wang X, Fan D, Cao X, et al. The role of reactive oxygen species in the rheumatoid arthritis-associated synovial microenvironment. Antioxidants (Basel) 2022;11:1153.
39. Erum U, Ahsan T, Khowaja D. Lipid abnormalities in patients with rheumatoid arthritis. Pak J Med Sci 2017;33:227–30.
40. Lakatos J, Harsagyi A. Serum total, HDL, LDL cholesterol, and triglyceride levels in patients with rheumatoid arthritis. Clin Biochem 1988;21:93–6.
41. Dursunoglu D, Evrengul H, Polat B, et al. Lp(a) lipoprotein and lipids in patients with rheumatoid arthritis: serum levels and relationship to inflammation. Rheumatol Int 2005;25:241–5.
42. Botta E, Merono T, Saucedo C, et al. Associations between disease activity, markers of HDL functionality and arterial stiffness in patients with rheumatoid arthritis. Atherosclerosis 2016;251:438–44.
43. Ormseth MJ, Stein CM. High-density lipoprotein function in rheumatoid arthritis. Curr Opin Lipidol 2016;27:67–75.
44. Barter PJ, Nicholls S, Rye KA, et al. Antiinflammatory properties of HDL. Circ Res 2004;95:764–72.
45. Brites F, Martin M, Guillas I, et al. Antioxidative activity of high-density lipoprotein (HDL): mechanistic insights into potential clinical benefit. BBA Clin 2017;8:66–77.
46. Fisher EA, Feig JE, Hewing B, et al. High-density lipoprotein function, dysfunction, and reverse cholesterol transport. Arterioscler Thromb Vasc Biol 2012;32:2813–20.
47. Giacaglia MB, Felix VP, Santana MF, et al. The composition of the HDL particle and its capacity to remove cellular cholesterol are associated with a reduced risk of developing active inflammatory rheumatoid arthritis. Int J Mol Sci 2024;25:10980.
48. Botta EE, Pierini F, Martin M, et al. Modifications on lipid profile

and high-density lipoprotein function related to treatment with tofacitinib in female patients with rheumatoid arthritis: impact of previous therapy with biological agents. J Clin Lipidol 2025 Feb 21 [Epub]. https://doi.org/10.1016/j.jacl.2025.02.013.
49. Yan J, Yang S, Han L, et al. Dyslipidemia in rheumatoid arthritis: the possible mechanisms. Front Immunol 2023;14:1254753.
50. Behl T, Kaur I, Sehgal A, et al. The lipid paradox as a metabolic checkpoint and its therapeutic significance in ameliorating the associated cardiovascular risks in rheumatoid arthritis patients. Int J Mol Sci 2020;21:9505.
51. Venetsanopoulou AI, Pelechas E, Voulgari PV, et al. The lipid paradox in rheumatoid arthritis: the dark horse of the augmented cardiovascular risk. Rheumatol Int 2020;40:1181–91.
52. Choi W, Kang JH, Park JY, et al. Elevated triglyceride levels are associated with increased risk for major adverse cardiovascular events in statin-naïve rheumatoid arthritis patients: a nationwide cohort study. Semin Arthritis Rheum 2023;63:152274.
53. Rodríguez-Carrio J, Alperi-Lopez M, Lopez P, et al. High triglycerides and low high-density lipoprotein cholesterol lipid profile in rheumatoid arthritis: a potential link among inflammation, oxidative status, and dysfunctional high-density lipoprotein. J Clin Lipidol 2017;11:1043–54.
54. Esmaillzadeh A, Azadbakht L. Increased levels of inflammation among women with enlarged waist and elevated triglyceride concentrations. Ann Nutr Metab 2010;57:77–84.
55. Nattagh-Eshtivani E, Gheflati A, Barghchi H, et al. The role of Pycnogenol in the control of inflammation and oxidative stress in chronic diseases: molecular aspects. Phytother Res 2022;36:2352–74.
56. Nattagh-Eshtivani E, Pahlavani N, Ranjbar G, et al. Does propolis have any effect on rheumatoid arthritis?: a review study. Food Sci Nutr 2022;10:1003–20.
57. Bahari H, Taheri S, Rashidmayvan M, et al. The effect of royal jelly on liver enzymes and glycemic indices: a systematic review and meta-analysis of randomized clinical trials. Complement Ther Med 2023;77:102974.
58. Putera HD, Doewes RI, Shalaby MN, et al. The effect of conjugated linoleic acids on inflammation, oxidative stress, body composition and physical performance: a comprehensive review of putative molecular mechanisms. Nutr Metab (Lond) 2023;20:35.
59. Esmaeili F, Maleki V, Kheirouri S, et al. The effects of taurine supplementation on metabolic profiles, pentosidine, soluble receptor of advanced glycation end products and methylglyoxal in adults with type 2 diabetes: a randomized, double-blind, placebo-controlled trial. Can J Diabetes 2021;45:39–46.
60. Hajizadeh-Sharafabad F, Alizadeh M, Mohammadzadeh MH, et al. Effect of Gundelia tournefortii L. extract on lipid profile and TAC in patients with coronary artery disease: a double-blind randomized placebo controlled clinical trial. J Herb Med 2016;6:59–66.
61. Saleh-Ghadimi S, Kheirouri S, Golmohammadi A, et al. Effect of flaxseed oil supplementation on anthropometric and metabolic indices in patients with coronary artery disease: a double-blinded randomized controlled trial. J Cardiovasc Thorac Res 2019;11:152–60.

https://doi.org/10.24171/j.phrp.2025.0036

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0009
Osong Public Health Res Perspect 2025;16(3):261-269

Original Article

The level of household food insecurity is associated with the risk of infectious diseases among toddlers in Indonesia: a cross-sectional study

Siti Masitoh[1], Tri Wurisastuti[1], Woro Riyadina[1], Sudarto Ronoatmodjo[2]

[1]Center for Research of Public Health and Nutrition, The National Research and Innovation Agency, Cibinong Science Center, West Java, Indonesia
[2]Department of Epidemiology, Faculty of Public Health, Universitas Indonesia, Kampus UI Depok, West Java, Indonesia

ABSTRACT

Objectives: This study investigated the correlation between food insecurity and infectious diseases among toddlers in Indonesia.
Methods: This research employed a cross-sectional study design using data from the 2021 Indonesian Nutrition Status Survey, which included a sample of 84,115 toddlers. Food insecurity levels were assessed using the food insecurity experience scale, developed by the Food and Agriculture Organization. The dependent variable, infectious disease, was defined as a toddler experiencing 1 or more of the following conditions: acute respiratory infections, diarrhea, pneumonia, measles, or worms. Multivariate analysis was conducted using multiple logistic regression to determine adjusted odds ratios (ORs).
Results: The findings indicate that 23.78% of toddlers experienced at least 1 infectious disease. While more than half of the respondents were food secure, 26.5% faced mild food insecurity, 13.6% moderate food insecurity, and 6.5% severe food insecurity. Toddlers from households experiencing mild, moderate, and severe food insecurity had higher risks of infectious diseases compared to those from food-secure households, with adjusted ORs of 1.367 (95% confidence interval [CI], 1.308–1.428), 1.490 (95% CI, 1.399–1.588), and 1.500 (95% CI, 1.374–1.638), respectively.
Conclusion: In conclusion, more severe food insecurity is correlated with an increased risk of toddlers suffering infectious diseases.

Keywords: Child; Communicable diseases; Food insecurity; Infant

Received: January 16, 2025
Revised: March 20, 2025
Accepted: April 24, 2025

Corresponding author:
Siti Masitoh
Center for Research of Public Health and Nutrition, The National Research and Innovation Agency, Meatrpo Building, Cibinong Science Center, Jl. Raya Jakarta-Bogor KM 46, Cibinong, Cibinong, West Java, Indonesia
E-mail: siti091@brin.go.id

Introduction

In 2023, an estimated 4.8 million children under the age of five died worldwide. This marks a 61 percent reduction in the global under-five mortality rate, which fell from 94 deaths per 1,000

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

live births in 1990 to 37 per 1,000 in 2023. Despite these gains, infectious diseases—particularly pneumonia, diarrhoea, and malaria—remain among the foremost causes of mortality in this age group, alongside complications related to preterm birth and intrapartum events [1,2]. Indonesia is among the top 10 countries with the highest under-5 mortality rates [2]. In Indonesia, the under-5 mortality rate is 32 deaths per 1,000 live births, representing a decrease from 40 per 1,000 live births in 2012 [3]. However, this rate remains high and far from achieving the target of 25 deaths per 1,000 live births set by the Sustainable Development Goals (SDGs) [2].

One of the primary reasons for Indonesia's high under-5 mortality rate is the prevalence of infectious diseases, which account for nearly 83% of deaths in this age group [4]. In 2021, pneumonia and diarrhea continued to be the leading causes of post-neonatal mortality, contributing to 14.4% and 14% of deaths, respectively. Congenital anomalies also accounted for 10.6% of deaths. Similarly, among children aged 12 to 59 months, diarrhea contributed to 10.3% of deaths, pneumonia to 9.4%, congenital diseases to 10%, and other causes such as neonatal complications, injuries, measles, and malaria in endemic regions [4,5].

Infectious diseases in toddlers are associated not only with under-5 mortality rates but also with stunting. Indonesia's stunting rate has steadily declined from 31% in 2018 to 24.4% in 2021, with the latest data from 2022 indicating a further reduction to 21.6% [4,6,7]. However, stunting remains a critical issue, as the current rate still exceeds the WHO standard of 20% [8]. Indonesia aims to reduce stunting prevalence to 14% by 2024 [9]. Given the significant impact of infectious diseases on stunting and under-5 mortality, reducing the prevalence of infectious diseases among toddlers is crucial for Indonesia.

According to United Nations International Children's Emergency Fund (UNICEF)'s 1998 framework, a bidirectional relationship exists between nutritional status and infectious diseases, wherein poor nutrition increases susceptibility to infections, and infections further compromise nutritional status. Additionally, poverty, food insecurity, and educational attainment serve as indirect factors influencing both infectious diseases and nutritional status [10]. Among the various factors contributing to infectious diseases, this study focuses primarily on food insecurity as the independent variable. A household is considered food insecure when it lacks sufficient safe and nutritious food for normal growth and development and to maintain an active and healthy life. This insecurity may result from the unavailability of food and/or insufficient resources to obtain it [11]. Data from the Central Bureau of Statistics indicate that the prevalence of moderate or severe food insecurity, measured using the food insecurity experience scale (FIES), has been steadily decreasing, from 5.42% in 2019 to 4.02% in 2024 [12].

HIGHLIGHTS

- One of the contributors to the high under-5 mortality rate in Indonesia is infectious diseases, accounting for approximately 83% of cases.
- Infectious diseases directly influence nutritional status, while poverty, food insecurity, and educational attainment act as indirect factors that can impact both infectious diseases and nutritional outcomes.
- The risk of infectious diseases among toddlers increases with higher levels of household food insecurity, even after adjusting for factors such as age, immunization status, region, and socioeconomic status.

Food insecurity leads to malnutrition, compromising toddlers' immunity and making them more susceptible to diseases. Malnutrition can increase the production of pro-inflammatory cytokines, which inhibit immune responses [13]. Additionally, malnutrition resulting from protein deficiency may cause dysfunction of T cells and B cells, key components in the adaptive immune response [14]. Consequently, malnourished toddlers are more susceptible to infectious diseases such as diarrhea and pneumonia, which are leading causes of childhood mortality in developing countries [15].

A study conducted in Brazil demonstrated a significant relationship between morbidity and food insecurity. Children from severely food-insecure families were more likely to experience coughing (adjusted odds ratio [OR], 1.79) and to be hospitalized due to diarrhea (adjusted OR, 2.55) [16]. Similarly, research in the United States found that children living in food-insecure households were 1.9 times more likely to have poor health status and 1.31 times more likely to be hospitalized since birth compared to children from food-secure households [17]. Similar results were observed among infants in Ghana [18].

Although existing evidence suggests an association between food insecurity and infectious diseases among toddlers in countries such as Brazil, the United States, and Ghana, these studies focused on specific diseases—such as cough, fever, diarrhea, and pneumonia in Brazil; general poor health status in the United States; and respiratory infections restricted to rural areas in Ghana. Furthermore, research on this relationship remains limited in low- and middle-income countries, particularly in Indonesia. Given that the impact of food insecurity may differ between developed and developing countries, context-specific research is essential.

https://doi.org/10.24171/j.phrp.2025.0009

Therefore, this study aims to comprehensively investigate the association between food insecurity and the risk of infectious diseases—including acute respiratory infections (ARIs), diarrhea, pneumonia, measles, and helminthiasis—among toddlers in Indonesia.

Materials and Methods

This study employs a cross-sectional design utilizing secondary data obtained from the 2021 Indonesian Nutritional Status Survey (Survei Status Gizi Indonesia, SSGI), covering samples from 34 provinces. The sample included children aged 0 to 59 months who participated in the SSGI and met specific inclusion criteria. Inclusion criteria were children aged 0–59 months who had complete data on infectious diseases (ARIs, diarrhea, pneumonia, pulmonary tuberculosis, measles, and helminthiasis) and comprehensive data on all relevant research variables. A total of 84,115 children met these criteria and were included in the analysis.

Variables

The research applies a causal model, with food insecurity as the primary independent variable measured by the FIES, and infectious disease status as the dependent/outcome variable. The analysis also includes covariates such as child characteristics (age, gender, low birth weight status, immunization status), parental characteristics (mother's education, father's education, mother's employment status), and household characteristics (area of residence, wealth quintile, household size, number of children under 5, water source, and sanitation).

The dependent/outcome variable, infectious disease status, was defined based on whether a child had experienced 1 or more of 6 diseases (ARIs, diarrhea, pneumonia, pulmonary tuberculosis, measles, and helminthiasis) at least once, based on diagnostic and symptom history. Children were categorized into 2 groups: those without infectious diseases and those with 1 or more infectious diseases.

Food insecurity, the primary independent variable, is defined as limited or uncertain availability of nutritionally adequate and safe food or limited or uncertain ability to acquire acceptable food in socially acceptable ways [19]. The FIES measures the percentage of individuals within a national population who have experienced moderate or severe food insecurity at least once within the past 12 months. The FIES instrument has been validated in Indonesia through Rasch model analysis, showing strong reliability (Rasch reliability, 0.77), good item fit (infit statistics, 0.7–1.3), and no significant residual correlations (all below the 0.4 threshold). These results confirm that FIES is a valid and reliable instrument for assessing food insecurity within the Indonesian context [20]. The FIES instrument comprises 8 sequentially ordered questions describing progressively increasing severity of food insecurity based on respondents' experiences: (1) worrying about not having enough food due to lack of money or other resources, (2) being unable to eat healthy and nutritious food due to lack of money or other resources, (3) eating only a little due to lack of money or other resources, (4) skipping a meal on a specific day due to lack of money or other resources, (5) eating less than needed due to lack of money or other resources, (6) running out of food due to lack of money or other resources, (7) feeling hungry but not eating due to lack of money or other resources, and (8) not eating for an entire day due to lack of money or other resources. Respondents answered "yes" (coded as 1) or "no" (coded as 0) to each of the 8 items, yielding a total score ranging from 0 to 8. Food insecurity levels were categorized as follows: (1) food secure (score 0); (2) mild food insecurity if respondents answered "yes" to items 1, 2, or 3; (3) moderate food insecurity if respondents answered "yes" to items 4, 5, or 6; and (4) severe food insecurity if respondents answered "yes" to items 7 or 8. This categorization is based on the assumption that the FIES questions progressively represent increasing severity of food insecurity, with questions 1–3 indicating mild, 4–6 moderate, and 7–8 severe food insecurity [21].

Data Analysis

Data analysis was conducted in sequential stages: univariate, bivariate, and multivariate. Univariate analysis described the characteristics of each variable. Bivariate analysis examined relationships between independent and dependent variables, calculating crude ORs. Subsequently, multivariate analysis using multiple logistic regression was performed to obtain adjusted ORs with 95% confidence intervals (CIs). Missing data were handled by exclusion, where participants with missing data on key variables were omitted from analysis. A sensitivity analysis was also conducted to assess potential impacts of excluding missing data, ensuring this approach did not introduce bias or compromise the validity of the study results.

Ethical Consideration

This study received ethical approval from the Ethics Committee of Faculty of Public Health, Universitas Indonesia (No: Ket-165/UN2.F10.D11/PPM.00.02/2023).

Results

The characteristics of the study respondents are presented

in Table 1. Among the respondents, 23.78% of children under 5 experienced at least 1 infectious disease. More than half of the respondents were food secure; however, 26.5% of households experienced mild food insecurity, 13.6% experienced moderate food insecurity, and 6.5% experienced severe food insecurity. Age distribution among respondents was relatively balanced: 0–11 months (18.8%), 12–23 months (21.7%), 24–35 months (21.1%), 36–47 months (20.3%), and 48–59 months (18.1%). Slightly more respondents were male (51.2%) than female. Most children had a normal birth weight (93.3%), normal nutritional status (76.6%), and the majority had received complete immunizations appropriate for their age (60.8%).

Parental education levels were evenly distributed, with slightly higher proportions of mothers and fathers having attained higher education. A higher proportion of mothers were unemployed (65.7%). More respondents resided in rural areas (52.2%). The distribution across wealth quintile index was nearly even, with the highest proportion in the fifth quintile (wealthiest group, 21.6%) and the lowest in the first quintile (poorest group, 17.3%). Over half of the respondents belonged to households consisting of 4 members or fewer (50.8%). Most households had only 1 child under 5 (78.1%), access to safe drinking water (87.3%), and proper sanitation (85.7%).

Table 2 presents the bivariate analysis examining associations between risk factor variables and infectious diseases. Children from households experiencing mild food insecurity had a 1.496 times greater likelihood (95% CI, 1.441–1.554) of contracting infectious diseases compared to children from food-secure households. The likelihood increased to 1.708 times (95% CI, 1.628–1.791) among children from households experiencing moderate food insecurity, and further increased to 1.777 times (95% CI, 1.666–1.896) among those experiencing severe food insecurity, compared to children from food-secure households. The relationship between the level of food insecurity and infectious disease status was statistically significant ($p = 0.001$).

These findings indicate that the risk of contracting infectious diseases among children under 5 increases progressively with higher levels of household food insecurity. Additionally, Table 2 demonstrates statistically significant relationships between all risk factor variables and infectious diseases in children, with the exception of household size.

Based on Table 3, the final multivariate analysis confirms that the risk of infectious diseases among children under 5 increases with household food insecurity severity. Specifically, children from households with mild food insecurity were 1.367 times more likely (95% CI, 1.308–1.428) to suffer infectious diseases. This risk further increased to 1.490 times (95% CI,

Table 1. Characteristics of respondents ($n = 84,115$)

Characteristic	Value
Infectious disease status	
No infectious disease	58,893 (70.0)
Has infectious disease	20,006 (23.8)
Level of food insecurity	
Food-secure	44,917 (53.4)
Mild insecurity	22,331 (26.5)
Moderate insecurity	11,404 (13.6)
Severe insecurity	5,463 (6.5)
Age (mo)	
0–11	15,786 (18.8)
12–23	18,283 (21.7)
24–35	17,764 (21.1)
36–47	17,037 (20.3)
48–59	15,245 (18.1)
Sex	
Female	41,008 (48.8)
Male	43,107 (51.2)
Low birth weight status	
Normal weight	78,498 (93.3)
Low birth weight	5,617 (6.7)
Nutritional status	
Not stunted	64,428 (76.6)
Stunted	19,687 (23.4)
Immunization status	
Complete	51,149 (60.8)
Incomplete	32,966 (39.2)
Mother's education level	
High	43,362 (51.6)
Low	40,753 (48.4)
Father's education level	
High	42,853 (50.9)
Low	41,262 (49.1)
Mother's employment status	
Not working	55,299 (65.7)
Working	28,816 (34.3)
Area of residence	
Urban	40,187 (47.8)
Rural	43,928 (52.2)
Wealth quintile index	
Quintile 1	14,554 (17.3)
Quintile 2	17,866 (21.2)
Quintile 3	16,173 (19.2)
Quintile 4	17,327 (20.6)
Quintile 5	18,195 (21.6)
No. of household members	
≤4	42,730 (50.8)
>4	41,385 (49.2)
No. of toddlers	
1	65,659 (78.1)
≥2	18,456 (21.9)
Access to drinking water source	
Adequate	73,399 (87.3)
Inadequate	10,716 (12.7)
Access to sanitation	
Adequate	72,058 (85.7)
Inadequate	12,057 (14.3)

Data are presented as *n* (%).

https://doi.org/10.24171/j.phrp.2025.0009

Table 2. Relationship between risk factors and infectious diseases among toddlers

Variable	Infectious diseases	
	Crude OR (95% CI)	*p*
Level of food insecurity		
Food-secure (ref.)	1	
Mild insecurity	1.496 (1.441–1.554)	0.001
Moderate insecurity	1.708 (1.628–1.791)	0.001
Severe insecurity	1.777 (1.666–1.896)	0.001
Age (mo)		
48–59 (ref.)	1	
0–11	0.934 (0.884–0.987)	0.015
12–23	1.451 (1.379–1.527)	0.001
24–35	1.269 (1.205–1.336)	0.001
36–47	1.099 (1.043–1.160)	0.001
Sex		
Female (ref.)	1	
Male	1.052 (1.019–1.086)	0.002
LBW status		
Normal weight (ref.)	1	
LBW	1.104 (1.036–1.177)	0.002
Immunization status		
Complete	1	
Incomplete	0.992 (0.960–1.025)	0.628
Nutritional status		
Not stunted (ref.)	1	
Stunted	1.221 (1.176–1.267)	0.001
Mother's education level		
High (ref.)	1	
Low	1.265 (1.225–1.307)	0.001
Father's education level		
High (ref.)	1	
Low	1.237 (1.198–1.277)	0.001
Mother's employment status		
Not working (ref.)	1	
Working	0.890 (0.861–0.921)	0.001
Area of residence		
Urban (ref.)	1	
Rural	1.269 (1.219–1.311)	0.001
Wealth quintile index		
Quintile 5 (ref.)	1	
Quintile 1	1.875 (1.779–1.976)	0.001
Quintile 2	1.600 (1.522–1.682)	0.001
Quintile 3	1.410 (1.338–1.485)	0.001
Quintile 4	1.226 (1.164–1.291)	0.001
No. of household members		
≤4 (ref.)	1	
>4	0.992 (0.960–1.024)	0.610
No. of toddlers		
1 (ref.)	1	
≥2	0.854 (0.821–0.889)	0.001
Access to drinking water source		
Adequate (ref.)	1	
Inadequate	1.193 (1.138–1.250)	0.001
Access to sanitation		
Adequate (ref.)	1	
Inadequate	1.348 (1.290–1.409)	0.001

OR, odds ratio; CI, confidence interval; ref., reference; LBW, low birth weight.

1.399–1.588) for children from households with moderate food insecurity, and to 1.500 times (95% CI, 1.374–1.638) for children from households experiencing severe food insecurity, compared to children from food-secure households. Subgroup analysis also revealed that children living in rural areas faced a higher risk than those in urban areas. The association between food insecurity level and infectious disease status remained statistically significant ($p = 0.001$).

Discussion

Food insecurity emerged as a significant risk factor in this study. The multivariate analysis demonstrated a dose-response relationship between the level of household food insecurity and the risk of infectious diseases among toddlers. Compared to children from food-secure households, those from households experiencing mild food insecurity had a 36.7% higher likelihood (OR, 1.367; 95% CI, 1.308–1.428) of contracting infectious diseases. This risk increased to 49.0% (OR, 1.490; 95% CI, 1.399–1.588) among children from moderately food-insecure households, and further rose to 50.0% (OR, 1.500; 95% CI, 1.374–1.638) for those from severely food-insecure households. These findings are consistent with studies from Brazil and the United States, which reported significantly higher risks of infectious diseases among children from severely food-insecure households [16,17]. Similar results have been observed in other low- and middle-income countries, such as India, Pakistan, Bangladesh, Nepal, and Sri Lanka, where food insecurity—compounded by poverty and poor sanitation—contributes substantially to high rates of infectious diseases among children [22].

While the findings of this study align with previous international research, Indonesia's unique geographic and socioeconomic conditions may influence the strength and extent of the observed association. As an archipelagic nation with substantial disparities in healthcare access, sanitation, and nutritional conditions, the health impact of food insecurity on children may be especially pronounced in certain regions [23]. The subgroup analysis revealed that the risk of infectious diseases was notably higher among children residing in rural areas compared to their urban counterparts.

To gain deeper insight into this relationship, it is crucial to consider the interconnected pathways through which food insecurity affects infectious diseases. These pathways include socioeconomic constraints, malnutrition, impaired immune function, limited healthcare access, psychological distress, and environmental enteric dysfunction (EED). Socioeconomic status plays a pivotal role since food insecurity often reflects broader poverty conditions. Studies

Table 3. Final model relationship between risk factors variables and infectious diseases among toddlers

Variable	Adjusted odds ratio (95% confidence interval)		
	Overall	Urban	Rural
Level of food insecurity[a)]			
Food-secure (ref.)	1	1	1
Mild insecurity	1.367 (1.308–1.428)	1.120 (1.13–1.274)	1.557 (1.477–1.639)
Moderate insecurity	1.490 (1.399–1.588)	1.400 (1.299–1508)	1.664 (1.555–1.780)
Severe insecurity	1.500 (1.374–1.638)	1.502 (1.361–1.658)	1.614 (1.471–1.770)

Ref., reference.
[a)]Controlled for age, immunization status, area of residence, wealth quintile index, and the interaction between food insecurity and area of residence.

conducted in Ethiopia revealed that children from the lowest wealth quintiles were significantly more likely to experience food insecurity [24]. This study confirmed a similar trend, with households in the poorest quintile exhibiting the highest risks of infectious diseases in a dose-response relationship. Research from Bangladesh supports this observation, demonstrating that children from poorer households were 1.3 times more likely to suffer from respiratory infections due to limited access to adequate food and healthcare [25,26].

Malnutrition is another critical pathway linking food insecurity to increased infection susceptibility. Deficiencies in essential nutrients such as protein, vitamin A, iron, and zinc can weaken immune responses, decrease immune cell production, diminish antibody activity, and impair mucosal defenses, making individuals more vulnerable to infections and delaying recovery [27,28]. A study conducted in Indonesia reported that toddlers from food-insecure families had a 1.24 times higher risk of stunting, with the risk rising to 1.39 among those from severely food-insecure households [29]. Similarly, Malaysian studies have indicated that food insecurity is associated with a 2.15 times greater risk of underweight status [30]. Food insecurity can be linked to malnutrition, referring to the loss and reduction in both quantity and quality of food due to economic constraint [31].

Limited healthcare access exacerbates the adverse health outcomes linked to food insecurity. Households facing financial constraints may prioritize basic food needs over pediatric healthcare, leading to reduced healthcare utilization. A study conducted in the United States found that individuals from food-insecure households were significantly less likely to have a regular healthcare provider or receive home health visits compared to their food-secure counterparts [32,33]. Due to data limitations and potential biases arising from the coronavirus disease 2019 pandemic, healthcare utilization variables were not included in this study.

Psychological distress, particularly among mothers, further mediates the relationship between food insecurity and child health outcomes. Mothers experiencing food insecurity often report high stress and anxiety levels, impairing their caregiving abilities [34]. According to UNICEF's conceptual framework, inadequate maternal and child care practices represent indirect factors contributing to malnutrition and increased infection risks [10]. An analysis of data from the Indonesia Family Life Survey highlighted a significant association between food insecurity and depressive symptoms among mothers [35]. Similar findings have been reported in African countries and India, where women from food-insecure households exhibit heightened levels of anxiety and depression, potentially affecting their caregiving practices [36].

Another critical factor linking food insecurity with infectious diseases is EED, a chronic intestinal disorder resulting from repeated pathogen exposure and inadequate sanitation. EED causes intestinal inflammation, increased gut permeability, and impaired nutrient absorption, further worsening malnutrition and increasing susceptibility to infections, particularly diarrhea and respiratory illnesses [37]. This harmful cycle illustrates that food insecurity not only leads to undernutrition but also exacerbates vulnerability to infectious diseases through impaired immune function, gut dysfunction, inadequate healthcare access, and psychological distress. These findings underscore the urgency for comprehensive interventions addressing both nutritional and socioeconomic determinants of child health.

The results of this study have significant implications for achieving the SDGs, particularly SDG 2 (zero hunger) and SDG 3 (good health and well-being). Tackling food insecurity through targeted assistance programs and integrating nutritional support with immunization initiatives aligns closely with global efforts aimed at reducing malnutrition and improving child health outcomes. Strengthening such interventions can significantly contribute toward meeting SDG targets by enhancing access to nutritious food, reducing childhood morbidity, and supporting overall health equity.

Addressing food insecurity necessitates comprehensive

https://doi.org/10.24171/j.phrp.2025.0009

strategies targeting both direct and underlying factors. Government-led initiatives, such as food assistance programs, nutritional supplementation, and economic support, are critical for mitigating the adverse effects of food insecurity. In Indonesia, the Bantuan Pangan Non-Tunai program has been implemented to improve household access to nutritious food, essential for preventing stunting and reducing infectious diseases among children [38], the Program Keluarga Harapan provides financial assistance to low-income families, enabling them to prioritize healthcare and better nutrition for their children [39].

Moreover, ensuring universal immunization coverage represents another essential intervention for reducing infection risks among children from food-insecure households. Research indicates that children with incomplete immunization schedules are at higher risk of severe infections and hospitalizations, particularly in resource-limited settings [40]. Integrating immunization efforts with food security programs could enhance outreach and significantly improve child health outcomes.

Strengths and Limitation

This study has several notable strengths. First, it utilized a relatively large sample derived from a robust sample-size calculation, ensuring high statistical power. The sample was selected using multistage random sampling, enabling generalization of the findings to the national level. Additionally, the research employed a validated instrument to measure food insecurity—the FIES—which has been utilized by the Central Bureau of Statistics in the National Socioeconomic Survey since 2017. Interviews were conducted by trained enumerators, which helps to minimize interviewer-related bias.

Despite these strengths, the study has several limitations, primarily related to its reliance on secondary data from the 2021 Indonesian Nutritional Status Survey, restricting the range of available variables. One potential limitation arises from the reliance on self-reported food insecurity data, which may introduce recall bias or social desirability bias. Respondents might underreport or overreport their food insecurity experiences due to memory limitations or the desire to provide socially acceptable answers. Such biases could lead to misclassification, potentially resulting in an underestimation or overestimation of the true association between food insecurity and infectious diseases in children.

Although multivariate analysis was employed to control for confounding variables, several relevant factors (e.g., dietary intake and healthcare access) could not be included due to data limitations. Specifically, dietary intake data were available only for children aged 0 to 23 months, and healthcare access data were limited to sick children. Additionally, the cross-sectional design of the study introduces temporal ambiguity, preventing determination of whether exposure (food insecurity) preceded the outcome (infectious diseases). Therefore, the study cannot establish causality between food insecurity and infectious diseases. Future research employing more robust designs, such as longitudinal or cohort studies, is necessary to address these limitations.

Conclusion

Based on the findings of this study, it can be concluded that the risk of toddlers suffering from infectious diseases increases progressively with the severity of household food insecurity, even after adjusting for confounders such as age, immunization status, area of residence, and wealth quintile index. Compared to toddlers from food-secure households, those from households experiencing mild food insecurity had a 1.367 times (95% CI, 1.308–1.428) higher risk of infectious diseases. This risk increased to 1.490 times (95% CI, 1.399–1.588) for toddlers in moderately food-insecure households, and further increased to 1.500 times (95% CI, 1.374–1.638) for those experiencing severe food insecurity.

Notes

Ethics Approval

This study was approved by the Ethics Committee of FKM UI (No: Ket-165/UN2.F10.D11/PPM.00.02/2023).

Conflicts of Interest

The authors have no conflicts of interest to declare.

Funding

None.

Availability of Data

The datasets are not publicly available but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: SM, SR; Data curation: SM, TW; Formal analysis: all authors; Investigation: SR; Methodology: SM, SR, WR; Project administration: SM; Supervision: WR, SR; Validation: SM, WR, SR; Visualization: SM, TW; Writing–original draft: all authors; Writing–review & editing: all authors. All authors read and approved the final manuscript.

References

1. UNICEF (US). Under-fiver mortality [Internet]. UNICEF; 2025 [cited 2025 May 17]. Available from: https://data.unicef.org/topic/child-survival/under-five-mortality/#:~:text=The global under-five mortality,a matter of urgent concern.
2. World Health Organization (WHO) (CH). Child mortality (under 5

years) [Internet]. WHO; 2022 [cited 2023 Feb 8]. Available from: https://www.who.int/news-room/fact-sheets/detail/child-mortality-under-5-years.
3. Statistics Indonesia (ID). Results of the 2020 population census long form: Badan Pusat Statistik: No. 09/01/Th. XXVI, 30 Januari 2023 [Internet]. Badan Pusat Statistik; 2023 [cited 2023 Jun 23]. Available from: https://www.bps.go.id/id/publication/2023/01/27/ffb5939b4393e5b1146a9b91/hasil-long-form-sensus-penduduk-2020.html. Indonesian.
4. Fikawati S. Child and adolescent nutrition. Rajawali Press; 2017. Indonesian.
5. Ministry of Health Republic of Indonesia (ID). Indonesia health profile 2021 [Internet]. Ministry of Health Republic of Indonesia; 2021 [cited 2023 Jul 23]. Available from: https://kemkes.go.id/id/profil-kesehatan-indonesia-2021. Indonesian.
6. UNICEF Indonesia (ID). Health provides the best chance for children to survive [Internet]. UNICEF Indonesia; 2023 [cited 2023 Feb 8]. Available from: https://www.unicef.org/indonesia/id/kesehatan. Indonesian.
7. Health Policy Development Agency MoH Indonesia (ID). Pocket book of the Indonesian Nutrition Status Study (SSGI) 2022 [Internet]. Health Policy Development Agency MoH Indonesia; 2022 [cited 2023 Jun 15]. Available from: https://www.badankebijakan.kemkes.go.id/buku-saku-hasil-studi-status-gizi-indonesia-ssgi-tahun-2021/. Indonesian.
8. World Health Organization (WHO) (CH). Nutrition Landscape Information System (NLIS): Stunting, wasting, overweight and underweight [Internet]. WHO; 2023 [cited 2023 Jun 18]. Available from: https://apps.who.int/nutrition/landscape/help.aspx?menu=0&helpid=391&lang=EN.
9. Government of Indonesia (ID). Presidential regulation of the Republic of Indonesia number 72 of 2021 on the acceleration of stunting reduction [Internet]. Government of Indonesia; 2021 [cited 2022 Feb 10]. Available from: https://peraturan.bpk.go.id/Details/174964/perpres-no-72-tahun-2021. Indonesian.
10. UNICEF (US). The state of the world's children 1998 [Internet]. Oxford University Press; 1998 [cited 2022 Jun 8]. Available from: https://www.unicef.org/media/84766/file/SOWC-1998.pdf.
11. Food and Agriculture Organization of the United Nations (IT). Hunger and food insecurity [Internet]. Food and Agriculture Organization of the United Nations; 2023 [cited 2023 Feb 9]. Available from: https://www.fao.org/hunger/en/.
12. Statistics Indonesia (ID). Prevalence of population with moderate or severe food insecurity, based on the food insecurity experience scale (percent) [Internet]. Badan Pusat Statistik; 2022 [cited 2025 May 28]. Available from: https://www.bps.go.id/id/statistics-table/2/MTQ3NCMy/prevalensi-penduduk-dengan-kerawanan-pangan-sedang-atau-berat-berdasarkan-pada-skala-pengalaman-kerawanan-pangan.html. Indonesian.
13. Amorim Sacramento L, Gonzalez-Lombana C, Scott P. Malnutrition disrupts adaptive immunity during visceral leishmaniasis by enhancing IL-10 production. PLoS Pathog 2024;20:e1012716.
14. Bourke CD, Berkley JA, Prendergast AJ. Immune dysfunction as a cause and consequence of malnutrition. Trends Immunol 2016;37:386–98.
15. Fassikaw Kebede B, Nemera Eticha B, Tsehay K. Treatment cure rate and associated factors of sever acute malnutrition 6-59 month children treated in therapeutic center in Pawe General Hospital: Nourth Wet Ethiopia. J Nutr Med Diet Care 2021;7:51.
16. Gubert MB, Spaniol AM, Bortolini GA, et al. Household food insecurity, nutritional status and morbidity in Brazilian children. Public Health Nutr 2016;19:2240–5.
17. Cook JT, Frank DA, Levenson SM, et al. Child food insecurity increases risks posed by household food insecurity to young children's health. J Nutr 2006;136:1073–6.
18. Ohemeng A, Marquis GS, Lartey A. Household food insecurity is associated with respiratory infections among 6-11-month old infants in rural Ghana. Pediatr Infect Dis J 2015;34:821–5.
19. USDA (US). Food security in the U.S.: measurement [Internet]. USDA; 2022 [cited 2022 Nov 30]. Available from: https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/measurement/.
20. Kadir K, Prasetyo OR, Rudiana E. Food insecurity experience scale measurement of agricultural households in indonesia: analysis of the agricultural integrated survey results [Internet]. Munich Personal RePEc Archive; 2024 [cited 2022 Jun 26]. Available from: https://mpra.ub.uni-muenchen.de/119416/.
21. Wambogo EA, Ghattas H, Leonard KL, et al. Validity of the food insecurity experience scale for use in sub-Saharan Africa and characteristics of food-insecure individuals. Curr Dev Nutr 2018;2:nzy062.
22. Cook JT, Frank DA, Berkowitz C, et al. Food insecurity is associated with adverse health outcomes among human infants and toddlers. J Nutr 2004;134:1432–8.
23. Bishwajit G, Ide S, Ghosh S. Social determinants of infectious diseases in South Asia. Int Sch Res Notices 2014;2014:135243.
24. Ayuningtyas D, Hapsari D, Rachmalina R, et al. Geographic and socioeconomic disparity in child undernutrition across 514 districts in Indonesia. Nutrients 2022;14:843.
25. Tebeje NB, Biks GA, Abebe SM, et al. Magnitude of child food insecurity, its association with child immunization and huosehold wealth status, and coping strategies in Dabat demographic and surveillance system north west Ethiopia. Int J Pediatr 2020;2020:3746354.
26. Azad KM. Risk factors for acute respiratory infections (ARI) among under-five children in Bangladesh. J Sci Res 2009;1:72–81.
27. Siziya S, Muula AS, Rudatsikira E. Correlates of diarrhoea among children below the age of 5 years in Sudan. Afr Health Sci 2013;13:376–83.
28. Anderson LC, Tegegn A, Tessema F, et al. Food insecurity, childhood illness and maternal emotional distress in Ethiopia. Public Health

https://doi.org/10.24171/j.phrp.2025.0009

Nutr 2012;15:648–55.
29. Masitoh S, Ronoatmodjo S, Nurokhmah S. The correlation between food insecurity level and stunting in Indonesia. J Public Health 2022; 13:385–98.
30. Ali Naser I, Jalil R, Wan Muda WM, et al. Association between household food insecurity and nutritional outcomes among children in Northeastern of Peninsular Malaysia. Nutr Res Pract 2014;8:304–11.
31. Reis M. Food insecurity and the relationship between household income and children's health and nutrition in Brazil. Health Econ 2012;21:405–27.
32. Kushel MB, Gupta R, Gee L, et al. Housing instability and food insecurity as barriers to health care among low-income Americans. J Gen Intern Med 2006;21:71–7.
33. Bhargava V, Lee JS. Food insecurity and health care utilization among older adults. J Appl Gerontol 2017;36:1415–32.
34. Hadley C, Patil CL. Food insecurity in rural Tanzania is associated with maternal anxiety and depression. Am J Hum Biol 2006;18:359–68.
35. Isaura ER, Chen YC, Adi AC, et al. Association between depressive symptoms and food insecurity among Indonesian adults: results from the 2007-2014 Indonesia family life survey. Nutrients 2019;11: 3026.
36. Hadley C, Tegegn A, Tessema F, et al. Food insecurity, stressful life events and symptoms of anxiety and depression in east Africa: evidence from the Gilgel Gibe growth and development study. J Epidemiol Community Health 2008;62:980–6.
37. Jones KD, Thitiri J, Ngari M, et al. Childhood malnutrition: toward an understanding of infections, inflammation, and antimicrobials. Food Nutr Bull 2014;35(2 Suppl):S64–70.
38. Agri TA, Ramadanti T, Adriani WA, et al. Advancing balanced growth to address SDG 2 challenges in the effort to reduce stunting in Indonesia [Internet]. Universitas Pembangunan Nasional "Veteran" Jakarta; 2024 [cited 2023 Jun 15]. Available from: https://conference.upnvj.ac.id/index.php/ncols/article/view/2972. Indonesian.
39. Ministry of Social Affairs of the Republic of Indonesia (ID). Regulation of the Minister of Social Affairs of the Republic of Indonesia Number 1 of 2018 on the Family Hope Program [Internet]. Ministry of Social Affairs of the Republic of Indonesia; 2018 [cited 2022 Jul 26]. Available from: https://peraturan.bpk.go.id/Details/129999/permensos-no-1-tahun-2018. Indonesian.
40. Atnafu Gebeyehu N, Abebe Gelaw K, Asmare Adella G, et al. Incomplete immunization and its determinants among children in Africa: systematic review and meta-analysis. Hum Vaccin Immunother 2023; 19:2202125.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0080
Osong Public Health Res Perspect 2025;16(3):270-284

Original Article

Analysis of influenza-like illness trends in Saudi Arabia: a comparative study of statistical and deep learning techniques

Fathelrhman EL Guma

Department of Mathematics, Faculty of Science, Al-Baha University, Al-Baha, Saudi Arabia

ABSTRACT

Objectives: To develop and evaluate forecasting models using the Holt-Winters statistical approach and the long short-term memory (LSTM) deep learning method for weekly seasonal influenza-like illness (ILI) incidences in Saudi Arabia. The study compares model performance and assesses the predictive value added by incorporating region-specific exogenous variables within Middle Eastern epidemiological modeling.

Methods: This study compared the performance of Holt-Winters and LSTM models in forecasting weekly ILI cases in Saudi Arabia, using data collected from 2017 to 2022. Time series analysis integrated exogenous variables including climatic conditions and population mobility trends. The Holt-Winters model employed both additive and multiplicative seasonal components. Model performance was evaluated using root mean squared error (RMSE), mean absolute percentage error, and R^2.

Results: The best-performing model, LSTM with exogenous variables, achieved an RMSE of 28.55, mean absolute error (MAE) of 0.14, R^2 of 0.96, and percent bias (PBIAS) of +2.1%, indicating negligible systematic error. The LSTM model without exogenous variables demonstrated slightly lower accuracy (RMSE of 34.07, MAE of 0.18, R^2 of 0.93, PBIAS of +5.8%), indicating strong predictive capability but less precision in determining peak ILI cases. The Holt-Winters model effectively captured seasonal and long-term trends, but showed a moderate performance with an RMSE of 82.57, MAE of 0.38, R^2 of 0.58, and a high PBIAS of +14.2%, revealing significant unexplained variability during periods of high incidence fluctuation.

Conclusion: This study highlights the respective strengths and limitations of statistical and machine learning approaches for ILI forecasting.

Keywords: Forecasting; Influenza, human; Neural networks, computer; Public health

Received: March 11, 2025
Revised: April 21, 2025
Accepted: April 25, 2025

Corresponding author:
Fathelrhman EL Guma
Department of Mathematics, Faculty of Science, Al-Baha University, Al-Aqiq 65931, Saudi Arabia
E-mail: fguma@bu.edu.sa

Introduction

The rapid spread of influenza-like illnesses (ILIs) poses a significant public health challenge

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

worldwide. According to the World Health Organization (WHO), approximately 1 billion cases of ILIs occur annually, including 3 to 5 million severe cases, resulting in an estimated 290,000 to 650,000 respiratory deaths globally [1,2]. In developing countries, ILIs account for over 99% of deaths among children under 5 years of age. These illnesses typically spread as seasonal epidemics during winter; however, in tropical regions, they circulate year-round. Common symptoms include cough, headache, severe malaise, and runny nose, and transmission occurs primarily through respiratory droplets [3]. Each year, over 2 million individuals gather in confined spaces for the Hajj pilgrimage, significantly increasing infection transmission during this period [4]. Seasonal variability and mass gatherings, especially during the Hajj season, complicate the monitoring and forecasting of ILIs in Saudi Arabia [5–8].

HIGHLIGHTS

- This study compared the Holt-Winters statistical model and the long short-term memory (LSTM) deep learning model for forecasting influenza-like illnesses in Saudi Arabia.
- LSTM outperformed Holt-Winters across all evaluation metrics, achieving a root mean squared error (RMSE) of 34.07 and an R^2 of 0.93, compared to Holt-Winters' RMSE of 82.57 and R^2 of 0.58.
- The LSTM model better captured nonlinear temporal patterns, whereas Holt-Winters struggled with rapid fluctuations in incidence and seasonal variability.
- Incorporating exogenous variables such as climate factors, mobility patterns, and vaccination rates further improved the predictive accuracy of the LSTM model.
- Future research should explore hybrid models integrating statistical and machine learning approaches for enhanced epidemic forecasting.

Traditional Forecasting

Forecasting is essential for combating the spread of ILIs as it facilitates effective resource allocation and the timely implementation of appropriate public health interventions. Accurate forecasting reduces uncertainty by predicting flu activity in advance, including when and where outbreaks are likely to occur. This predictive capability supports efficient resource allocation and improved planning before flu-related hospitalizations surge [9]. Previously, influenza surveillance data collected from clinics, diagnostic laboratories, general practitioners, and public health organizations were used to track ILIs and other respiratory infections. However, this approach often led to delays of 1 week or more in report issuance, with frequent retrospective modifications [10]. Additionally, unreported and unnoticed cases posed significant challenges in accurately predicting influenza incidence. The 2014 WHO Global Epidemiological Surveillance Standards for Influenza propose a "sentinel surveillance" approach, involving the regular collection of epidemiological and virological data from selected monitoring sites [11,12]. Although this approach is typically efficient in gathering high-quality data, it has several inherent limitations. First, it only captures influenza cases in individuals who seek medical attention, representing merely a fraction of the actual disease burden. Additionally, there is typically a delay between symptom onset and the decision to seek healthcare, compounded by a further delay of 1 to 2 weeks between data collection and its reporting. Moreover, the effectiveness of sentinel surveillance heavily depends on factors such as the availability of quality laboratory resources and skilled personnel, which may not always be accessible. Lastly, the selection of sentinel sites often prioritizes locations serving large, easily accessible populations, introducing potential sampling bias [13]. Mathematical models for predicting infectious disease outbreaks generally fall into 2 primary categories. The first category includes mechanistic models, which predict the future course of an ongoing epidemic. These models rely on understanding disease transmission mechanisms and their influencing factors, such as the depletion of susceptible populations or climate-related impacts on transmission rates. Mechanistic models may function at the population level (e.g., compartmental models) or the individual level (e.g., agent-based models) [14]. In the United States, Ginsberg et al. [15] were the first to use Google query data to predict ILI rates, creating an influenza trend model that monitored real-time data from millions of search engine queries. However, this model exhibited limitations common to many previous surveillance systems. Other research teams have subsequently utilized various online platforms, including Twitter, Weibo, Baidu, Yahoo, and Google, to develop models for improved accuracy in identifying influenza trends [16]. These limitations of traditional methods highlight the need for more advanced approaches, particularly in managing complex patterns and nonlinear dynamics.

Advancing Machine Learning Models

Recent advancements in machine learning have demonstrated exceptional performance in predicting influenza trends, particularly through models that analyze high-dimensional

and nonlinear data. Notable examples include recurrent neural networks (RNNs), random forests, and support vector machines [17–19]. Prior research has confirmed the effectiveness of machine learning methods for predicting various epidemic diseases across different countries. Among these, deep learning techniques—especially long short-term memory (LSTM) networks—have emerged as powerful tools for forecasting complex time series data [20]. LSTM models excel at modeling long-term temporal dependencies and irregular patterns, significantly surpassing traditional forecasting methods like seasonal autoregressive integrated moving average (SARIMA) and Holt-Winters in accurately predicting ILI trends [21,22]. Bidirectional LSTM networks extend the capabilities of LSTM by capturing patterns in both forward and backward temporal directions, providing superior predictive accuracy for both short- and long-term trends, particularly when temporal correlations are complex [23]. However, achieving optimal performance with these advanced models requires data of sufficient quality and quantity, which can be a notable limitation.

Despite these technological advancements, Saudi Arabia and other Middle Eastern countries continue to underutilize deep learning models for ILI forecasting. Most research in the region still relies heavily on conventional statistical methods, which effectively identify seasonal trends but fall short of capturing intricate asymmetrical dynamics [24]. Moreover, earlier studies generally overlooked exogenous variables—such as vaccination rates, climatic conditions, and population mobility—which are crucial for enhancing predictive accuracy. One study in the Asia-Pacific region incorporated migration trends and demographic data to significantly improve predictive performance [25]. Additionally, other researchers have demonstrated the advantages of employing hybrid forecasting models, combining statistical and machine learning methods for predicting ILI trends across various regions [26,27]. Gomez-Cravioto et al. [28] and Budiharto [18] highlighted the efficiency of LSTM models in predicting nonlinear patterns associated with seasonal variability. By integrating key factors such as climate variations, demographic patterns, population migration, and vaccination rates, the current study addresses existing knowledge gaps and aims to enhance the forecasting accuracy of ILIs in Saudi Arabia.

Emerging Machine Learning in the Middle East

Various methods have been applied by Middle Eastern countries for forecasting ILIs, but these efforts remain in an ongoing developmental stage [24]. A limited number of studies have employed traditional forecasting methods, such as autoregressive integrated moving average (ARIMA) and SARIMA models, particularly during pandemics. Recently, a novel extreme gradient boosting (XGBoost) model demonstrated improved accuracy in forecasting monthly influenza case numbers in Saudi Arabia, surpassing traditional methods in pandemic scenarios [29]. Another study conducted in Syria compared multiple forecasting models using weekly ILI data collected via the Early Warning Alert and Response System. Although this study observed forecasting improvements, it noted significant limitations regarding data quality and availability [30]. Table 1 provides a comparative overview of traditional and machine learning methods. During the coronavirus disease 2019 (COVID-19) pandemic, Saudi Arabia leveraged machine-learning models for outbreak prediction, enhancing early disease detection, screening, prognosis, and facilitating the development of proactive strategies to mitigate disease impacts [8]. These findings underscore the underutilized potential of advanced machine-learning approaches in Middle Eastern countries, highlighting the need to address this gap.

Hybrid models, which integrate the strengths of traditional statistical approaches and machine learning techniques, offer the potential to develop more accurate and interpretable

Table 1. Comparison of traditional and machine learning models

Features	Traditional models	Machine learning models
Model types	SARIMA, exponential smoothing, Holt-Winters	Random forests, support vector machines, recurrent neural networks (e.g., LSTM, BiLSTM, ConvLSTM)
Data handling	Primarily univariate time series data	Capable of handling high-dimensional, nonlinear, and multivariate data
Trend and pattern capture	Effective in identifying linear and seasonal trends	Superior in capturing complex nonlinear dynamics and long-term dependencies
Predictive accuracy	Adequate for short-term forecasts; may struggle with irregular patterns	Demonstrated higher accuracy in various studies, especially for long-term and complex pattern forecasting
Applications in the Middle East	Widely used in regional studies; limited integration of exogenous variables	Underutilized; potential for enhanced accuracy by incorporating factors like climate, mobility, and vaccination rates

SARIMA, seasonal autoregressive integrated moving average; LSTM, long short-term memory; BiLSTM, bidirectional LSTM.

https://doi.org/10.24171/j.phrp.2025.0080

forecasting systems. Despite advancements in ILI forecasting, there has been limited comparative analysis between traditional statistical methods—such as the Holt-Winters model—and advanced deep learning techniques, particularly LSTM networks, in Saudi Arabia. Most research has focused exclusively on individual modeling approaches, often neglecting the influence of seasonal fluctuations and mass gatherings, such as Hajj and Umrah, on forecasting performance. An improved understanding of the factors that improve predictive accuracy could significantly optimize public health interventions and enhance early warning systems.

The aims of this study were as follows: (1) To compare the performance of the traditional Holt-Winters statistical model and the deep learning-based LSTM model in forecasting weekly seasonal ILI incidences in Saudi Arabia. (2) To evaluate the predictive accuracy of both models using the root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) metrics. (3) To enhance public health awareness by identifying an adaptable, data-driven forecasting model tailored to the unique epidemiological trends and public health challenges in Saudi Arabia.

Materials and Methods

Study Design and Data Collection

This study evaluates the performance of the Holt-Winters and LSTM models for forecasting weekly ILI cases in Saudi Arabia. Data covering a 6-year period from 2017 to 2022 were obtained from the Saudi Ministry of Health and the WHO. Weekly data were collected from all regions of Saudi Arabia. Exogenous variables, such as climatic conditions, population mobility trends, and vaccination rates, were included to improve predictive accuracy.

Holt Linear Model

The Holt-Winters forecasting method enhances univariate time series modeling [31–33]. As a statistical approach, the Holt-Winters method effectively captures seasonal and trend components, making it particularly suitable for short-term forecasting of seasonal data due to its computational efficiency and ease of implementation. Depending on the characteristics of the seasonal component, either additive or multiplicative Holt-Winters models are utilized [34–36]. Additive models are appropriate for data exhibiting linear or stable exponential trends without significant growth over time [37].

The additive Holt-Winters model is mathematically represented as follows [29,30]:

Forecasting:

$$(1)\ \widehat{x}_{t+h|t} = L_t + h.T_t + S_{t-m+1+(h-1)(m)},$$

where:
$\widehat{x}_{t+h|t}$: Forecasted value for time $t+h$,
L_t: Level component at time t,
T_t: Trend component at time t,
$S_{t-m+1+(h-1)(m)}$: Seasonal component for the forecasted period,
m: Length of the seasonal cycle.
Level (L_t):

$$(2)\ L_t = \alpha(x_t - S_{t-m}) + (1-\alpha)(L_{t-1} + T_{t-1}),$$

Trend (T_t):

$$(3)\ T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1},$$

Seasonality (S_t):

$$(4)\ S_t = \gamma(x_t - L_t) + (1-\gamma)S_{t-m},$$

where:
x_t: Observed value at time t,
S_{t-m}: Seasonal component from m periods ago,
a, β, γ: Smoothing parameters for level, trend, and seasonality,
$0 < a, \beta, \gamma < 1$,

The multiplicative model is formally expressed as follows [24,25]:

Forecasting:

$$(5)\ \widehat{x}_{t+h|t} = (L_t + h.T_t).S_{t-m+1+(h-1)()m},$$

$\widehat{x}_{t+h|t}$: Forecasted value for time $t+h$,
L_t: Level component at time t,
T_t: Trend component at time t,
$S_{t-m+1+(h-1)(m)}$: Seasonal component for the forecasted period,
m: Length of the seasonal cycle.
Level (L_t):

$$(6)\ L_t = \alpha\frac{x_t}{S_{t-m}} + (1-\alpha)(L_{t-1} + T_{t-1}),$$

Trend (T_t):

$$(7)\ T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1},$$

Seasonality (S_t):

(8) $S_t = \gamma \frac{x_t}{L_t} + (1-\gamma) S_{t-m},$

where:
x_t: Observed value at time t,
S_{t-m}: Seasonal component from m periods ago,
α, β, γ: Smoothing parameters for level, trend and seasonality, $0<\alpha,\beta,\gamma<1$

Long Short-Term Memory

LSTM is a specific type of RNN. It enhances traditional RNNs through an improved structure containing 3 gate mechanisms—forget, input, and output gates—as well as a cell state [38]. This enhanced structure is crucial for epidemiological forecasting, as it effectively manages abrupt changes and irregular patterns driven by climatic variations, public health interventions, or mass gatherings. LSTM models overcome many limitations of traditional statistical methods by better capturing long-term dependencies and nonlinear interactions. Figure 1 illustrates the architecture of the LSTM model. The following paragraphs detail the architecture and gate mechanisms of the LSTM model [25].

The forget gate reviews the current time step's input information, denoted as s_t, and the previous time step's output information, denoted as h_{t-1}. When $f_t=0$, the gate discards the read information. Conversely, when $f_t=1$, it retains the read information. The formula for f_t is as follows [39]:

(9) $f_t = \sigma(w_f \cdot [h_{t-1}. x_t] + b_f),$

This gate determines which new input information to store in the neuron. It creates a candidate cell state $\overline{C}_t$, and the input gate updates the candidate cell state. The new information is subsequently added to the cell state. The specific formula is as follows [40]:

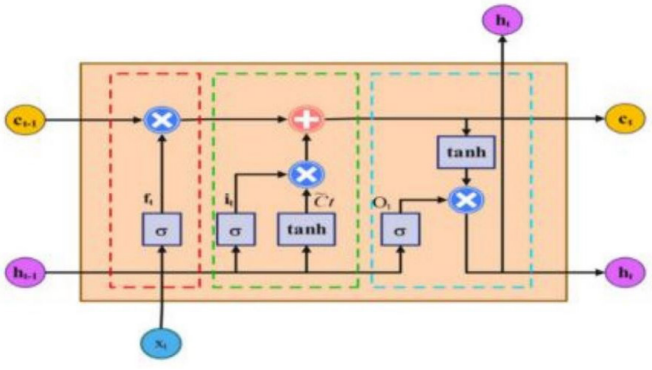


Figure 1. The architecture of the long short-term memory model and its gate structures.

(10) $\overline{C}_t = \tanh(w_c. [h_{t-1} \cdot x_t] + b_c),$

(11) $i_t = \sigma(w_i. [h_{t-1} \cdot x_t] + b_i),$

(12) $c_t = f_t. c_{t-1} + i_t \cdot \overline{C}_t,$

In the above formulas, w_c is the weight matrix for the cell state, b_c is the bias coefficient for the cell state, w_i is the weight matrix for the input gate, and b_i is the bias coefficient for the input gate.

The output gate determines the final output h_t using the cell state. It starts by processing the current input information x_t and the previous output information h_{t-1}. Then, it multiplies these values by the cell state processed by the tanh layer to obtain the final output h_t. The specific formula is as follows [41]:

(13) $o_t = \sigma(w_o. [h_{t-1} \cdot x_t] + b_o),$

(14) $h_t = o_t \cdot \tanh(c_t),$

in these formulas, w_o is the weight matrix for the output gate, and b_o is the bias coefficient for the output gate.

Both models are consistently employed throughout this study, leveraging combined statistical and machine learning strengths. Combining Holt-Winters and LSTM methods potentially enhances forecasting accuracy by better adjusting for seasonal variations. Optimizing hyperparameters is crucial for enhancing model performance, particularly when predicting unseen or new data [42,43]. Hyperparameter tuning involves selecting and adjusting parameters such as the optimizer type, number of epochs, number of neurons, activation functions, and loss functions. Appropriate optimization levels significantly improve the accuracy and generalizability of forecasting outcomes.

Cross-Validation for Model Optimization

Cross-validation is a systematic method used to explore combinations of hyperparameters, helping to identify the optimal model configuration. This approach is especially valuable for time-sensitive applications such as disease forecasting [44]. It is widely employed to estimate prediction errors and focuses explicitly on preventing model overfitting. A commonly used cross-validation method for optimizing LSTM models is Monte Carlo cross-validation [45]. For LSTM models, the optimizer's learning rate (e.g., Adam or RMSprop) regulates the step size for updating model parameters (Figure 2). Overfitting is typically mitigated by incorporating a dropout rate, which randomly disables neurons during training. Additionally, the number of epochs and batch size

https://doi.org/10.24171/j.phrp.2025.0080

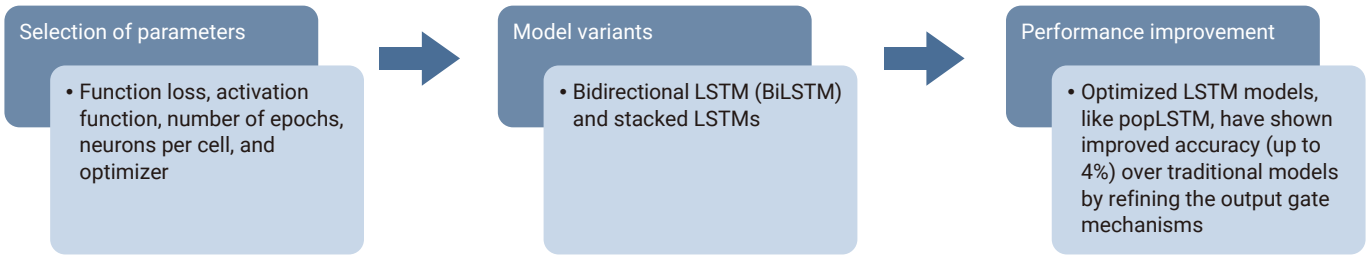


Figure 2. Hyperparameter tuning techniques.
LSTM, long short-term memory; popLSTM, proposed optimized LSTM.

play crucial roles: increasing epochs can capture complex patterns but risk overfitting, whereas batch size influences both training duration and model generalization [42].

Frequently used metrics for evaluating prediction accuracy include MAE, mean squared error (MSE), RMSE, and the coefficient of determination (R^2) [42,43]. The MAE quantifies the average magnitude of prediction errors, representing the typical error magnitude expected from forecasts [46–50]:

$$(15)\ \mathrm{MAE} = \frac{\sum_{j=1}^{n}|Z_j - Z|}{n}$$

Where: n is the number of errors, and $|Z_j\text{-}Z|$ the absolute errors.

Eq. (15) defines the RMSE, which represents the standard deviation of the prediction errors.

$$(16)\ \mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\chi_i - \widehat{\chi}_i\right)^2}{n}}$$

R^2 quantifies the proportion of the predicted weekly ILI cases that can be explained by the predictor variables, as defined in Eq. (16).

$$(17)\ R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\chi_i - \widehat{\chi}_i\right)^2}{\sum_{i=1}^{n}\left(\chi_i - \overline{\overline{\chi_l}}\right)^2}$$

where $\widehat{\chi}_i$ represents the predicted value of the i^{th} sample and *represents* the matching real value for the whole n samples.

PBIAS (%)

The percent bias (PBIAS) quantifies the amount of bias in the model by assessing the cumulative discrepancy between observed and predicted values relative to the total observed values:

$$(18)\ \mathrm{PBIAS} = \frac{\sum_{i=1}^{n}\left(\chi_i - \widehat{\chi}_i\right)}{\sum_{i=1}^{n}(\chi_i)}$$

where,
x_i is the observed value,
$\widehat{\chi}_i$ is the predicted value,
n is the number of observations.
The magnitude of the errors that may be predicted from the average forecast is denoted by PBIAS [51–53].

Willmott's Index of Agreement

Willmott's index of agreement (WI) is a dimensionless statistic used to evaluate the magnitude and direction of predictive error across various models [52,54,55]. It is particularly useful for comparing predictive accuracy between models:

$$(19)\ WI = \frac{\sum_{i=1}^{n}\left(\widehat{\chi}_i - \chi_i\right)^2}{\sum_{i=1}^{n}\left(\left|\widehat{\chi}_i - \overline{\chi_i}\right| + |\chi_i - \overline{\chi_i}|\right)^2}$$

where:
x_i is the observed value,
$\widehat{\chi}_i$ is the predicted value,
n is the number of observations.
$\overline{\chi_i}$ is the mean of observed values.
The workflow is presented in Figure 3.

Results

Data Characterization

Temporal epidemiological analysis was performed on ILI trends, highlighting variability in disease incidence across different seasons and years. Weekly data on ILIs in Saudi Arabia, spanning from January 2017 to December 2022, were collected from the Saudi Ministry of Health and the WHO. The dataset encompasses all regions of Saudi Arabia, providing comprehensive national-level insight into ILI dynamics.

Data Preparation

To ensure data accuracy, consistency, and readiness for analysis, the following pre-processing steps were implemented: (1) Missing data points were imputed using interpolation. (2) Data were normalized using min-max scaling to ensure compatibility with the LSTM algorithms, rescaling ILI case counts to (0, 1). This supports faster convergence in LSTM. (3) Data were split into training (80%), validation (10%), and testing (10%) sets.

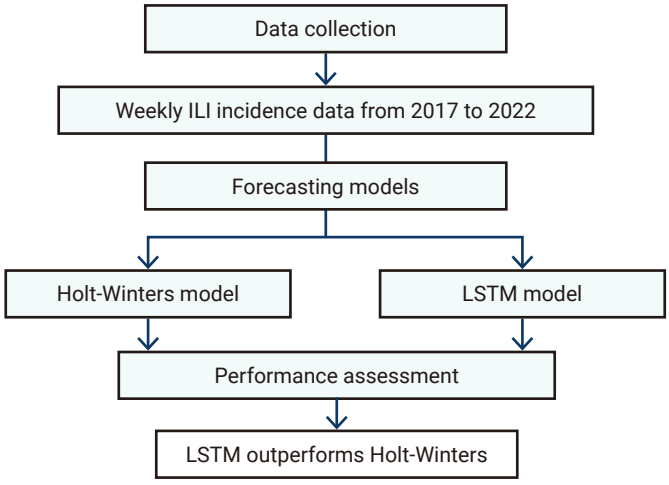


Figure 3. Representation of the entire workflow.
LSTM, long short-term memory; ILI, influenza-like illness.

Characteristics of weekly ILI cases

Figure 4 illustrates clear, recurrent peaks at regular intervals, demonstrating seasonal patterns in weekly ILI incidence. The peaks correspond with defined seasons, particularly winter, and significant mass events such as the Hajj pilgrimage season. A noticeable decline in ILI cases occurred between 2020 and 2021, reflecting stringent preventive measures implemented in response to COVID-19. This decline reflects the impact of public health interventions on disease spread.

Forecasting weekly ILI cases utilizing the Holt-Winters model

The Holt-Winters forecasting models were developed and assessed using additive and multiplicative seasonal components based on the training dataset. A systematic grid search identified the optimal model configuration, selecting the model with the lowest RMSE for the test dataset. Comprehensive evaluation identified the Holt-Winters multiplicative model as the best-performing configuration, and multiple metrics were computed to evaluate its performance (Table 2). These metrics affirmed the model's capability to capture seasonal trends effectively. The smoothing parameters for level, trend, and seasonality were automatically optimized based on data characteristics.

Figure 5 presents the actual and predicted values generated by the Holt-Winters model. The graph demonstrates that the model effectively captured general seasonal patterns and long-term trends; however, deviations are evident during

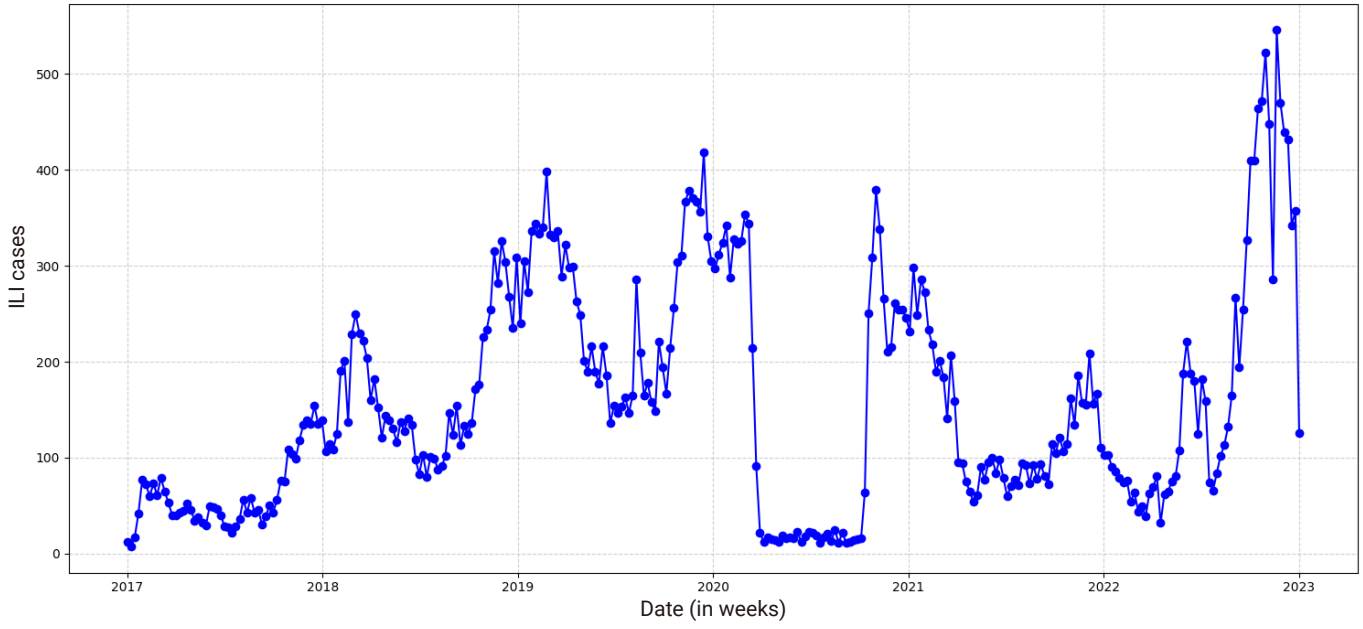


Figure 4. Weekly influenza-like illness (ILI) incidence in Saudi Arabia from 2017 to 2023, showing seasonal peaks and fluctuations.

https://doi.org/10.24171/j.phrp.2025.0080

Table 2. Evaluation of the performance of 2 models

Model	RMSE	MAPE	R^2	PBIAS	WI
Holt-Winters	82.57 (95% CI, 80.21–85.12)	0.38 (95% CI, 0.36–0.40)	0.58	+14.2%	105.79
LSTM	34.07 (95% CI, 32.45–35.62)	0.18 (95% CI, 0.16–0.19)	0.93	+5.8%	0.48
LSTM (with weather+mobility+vaccination)	28.55	0.14	0.96	+2.1%	

RMSE, root mean squared error; MAPE, mean absolute percentage error; PBIAS, percent bias; WI, Willmott's index of agreement; CI, confidence interval; LSTM, long short-term memory.

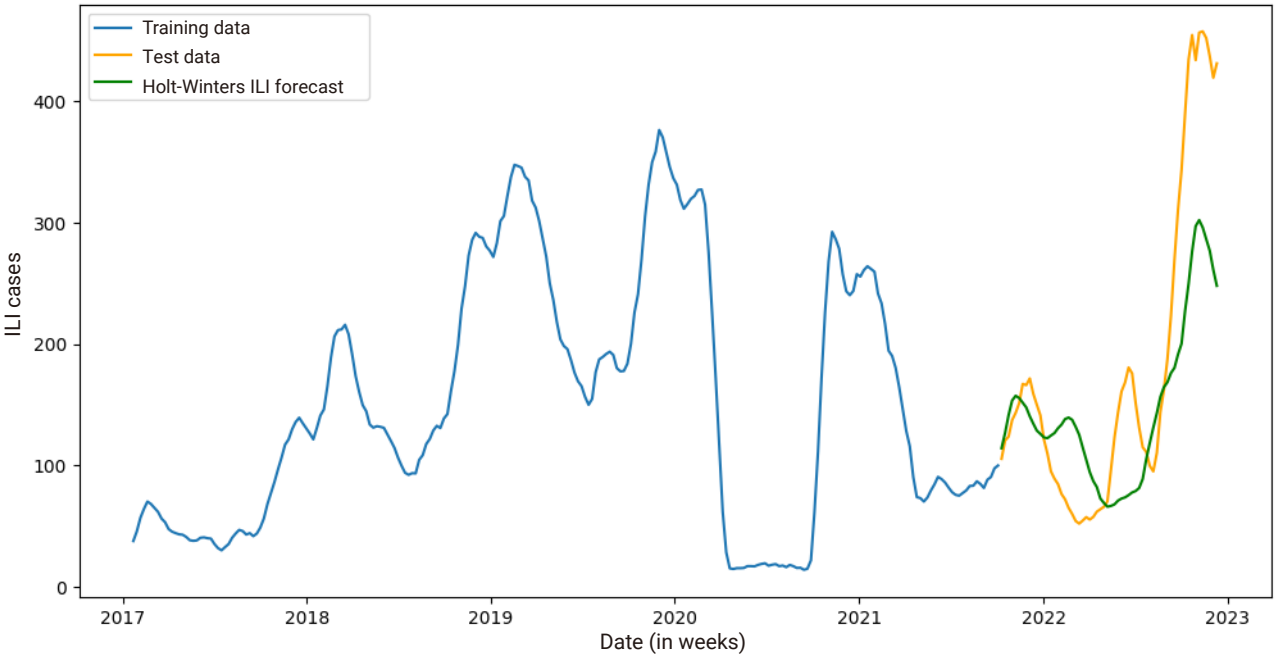


Figure 5. Holt-Winters forecasting of influenza-like illness (ILI) with training and test datasets.

periods of high volatility. Notably, the model substantially underestimated ILI cases during seasonal peaks, such as the Hajj season, and struggled to adapt to the sudden case declines during 2020–2021 attributed to COVID-19 prevention measures

Forecasting weekly ILI cases utilizing the LSTM model

For the LSTM model, data normalization was performed using the min-max Scaler, and sequences were generated using a 10-week timestep to provide adequate temporal context. The model was trained across 20 epochs using a batch size of 32, employing the Adam optimizer and MSE as the loss function to reduce forecasting inaccuracies. Model performance was evaluated, as summarized in Table 1. Figures 5 and 6 emphasize discrepancies between observed and predicted values, clearly showing that the Holt-Winters model struggled with high-variability periods, while the LSTM model maintained greater stability.

Figure 5 demonstrates the Holt-Winters model's forecast for weekly ILI cases from 2017 to 2023, comparing training and test datasets with predictions. Although the Holt-Winters model captured seasonality and general trends reasonably well, it struggled with sudden spikes during events such as the Hajj season and significant variability during the COVID-19 pandemic. These deviations reflect the model's inability to fully account for nonlinear patterns, suggesting that incorporating exogenous variables or using hybrid modeling approaches could enhance forecast performance.

Figure 6 shows that the LSTM model effectively captured both seasonal variations and long-term trends in ILI incidence. The graph compares observed data from the training set, test set, and LSTM predictions, clearly illustrating the model's capacity to discern patterns accurately and generate reliable predictions during the testing period. While the model generally exhibited strong performance, minor inconsistencies were observed during periods of rapid variation in ILI incidence, indicating opportunities for further

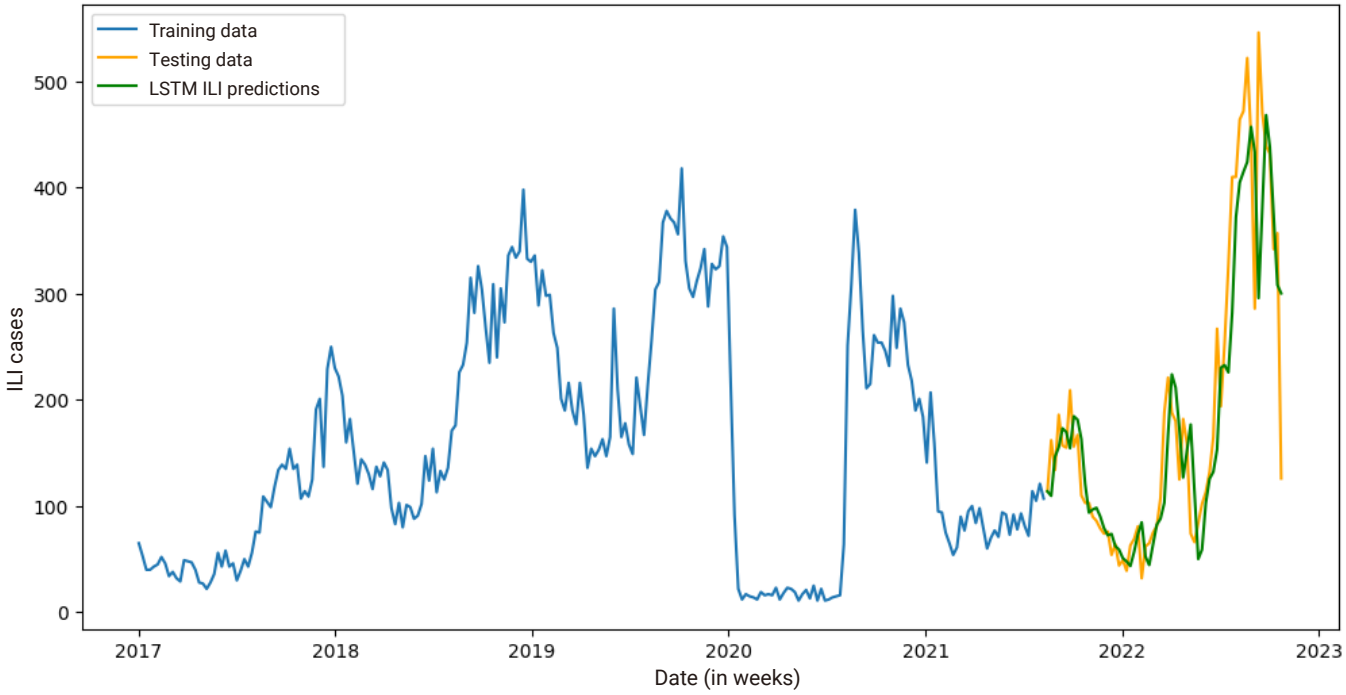


Figure 6. Long short-term memory (LSTM) forecasting of influenza-like illness (ILI) with training and test datasets.

improvement through feature engineering or additional hyperparameter optimization. LSTM predictions closely tracked actual values, showing only slight discrepancies during abrupt incidence changes. Incorporating exogenous variables such as mobility patterns, climatic factors, and vaccination rates, along with hyperparameter fine-tuning, could further enhance predictive accuracy. The LSTM model notably demonstrated superior accuracy in handling temporal dependencies and complex epidemiological patterns.

Figure 7 highlights a sudden peak during the Hajj season, illustrating that the LSTM model predicted these spikes more accurately than the Holt-Winters model. The LSTM's superior performance is attributed to its advanced capacity for handling temporal dependencies and retaining historical information. However, brief periods of high volatility showed slight deviations in LSTM predictions, suggesting that additional feature engineering may be beneficial (Table 2).

To comprehensively assess forecasting performance, statistical metrics were used, including RMSE, MAPE, R^2, PBIAS [111], and the WI [111]. RMSE and MAPE assess prediction accuracy by quantifying absolute and relative errors, respectively, while R^2 indicates the proportion of variance explained by the model. PBIAS provides insights into the direction and magnitude of systematic model bias, determining if predictions are generally overestimated or underestimated relative to observations, particularly beneficial for long-term forecasts. The WI evaluates overall predictive performance, quantifying the proximity between observed and predicted values, further illustrated by the Taylor diagram and box plot of prediction errors (Figures 8, 9). As summarized in Table 2, the LSTM model significantly outperformed the Holt-Winters model, with notably lower RMSE (34.07 vs. 82.57), reduced MAPE (0.18 vs. 0.38), higher R^2 (0.93 vs. 0.58), and substantially improved PBIAS (+5.8% vs. +14.2%) and WI (0.48 vs. 105.79) scores.

These findings strongly support the conclusion that LSTM models are more suitable than traditional statistical approaches for modeling complex, nonlinear, and seasonally varying time-series data related to ILI trends.

Discussion

This study compared the predictive performance of Holt-Winters and LSTM models for forecasting weekly ILI cases in Saudi Arabia, emphasizing the inclusion of exogenous variables such as climatic conditions, population mobility, and vaccination coverage. The results demonstrated that the LSTM model, particularly when incorporating exogenous variables, consistently outperformed the Holt-Winters model across all evaluation metrics. Specifically, the LSTM model achieved an RMSE of 28.55 and an R^2 of 0.96, highlighting its superior capability to capture nonlinear trends and significant peaks in disease incidence, particularly during high-variability

https://doi.org/10.24171/j.phrp.2025.0080

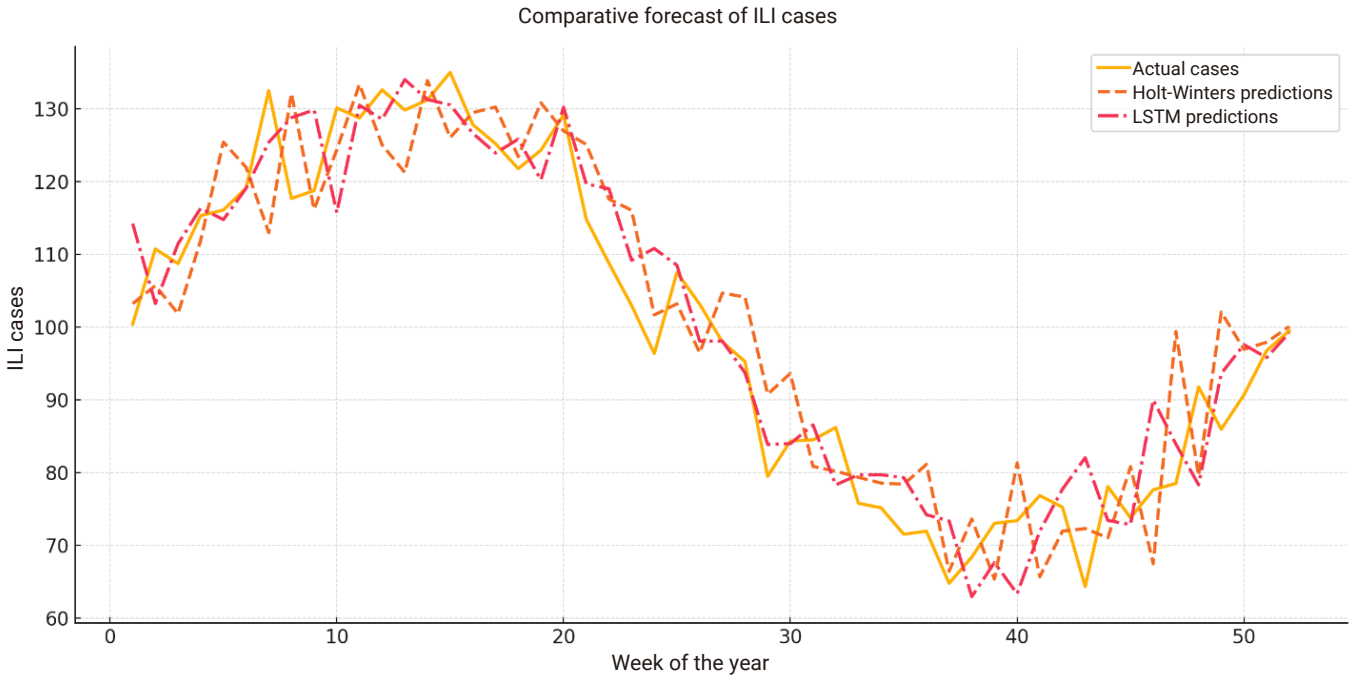


Figure 7. Comparative forecast of influenza-like illness (ILI) cases. LSTM, long short-term memory.

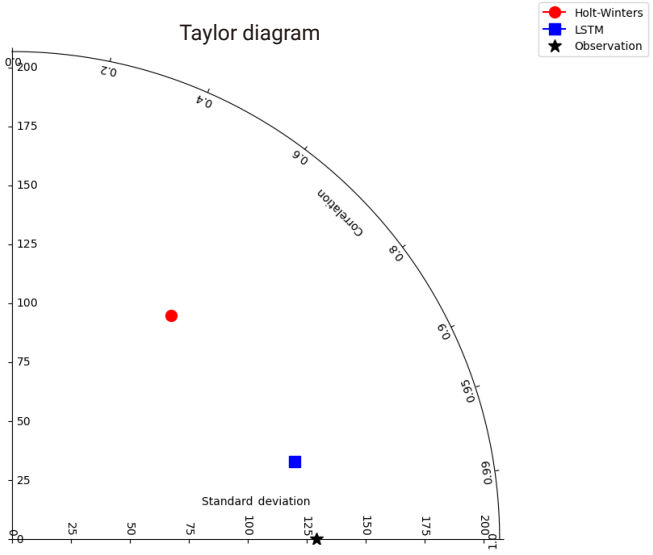


Figure 8. Taylor diagram. LSTM, long short-term memory.

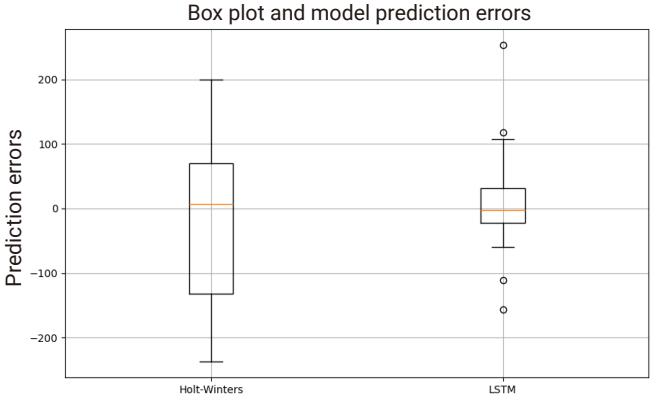


Figure 9. Box plot and model prediction error analysis. LSTM, long short-term memory.

periods like the Hajj season. Both additive and multiplicative seasonal components were methodically examined using a grid search methodology to determine the optimal model configuration. This systematic evaluation allowed exploration of multiple parameter combinations, ensuring the selection of the most effective model configuration by balancing computational efficiency and accuracy. Although the Holt-Winters model provides an intuitive framework for addressing seasonality, differences emerged in forecasts during periods of elevated ILI variability. These discrepancies highlight that, despite its proficiency in capturing general trends and predictable seasonal patterns, the Holt-Winters model struggles with nonlinear dynamics and abrupt changes typical of epidemiological data. A previous study indicated similar challenges, noting that the model may fail to predict sudden spikes or drops in disease incidence due to unforeseen events or interventions. Moreover, Holt-Winters models inherently assume consistent seasonal patterns, an assumption frequently invalid in epidemiological contexts where seasonality may vary considerably [34].

In contrast, the LSTM model demonstrated enhanced performance across all assessment criteria, including a statistically significant reduction in RMSE (34.07 vs. 82.57), reduced MAPE (0.18 vs. 0.38), and increased R^2 (0.93 vs. 0.58). These metrics underscore the practical advantage of LSTM, facilitating more accurate disease forecasts and enabling timely public health interventions. The improved metrics also reflect the LSTM model's strong capability in learning sequential dependencies and complex nonlinear interactions inherent in epidemiological time series data. By utilizing 2 stacked LSTM layers combined with fully connected layers, the model effectively captured intricate temporal patterns. Stacked LSTM architecture enhances the model's ability to recognize complex temporal dynamics and nonlinear data dependencies. Each additional layer of LSTM allows for deeper, more abstract representations of the data, significantly improving performance on forecasting tasks [56]. Despite these advancements, the LSTM model encountered challenges when handling sudden fluctuations in ILI incidence rates. Occasional discrepancies during periods of rapid change likely reflect issues such as limited data availability or inherent randomness in epidemiological data, indicating room for further model refinement.

Compared to related research, the current study offers several methodological improvements. For instance, Khan et al. [57] introduced a cloud-based modeling system for the influenza pandemic using a feed-forward propagation neural network (MSDII-FFNN) model for pandemic influenza forecasting using Internet of Things-generated data [57]. While comparable in its real-time monitoring objective, their approach differed methodologically by focusing more on infrastructure integration rather than methodological optimization. Unlike our study, Khan et al. [57] did not utilize standardized performance metrics or baseline controls, thus limiting the rigorous evaluation of their model's reliability and accuracy. Similarly, Alzahrani and Guma [58] utilized ARIMA, SARIMA, and XGBoost models to predict monthly influenza cases in Saudi Arabia. Their findings align with our results in demonstrating machine learning's superiority over traditional statistical methods. However, their study differed significantly in terms of scope and temporal resolution, employing monthly data and omitting exogenous variables, thus restricting their model's applicability for predicting rapid epidemiological changes such as those associated with the Hajj pilgrimage or unexpected outbreaks.

Meanwhile, the study by Olukanmi et al. [59] is similar to ours in employing deep learning techniques for direct weekly ILI predictions. Their approach, however, was distinguished by using digital behavior data from Google Trends rather than structured environmental data. While both studies highlight the strengths of LSTM models, they demonstrated the effectiveness of real-time public data in identifying early symptom-related searches, whereas our study emphasized epidemiological and environmental proxies to further improve predictive performance. These comparisons illustrate how different data inputs—such as digital search trends in Olukanmi et al. [59], cloud-embedded monitoring in Khan et al., and monthly aggregated case data in Alzahrani and Guma [58]—influence model outcomes. Our research positions itself between these methodologies by utilizing formal, real-world weekly data with additional exogenous variables, achieving a balance of timeliness, precision, and policy relevance. Developing interpretable hybrid models capable of integration into public health systems will be crucial for timely and accurate epidemic forecasting across diverse epidemiological contexts.

However, other limitations must be considered. Training deep learning models such as LSTM is resource-intensive, demanding substantial computational power, specialized hardware, and memory [43,60]. Conversely, the Holt-Winters model is more suited to real-time applications due to its computational efficiency and lower resource requirements. Nonetheless, this computational efficiency compromises prediction accuracy when dealing with complex, nonlinear epidemiological data.

Integrating advanced characteristics—such as population mobility patterns, climatic variations, and vaccination rates—can significantly enhance a model's ability to predict seasonal peaks and fluctuations associated with mass gatherings, notably the Hajj pilgrimage, thereby deepening our understanding of influenza transmission dynamics [61,62]. Certain models can dynamically adjust real-time forecasts according to changing environmental factors. Improved epidemiological models may focus more closely on population migration patterns to enhance disease transmission predictions [61,63]. Moreover, climatic factors like temperature and humidity have previously improved the accuracy of epidemiological models for COVID-19 transmission projections [64]. Similarly, predictive models in the United States have enhanced performance by explicitly incorporating population migration rates as exogenous variables [64]. Concentrating on these factors could allow machine learning models to predict changes in ILI incidence more accurately, ultimately facilitating more effective public health strategies.

When comparing machine learning models, interpretability and simplicity are key strengths of the Holt-Winters model, despite its performance decline in complex scenarios characterized by temporal dependencies or sudden shifts. Conversely, long-term machine learning models such as

https://doi.org/10.24171/j.phrp.2025.0080

LSTM demonstrate robust capabilities in representing these complexities but require careful hyperparameter tuning, substantial computational resources, and strategies to mitigate overfitting. These findings underscore the appropriateness of advanced machine learning methods, like LSTM, for predicting influenza and related illnesses within dynamic epidemic settings. In stable conditions, the Holt-Winters model remains beneficial due to its efficiency and interpretability. Future studies could explore hybrid methodologies combining statistical and machine learning models, capitalizing on their complementary strengths to enhance predictive performance. Previous research has demonstrated the effectiveness of hybrid models—for instance, combining LSTM with ARIMA in time-series forecasting has yielded improvements in predictive accuracy and adaptability within epidemiological modeling [62]. Additionally, hybrid deep learning approaches integrating convolutional neural networks with LSTMs have effectively predicted disease trends, enhancing both feature extraction and sequential learning capabilities [63]. These hybrid models could achieve even greater accuracy when enriched with exogenous variables such as population migration, vaccination rates, or climatic conditions.

Limitations and Future Work

This study's models had limited capacity to fully adjust for external factors affecting disease transmission, focusing predominantly on historical ILI case data. Incorporating exogenous variables such as population mobility, vaccination coverage, and environmental conditions (temperature and humidity) may further improve forecasting accuracy. Still, the models faced challenges accommodating abrupt shifts and inherent nonlinearities in the data, illustrating the broader challenge of relying solely on statistical approaches in highly volatile datasets. Nevertheless, the LSTM model consistently performed better across all evaluation metrics, reflecting its strong capability to capture temporal dependencies and nonlinear interactions. Lower RMSE and MAPE values combined with a higher R^2 indicate that LSTM is particularly suitable and effective for predicting ILI cases in dynamic and complex epidemic scenarios. Predictions occasionally deviated during rapid data fluctuations, indicating potential improvements through additional feature enrichment and hyperparameter tuning. Previous studies have shown that incorporating mobility data into COVID-19 models and environmental variables into epidemiological forecasts enhances predictive performance [61,63,64]. Future research efforts should focus on developing hybrid models combining statistical and machine learning approaches to leverage their respective strengths. Moreover, integrating exogenous variables such as meteorological data, migration patterns, vaccination coverage, and public health interventions could yield more precise and practically valuable predictions. The findings from this study contribute to improved methods for forecasting ILI cases, facilitating public health planning and enhancing intervention strategies for seasonal and emerging infectious diseases. For instance, the ARIMAX model successfully incorporates external variables into epidemiological predictions. Combining ARIMAX with LSTM models could further enhance predictive accuracy while retaining interpretability. Future work might also involve public health experts in model development, thereby improving adoption and utilization within healthcare systems. Additionally, the integration of real-time data streams and adaptive modeling techniques warrants further investigation to better accommodate evolving epidemiological trends.

Conclusion

This study explored the strengths and limitations of Holt-Winters and LSTM models for predicting weekly ILI cases in Saudi Arabia. The Holt-Winters model effectively captured seasonal patterns and general trends but struggled significantly with sudden shifts in incidence. In contrast, the LSTM model delivered superior prediction accuracy, demonstrating exceptional proficiency in modeling nonlinear patterns and temporal relationships. Nevertheless, the LSTM model remains sensitive to hyperparameter adjustments and requires considerable computational resources. Future research should explore hybrid modeling approaches, integrating deep learning and statistical methods alongside exogenous variables to improve predictive accuracy further. These advances will significantly enhance epidemic forecasting, benefiting public health planning and response strategies. Thus, this study reinforces that novel machine learning approaches, such as LSTM, generally yield superior predictive results for complex epidemiological scenarios, whereas classical statistical models like Holt-Winters remain faster, simpler, and valuable in more stable and predictable contexts.

Notes

Ethics Approval
Not applicable.

Conflicts of Interest
The author has no conflicts of interest to declare.

Funding
None.

Availability of Data
Data are available upon reasonable request.

References

1. Haider S, Hassan MZ. Seasonal influenza surveillance and vaccination policies in the WHO South-East Asian Region. BMJ Glob Health 2025; 10:e017271.
2. Fiandrino S, Bizzotto A, Guzzetta G, et al. Collaborative forecasting of influenza-like illness in Italy: the Influcast experience. Epidemics 2025;50:100819.
3. Babcock HM, Merz LR, Fraser VJ. Is influenza an influenza-like illness?: clinical presentation of influenza in hospitalized patients. Infect Control Hosp Epidemiol 2006;27:266–70.
4. Alsubaie N, EL Guma F, Boulehmi K, et al. Improving influenza epidemiological models under Caputo fractional-order calculus. Symmetry 2024;16:929.
5. Rashid H, Shafi S, Haworth E, et al. Influenza vaccine in Hajj pilgrims: policy issues from field studies. Vaccine 2008;26:4809–12.
6. Liu IT, Prasad V, Darrow JJ. Evidence for community face masking to limit the spread of SARS-CoV-2: a critical review. Health Matrix 2023; 33:1.
7. Alshahrani SM, Zahrani Y. Prevalence and predictors of seasonal influenza vaccine uptake in Saudi Arabia post COVID-19: a web-based online cross-sectional study. Vaccines (Basel) 2023;11:353.
8. Nagraj VP, Benefield AE, Williams D, et al. PLANES: plausibility analysis of epidemiological signals. PLoS One 2025;20:e0320442.
9. Nsoesie EO, Brownstein JS, Ramakrishnan N, et al. A systematic review of studies on forecasting the dynamics of influenza outbreaks. Influenza Other Respir Viruses 2014;8:309–16.
10. Lima S, Goncalves AM, Costa M. Time series forecasting using Holt-Winters exponential smoothing: an application to economic data. AIP Conf Proc 2019;2186:090003.
11. Luo J, Wang X, Fan X, et al. A novel graph neural network based approach for influenza-like illness nowcasting: exploring the interplay of temporal, geographical, and functional spatial features. BMC Public Health 2025;25:408.
12. Meyer AG, Lu F, Clemente L, et al. A prospective real-time transfer learning approach to estimate influenza hospitalizations with limited data. Epidemics 2025;50:100816.
13. Cobb NL, Collier S, Attia EF, et al. Global influenza surveillance systems to detect the spread of influenza-negative influenza-like illness during the COVID-19 pandemic: time series outlier analyses from 2015-2020. PLoS Med 2022;19:e1004035.
14. Guma FE. Comparative analysis of time series prediction models for visceral leishmaniasis: based on SARIMA and LSTM. Appl Math Inf Sci 2024;18:125–32.
15. Ginsberg J, Mohebbi MH, Patel RS, et al. Detecting influenza epidemics using search engine query data. Nature 2009;457:1012–4.
16. Chen Q, Zheng X, Shi H, et al. Prediction of influenza outbreaks in Fuzhou, China: comparative analysis of forecasting models. BMC Public Health 2024;24:1399.
17. Tsan YT, Chen DY, Liu PY, et al. The prediction of influenza-like illness and respiratory disease using LSTM and ARIMA. Int J Environ Res Public Health 2022;19:1858.
18. Budiharto W. Data science approach to stock prices forecasting in Indonesia during COVID-19 using long short-term memory (LSTM). J Big Data 2021;8:47.
19. Akter T, Hossain MF, Ullah MS, et al. Mortality prediction in COVID-19 using time series and machine learning techniques. Comput Math Methods 2024;2024:5891177.
20. Fotis G, Vita V, Ekonomou L. Machine learning techniques for the prediction of the magnetic and electric field of electrostatic discharges. Electronics 2022;11:1858.
21. Zhao D, Zhang R. A new hybrid model SARIMA-ETS-SVR for seasonal influenza incidence prediction in mainland China. J Infect Dev Ctries 2023;17:1581–90.
22. Pavlatos C, Makris E, Fotis G, et al. Enhancing electrical load prediction using a bidirectional LSTM neural network. Electronics 2023;12:4652.
23. Yang J, Yang L, Li G, et al. Influenza time series prediction models in a megacity from 2010 to 2019: based on seasonal autoregressive integrated moving average and deep learning hybrid prediction model. Chin Med J (Engl) 2024;137:2242–4.
24. Keshavamurthy R, Pazdernik KT, Ham C, et al. Meeting global health needs via infectious disease forecasting: development of a reliable data-driven framework. JMIR Public Health Surveill 2025;11:e59971.
25. Rauf HT, Lali MI, Khan MA, et al. Time series forecasting of COVID-19 transmission in Asia Pacific countries using deep neural networks. Pers Ubiquitous Comput 2023;27:733–50.
26. Darwish A, Rahhal Y, Jafar A. A comparative study on predicting influenza outbreaks using different feature spaces: application of influenza-like illness data from Early Warning Alert and Response System in Syria. BMC Res Notes 2020;13:33.
27. Talkhi N, Akhavan Fatemi N, Ataei Z, et al. Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: a comparison of time series forecasting methods. Biomed Signal Process Control 2021;66:102494.
28. Gomez-Cravioto DA, Diaz-Ramos RE, Cantu-Ortiz FJ, et al. Data analysis and forecasting of the COVID-19 spread: a comparison of recurrent neural networks and time series models. Cognit Comput 2021;1–12.
29. Awajan AM, Ismail MT, Al Wadi S. Improving forecasting accuracy for stock market data using EMD-HW bagging. PLoS One 2018;13: e0199582.
30. Swapnarekha H, Behera HS, Nayak J, et al. Multiplicative Holts Winter model for trend analysis and forecasting of COVID-19 spread in India. SN Comput Sci 2021;2:416.
31. Pongdatu GA, Putra YH. Seasonal time series forecasting using SARIMA and Holt Winter's exponential smoothing. IOP Conf Ser Mater Sci Eng 2018;407:012153.
32. Abdullah MM. Using the single-exponential-smoothing time series model under the additive Holt-Winters algorithm with decomposition and residual analysis to forecast the reinsurance-revenues dataset.

https://doi.org/10.24171/j.phrp.2025.0080

Paki J Stat Oper Res 2024;311–40.
33. Ince MN, Tasdemir C. Forecasting retail sales for furniture and furnishing items through the employment of multiple linear regression and Holt-Winters models. Systems 2024;12:219.
34. Jin Y, Wang R, Zhuang X, et al. Prediction of COVID-19 data using an ARIMA-LSTM hybrid forecast model. Mathematics 2022;10:4001.
35. Waqas M, Humphries UW. A critical review of RNN and LSTM variants in hydrological time series predictions. MethodsX 2024;13: 102946.
36. Michankow J, Sakowski P, Slepaczuk R. Mean absolute directional loss as a new loss function for machine learning problems in algorithmic investment strategies. J Comput Sci 2024;81:102375.
37. Makarovskikh T, Abotaleb M. Hyper-parameter tuning for long short-term memory (LSTM) algorithm to forecast a disease spreading. In: 2022 VIII International Conference on Information Technology and Nanotechnology (ITNT); 2022 May 23; Samara, Russian Federation. IEEE; 2022. p. 1–6.
38. Sediyono E, Wahyuni SN, Sembiring I. Optimizing the long short-term memory algorithm to improve the accuracy of infectious diseases prediction. IAES Int J Artif Intell 2024;13:2893–903.
39. Shejul K, Harikrishnan R, Gupta H. The improved integrated Exponential Smoothing based CNN-LSTM algorithm to forecast the day ahead electricity price. MethodsX 2024;13:102923.
40. Sembiring I, Wahyuni SN, Sediyono E. LSTM algorithm optimization for COVID-19 prediction model. Heliyon 2024;10:e26158.
41. Gowriswari S, Brindha S. Hyperparameters optimization using gridsearch cross validation method for machine learning models in predicting diabetes mellitus risk. In: 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT); 2022 Mar 10; Chennai, India. IEEE; 2022. p. 1–4.
42. Funk S, Camacho A, Kucharski AJ, et al. Assessing the performance of real-time epidemic forecasts: a case study of Ebola in the Western Area region of Sierra Leone, 2014-15. PLoS Comput Biol 2019;15: e1006785.
43. Chhabra A, Singh SK, Sharma A, et al. Sustainable and intelligent time-series models for epidemic disease forecasting and analysis. Sustain Technol Entrep 2024;3:100064.
44. Plevris V, Solorzano G, Bakas NP, et al. Investigation of performance metrics in regression analysis and machine learning-based prediction models. In: 8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2022); 2022 Jun 5-9; Norway. ECCOMAS Congress; 2022. p. 2022
45. Das P. PerMat: Performance metrics in predictive modeling [Internet]. PerMat; 2024 [cited 2025 Jun 2]. Available from: http://dx.doi.org/10.32614/cran.package.permat.
46. Ryan TP, Woodall WH. The most-cited statistical papers. J Appl Stat 2005;32:461–74.
47. MacDonald P. An influenza vaccination season like no other: present and future aspects. Pract Nurs 2021;32(Sup3a):S10–4.
48. Plenos M. Time series forecasting using holt-winters exponential smoothing: application to abaca fiber data. Sci J Wars Univ Life Sci 2022;22:17–29.
49. Xiao F. Time series forecasting with stacked long short-term memory networks. arXiv [Preprint] 2020 Nov 2. https://doi.org/10.48550/arXiv.2011.00697.
50. Wanli Z, Youjun T, Xiaomei M. From learning science to computer science: a scientometric review of deeper learning in foreign languages (1993-2024). SAGE Open 2025;15:21582440251322564.
51. Tao H, Abba SI, Al-Areeq AM, et al. Hybridized artificial intelligence models with nature-inspired algorithms for river flow modeling: a comprehensive review, assessment, and possible future research directions. Eng Appl Artif Intell 2024;129:107559.
52. Samantaray S, Sahoo A, Satapathy DP, et al. Suspended sediment load prediction using sparrow search algorithm-based support vector machine model. Sci Rep 2024;14:12889.
53. Samantaray S, Sahoo A, Yaseen ZM, et al. River discharge prediction based multivariate climatological variables using hybridized long short-term memory with nature inspired algorithm. J Hydrol 2025; 649:132453.
54. Mahakur V, Mahakur VK, Samantaray S, et al. Prediction of runoff at ungauged areas employing interpolation techniques and deep learning algorithm. HydroRes 2025;8:265–75.
55. Samantaray S, Ghose DK. Prediction of S12-MKII rainfall simulator experimental runoff data sets using hybrid PSR-SVM-FFA approaches. J Water Clim Chang 2022;13:707–34.
56. Semenza JC, Paz S. Climate change and infectious disease in Europe: impact, projection and adaptation. Lancet Reg Health Eur 2021;9:100230.
57. Khan MA, Abidi WU, Al Ghamdi MA, et al. Forecast the influenza pandemic using machine learning. Comput Mater Contin 2021; 66:331–40.
58. Alzahrani SM, Guma FE. Improving seasonal influenza forecasting using time series machine learning techniques. J Inf Syst Eng Manag 2024;9:30195.
59. Olukanmi SO, Nelwamondo FV, Nwulu NI. Utilizing google search data with deep learning, machine learning and time series modeling to forecast influenza-like illnesses in South Africa. IEEE Access 2021; 9:126822–36.
60. Shaikh F, Khattak HA. Hybrid forecasting model for smart grid using exogenous variables. In: 2024 International Conference on Emerging Trends in Smart Technologies (ICETST); 2024 Oct 10-11; Karachi, Pakistan. IEEE; 2024. p. 1–6.
61. Jiang N, Chu W, Li Y. Modeling, inference, and prediction in mobility-based compartmental models for epidemiology. arXiv [Preprint] 2024 Jun 17. https://doi.org/10.48550/arXiv.2406.12002.
62. Xu B, Zhu Q. Dynamic probabilistic latent variable model with exogenous variables for dynamic anomaly detection. In: 2023 American Control Conference (ACC); 2023 May 31; San Diego, USA. IEEE; 2023. p. 3945–50.
63. Hametner C, Kozek M, Bohler L, et al. Estimation of exogenous drivers

to predict COVID-19 pandemic using a method from nonlinear control theory. Nonlinear Dyn 2021;106:1111–25.

64. Toutiaee M, Li X, Chaudhari Y, et al. Improving COVID-19 forecasting using exogenous variables. arXiv [Preprint] 2021 Jul 20. https://doi.org/10.48550/arXiv.2107.10397.

https://doi.org/10.24171/j.phrp.2025.0080

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2025.0067
Osong Public Health Res Perspect 2025;16(3):285-291

Short Communication

Subacute sclerosing panencephalitis incidence following measles elimination efforts in the Republic of Korea

Young Hwa Lee[1], Jung Hye Byeon[2], Cho Ryok Kang[3], Young June Choe[1,2], Jong-Koo Lee[4]

[1]Allergy Immunology Center, Korea University, Seoul, Republic of Korea
[2]Department of Pediatrics, Korea University Anam Hospital Korea University College of Medicine, Seoul, Republic of Korea
[3]Department of Nursing, Wonkwang University, Iksan, Republic of Korea
[4]National Academy of Medicine of Korea, Seoul, Republic of Korea

ABSTRACT

Objectives: Subacute sclerosing panencephalitis (SSPE) is a rare but fatal neurodegenerative disease caused by persistent measles virus infection. After a significant measles outbreak in 2000–2001, the Republic of Korea implemented a nationwide measles elimination program, which led to a dramatic reduction in measles incidence. This study aimed to evaluate the impact of these measles elimination efforts on the incidence of SSPE in the Republic of Korea.
Methods: This nationwide, population-based retrospective cohort study identified patients newly diagnosed with measles and SSPE between 2007 and 2022, registered in the Health Insurance Review and Assessment Service (HIRA) and Korea Disease Control and Prevention Agency (KDCA) databases. Population-based incidence rates of measles and SSPE were calculated and compared annually.
Results: A total of 236 measles cases (HIRA data) and 1,168 measles cases (KDCA data), along with 2,736 SSPE cases, were diagnosed during the study period. Measles incidence significantly declined, reaching zero cases in 2021, while SSPE incidence displayed an upward trend, peaking in 2014. The mean age at SSPE onset was 21.2 years, with a marked male-to-female ratio of 13.0:1.
Conclusion: SSPE incidence was remarkably low in the post-outbreak period, likely attributable to successful measles control. This study underscores the critical importance of maintaining low measles incidence through sustained vaccination efforts, preventing SSPE and other measles-related complications.

Keywords: Disease outbreaks; Incidence; Measles; Republic of Korea; Subacute sclerosing panencephalitis

Received: March 3, 2025
Revised: April 1, 2025
Accepted: April 28, 2025

Corresponding author:
Jong-Koo Lee
National Academy of Medicine of Korea, 51 Seochojungang-ro, Seocho-gu, Seoul 06654, Republic of Korea
E-mail: docmohw@snu.ac.kr

Young Hwa Lee and Jung Hye Byeon contributed equally to the work as the co-first authors.

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Measles, a highly contagious viral disease, poses a substantial public health threat due to its potential for severe complications. Among these complications, subacute sclerosing panencephalitis (SSPE) is particularly devastating and often fatal. SSPE is caused by persistent measles virus infection within the central nervous system, which can reactivate years after the initial infection, leading to progressive neurological deterioration and ultimately death. The interval between primary measles infection and the onset of SSPE can vary significantly, ranging from several months to multiple decades, underscoring the long-term neurological impact of measles infection [1]. Additionally, previous studies have reported that SSPE predominantly affects young adults [2,3]. This epidemiological characteristic highlights the necessity of maintaining high measles vaccination coverage not only to protect children but also to mitigate long-term consequences in later life.

Although the introduction of effective measles vaccination programs has significantly reduced the incidence of measles and associated complications, including SSPE, the growing trend of vaccine hesitancy and declining herd immunity threaten the gains achieved through vaccination efforts. A resurgence of measles cases could result in a considerable increase in SSPE incidence, emphasizing the critical need for continued vaccination.

The Republic of Korea experienced a notable measles outbreak in 2000–2001 and subsequently implemented a nationwide measles elimination program, providing a unique opportunity to investigate the long-term impact of measles control on SSPE occurrence [4]. This study aimed to determine the incidence of SSPE following the 2000–2001 measles outbreak and evaluate the effectiveness of measles elimination efforts in preventing this devastating complication.

HIGHLIGHTS

- The Republic of Korea's measles elimination resulted in reduced subacute sclerosing panencephalitis (SSPE) incidence, though it peaked after the 2000–2001 measles outbreak. Achieving zero measles cases by 2021 correlated with low SSPE rates. SSPE onset predominantly occurred years after measles exposure, with a marked male predominance. This study emphasizes vaccination's long-term role in preventing SSPE and the importance of sustained measles control.

Materials and Methods

Study Design and Data Source

This nationwide, population-based retrospective cohort study utilized claims data from the National Health Insurance (NHI) system, which is administered by the Health Insurance Review and Assessment Service (HIRA). In the Republic of Korea, all measles cases must be reported to local community health centers, after which they are recorded by the Korea Disease Control and Prevention Agency (KDCA) and the Korean Statistical Information Service (KOSIS). The KOSIS database provided data on measles incidence covering the period from 1974 to 2022, which were included in this study. The NHI system offers universal health coverage to approximately 98% of the Korean population, ensuring that the analyzed sample is highly representative. Claims data from January 2007 through December 2022 were extracted from the HIRA database. This timeframe was selected to reflect the most recent epidemiological trends, thus ensuring an analysis based on current and representative public health data. Data on measles incidence were obtained from both HIRA and KDCA databases. HIRA provides information derived from insurance claims, whereas KDCA data are based on officially reported cases. Discrepancies between these datasets likely result from variations in case definitions and reporting practices. Notably, SSPE cases have only been captured in the HIRA database since 2007; consequently, data on SSPE prior to 2007 were unavailable. Cross-sectional analyses were conducted to evaluate annual temporal trends.

Study Population

This study analyzed national epidemic data for measles among individuals aged 0 to 20 years, as well as SSPE cases, using NHI claims data collected between January 2007 and December 2022. Cases were identified based on inpatient and outpatient claims data that included the relevant International Classification of Diseases, 10th Revision codes: B05 (measles) and A81.1 (SSPE). Patients were included in the study only if these disease codes appeared as the primary diagnosis on their claims. Cases where these codes were listed as accessory diagnoses were excluded. Population data for the Republic of Korea, imported from KOSIS census data provided by Statistics Korea, were utilized to calculate the incidence rates of measles and SSPE relative to the entire population.

Statistical Analysis

Population-based incidence rates for measles and SSPE were calculated per 100,000 individuals. Specifically, SSPE incidence was calculated as the annual number of newly diagnosed SSPE cases per 100,000 individuals in the

https://doi.org/10.24171/j.phrp.2025.0067

population. The chi-square test was used for categorical variables, while the independent t-test was employed for continuous variables. Statistical significance was set at $p < 0.05$. All statistical analyses were conducted using SAS software ver. 9.4 (Enterprise Guide; SAS Institute Inc.).

Ethics Approval

This study was approved by the Institutional Review Board of Korea University Anam Hospital (IRB No. 2023AN0395).

Results

During the observation period, the HIRA database recorded 236 cases of measles and 2,736 cases of SSPE. Surveillance data from the KDCA identified 1,168 measles cases. Reported SSPE cases varied substantially, ranging annually from approximately 93 cases (0.19 per 100,000 individuals) to 333 cases (0.66 per 100,000 individuals). A significant decline in reported measles cases was observed over the study period, from 442 cases (3.87 per 100,000 individuals) in 2014 to zero cases reported by both HIRA and KDCA since 2021 (Table 1). According to the KOSIS database, measles incidence gradually declined from 20.4 per 100,000 individuals in 1976, experienced a sharp increase to 69.4 per 100,000 individuals in 2000, decreased to 48.7 per 100,000 individuals in 2001, and ultimately reached zero in 2021. Conversely, SSPE incidence exhibited an upward trend, rising from 0.24 per 100,000 individuals in 2007 to 0.63 in 2013, and further increasing to 0.66 in 2014. Notably, the peak incidence of SSPE occurred approximately 12 to 14 years after the measles epidemic, with incidence declining thereafter (Figure 1). In Figure 1, the absence of SSPE cases before 2007 is due to the analysis being limited to data available from January 2007 onward, when HIRA's systematic collection of claims data commenced. The mean age at SSPE onset was 21.2 years, ranging from 6 to 30 years. The sex distribution was significantly skewed, exhibiting a male-to-female ratio of 13.0:1. Notably, the age at onset of SSPE increased following significant reductions or cessation of measles transmission (Table 2).

Among measles patients, 57.2% were male and 42.8% were female. In contrast, SSPE patients showed a distinct male predominance, with 92.8% male and 7.2% female. Measles patients had a mean age of 4.8 years, with a median age ranging from 1 to 7 years. Conversely, SSPE patients had a mean age of 21.2 years, with a median age ranging from 19 to 24 years. Healthcare coverage differed substantially between the 2 groups. Most measles patients (96.2%) were enrolled in the NHI system, with only 3.8% benefiting from medical assistance. In contrast, SSPE patients primarily depended on medical assistance, with 72.4% enrolled in this program and only 27.6% covered by NHI.

Table 1. Measles and SSPE cases by year of onset, the Republic of Korea, 2007–2022

Year	KDCA		HIRA		HIRA	
	No. of measles	Incidence rate per 100,000	No. of measles	Incidence rate per 100,000	No. of SSPE	Incidence rate per 100,000
2007	194	1.52	161	1.26	119	0.24
2008	2	0.02	30	0.24	97	0.20
2009	17	0.14	13	0.10	93	0.19
2010	114	0.93	13	0.11	105	0.21
2011	42	0.35	1	0.01	133	0.27
2012	3	0.03	1	0.01	246	0.49
2013	107	0.92	2	0.02	321	0.63
2014	442	3.87	6	0.05	333	0.66
2015	7	0.06	0	0.00	248	0.49
2016	18	0.17	2	0.02	221	0.43
2017	7	0.07	0	0.00	230	0.45
2018	15	0.15	0	0.00	169	0.33
2019	194	1.96	7	0.07	111	0.22
2020	6	0.06	0	0.00	113	0.22
2021	0	0.00	0	0.00	99	0.19
2022	0	0.00	0	0.00	98	0.19
Total cases	1,168		236		2,736	

Data from KDCA and HIRA.

SSPE, subacute sclerosing panencephalitis; KDCA, Korea Disease Control and Prevention Agency; HIRA, Health Insurance Review and Assessment Service.

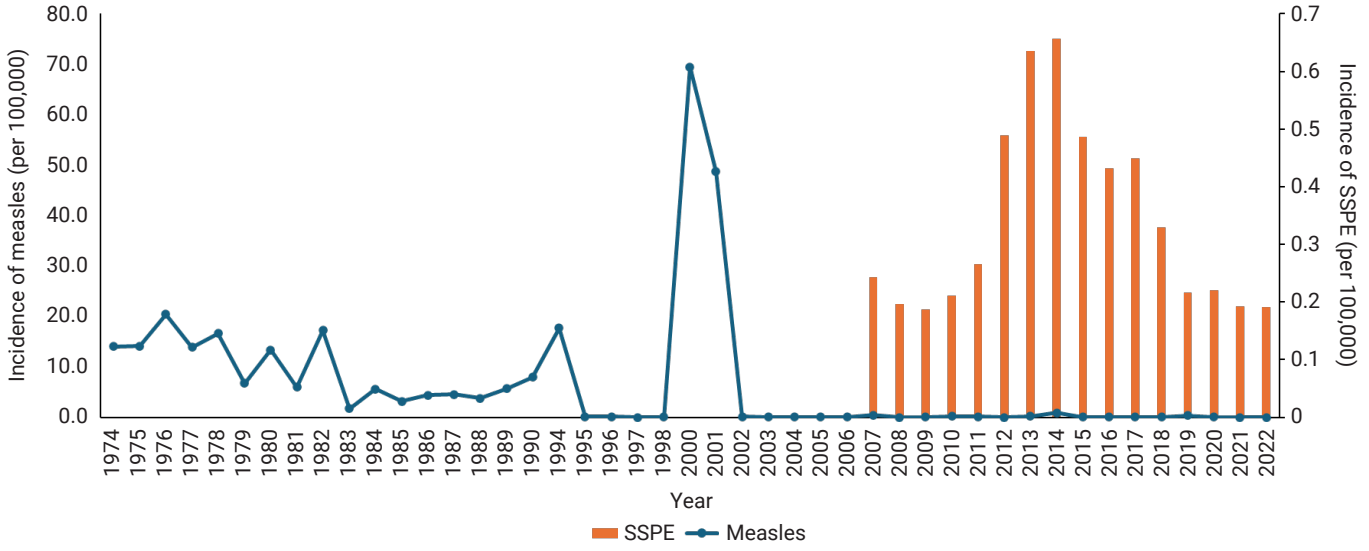


Figure 1. Incidence of subacute sclerosing panencephalitis (SSPE) patients and measles in the Republic of Korea, 1974–2022. Data for the years 1991, 1992, 1993, and 1999 were not available. Data on measles incidence were collected from Statistics Korea, and SSPE data were obtained from the Health Insurance Review and Assessment Service.

Table 2. The annual number of SSPE cases, age at onset, and the sex ratio, reported to the HIRA

Year	No. of SSPE cases reported[a]	Mean age at onset (range, y)[a]	Male:female
2007	119	13.3 (6–15)	4.2:1
2008	97	14.2 (7–16)	4.7:1
2009	93	15.1 (8–17)	5.2:1
2010	105	16.3 (9–18)	10.7:1
2011	133	17.3 (10–19)	7.9:1
2012	246	18.7 (11–20)	16.6:1
2013	321	19.8 (12–21)	17.9:1
2014	333	20.7 (13–22)	24.6:1
2015	248	21.7 (14–23)	19.7:1
2016	221	22.9 (15–24)	19.1:1
2017	230	24.0 (17–25)	12.5:1
2018	169	24.9 (17–26)	15.9:1
2019	111	25.8 (18–27)	21.2:1
2020	113	26.6 (19–28)	7.1:1
2021	99	28.0 (21–29)	32.0:1
2022	98	29.1 (21–30)	97.0:1
Calculated overall median (range)		21.2 (6–30)	13.0:1

SSPE, subacute sclerosing panencephalitis; HIRA, Health Insurance Review and Assessment Service.
[a]The mean age is reported to allow a direct comparison with the existing literature and to summarize the overall data distribution.

The treatment costs associated with SSPE patients were markedly higher than those for measles patients, averaging 328,315 Korean won (Table 3).

Discussion

The most significant finding of this study is the remarkably low incidence of SSPE in the Republic of Korea following the measles outbreak of 2000–2001. This observation underscores the profound impact of effective public health interventions in mitigating the long-term consequences of infectious diseases. The successful implementation of a nationwide measles elimination program, primarily driven by widespread vaccination, drastically curtailed measles transmission [5]. This reduction in measles incidence directly translated to a decreased population at risk of developing

https://doi.org/10.24171/j.phrp.2025.0067

Table 3. Characteristics of measles and SSPE patients from 2007 to 2022, reported to the HIRA

Characteristic	Measles (HIRA)	SSPE
Sex (*n*, %)		
Male	135 (57.2)	2,540 (92.8)
Female	101 (42.8)	196 (7.2)
Age (y)		
Mean ± SD	4.8 ± 4.9	21.2 ± 4.3
Median (IQR)	3 (1–7)	21 (19–24)
Type of insurance		
National health insurance	227 (96.2)	756 (27.6)
Medical aid beneficiary	9 (3.8)	1,980 (72.4)
Medical cost per capita (KRW) (median, IQR)	15,790 (11,780–38,455)	40,410 (8,670–61,720)
Medical cost per capita (KRW) (mean ± SD)	99,380 ± 251,972	328,315 ± 1,689,260
Total medical cost (KRW)	23,453,690	898,269,940

SSPE, subacute sclerosing panencephalitis; HIRA, Health Insurance Review and Assessment Service; SD, standard deviation; IQR, interquartile range; KRW, Korean won.

SSPE, a rare but devastating complication that can manifest years after the initial measles infection.

The temporal association observed between the measles outbreak and the subsequent peak in SSPE cases, followed by a sharp decline, further supports the causal link between measles control and SSPE prevention [4]. The findings of this study resonate with observations from other countries that have achieved high measles vaccination coverage [3]. For instance, in the United Kingdom, robust vaccination programs led to a dramatic reduction in SSPE incidence, from 5 to 10 cases per year in the 1990s to near elimination in recent years [6]. This remarkable progress highlights the global impact of measles vaccination in preventing not only acute measles infections but also protecting against the long-term neurological consequences of the virus. The virtual disappearance of SSPE in countries with sustained low measles incidence serves as strong evidence of the enduring benefits of vaccination, underscoring the importance of maintaining high immunization rates [7]. It is noteworthy that SSPE typically develops 7 to 10 years after measles infection, suggesting that SSPE incidence observed in a given year likely corresponds to measles cases from several years earlier.

In this study, we reported the incidence of SSPE per 100,000 individuals in the general population rather than per 100,000 measles cases. This approach reflects our aim to assess the population-level impact of measles and SSPE in the context of a nationwide measles elimination program. Reporting SSPE incidence at the population level provides a clearer understanding of the disease burden within the broader public health context and allows for a direct evaluation of measles control measures' effectiveness in preventing this rare but severe complication. Our findings emphasize the importance of maintaining low measles incidence to prevent SSPE. The study revealed no concerning trends or indications of increased SSPE risk in the post-outbreak period, emphasizing the effectiveness of sustained measles control efforts. This observation aligns with findings from other countries, such as Finland and the United Kingdom, where sustained low measles incidence has led to an almost complete disappearance of SSPE cases [6,8]. The temporal relationship between the 2000–2001 measles outbreak and the SSPE incidence peak around 2014 further emphasizes the delayed yet significant impact of measles infection on SSPE development [4]. This lag underscores the necessity of long-term surveillance even after successful measles control. The continued decline in SSPE cases following the peak provides reassurance about the sustained effectiveness of vaccination programs.

Our study supports previous research demonstrating a pronounced male predominance among SSPE patients. For instance, a German study covering the period from 2003 to 2009 reported similar sex-related disparities, suggesting males may be more susceptible to SSPE or its clinical manifestations [9]. Although the underlying mechanisms remain unclear, factors such as hormonal differences, genetic predispositions, or variations in immune responses might contribute to this phenomenon [10]. In our study, the male-to-female ratio was approximately 13:1, further highlighting this significant gender difference. Additionally, the observed pattern of SSPE incidence peaking after a measles outbreak and subsequently declining aligns with trends observed in other countries [11], reinforcing the strong association between effective measles control and SSPE prevention [12].

This study has some limitations. First, as a retrospective analysis utilizing national health insurance claims data,

it is inherently subject to potential coding errors and underreporting of SSPE cases. Second, the study did not assess measles vaccination status or the severity of primary measles infection, both of which could influence SSPE risk. Third, because the study relied solely on reported and claims data, it does not provide insights into the clinical long-term risk of SSPE development; moreover, since our data cover only from 2007 to 2022, they may not fully represent broader epidemiological trends observed since 1974. Although our study identifies a significant male predominance among SSPE cases, we could not compare this finding with the sex distribution of measles cases during the 2000–2001 outbreak due to data limitations. Consequently, further research is needed to determine whether this disparity reflects inherent sex-based susceptibility to SSPE or merely reflects demographic characteristics of the initial measles outbreak.

Despite these limitations, the study's strengths include its large nationwide sample size and comprehensive coverage through the national health insurance database, minimizing selection bias and providing a representative depiction of SSPE incidence in the Republic of Korea. Measles immunization primarily targets reducing measles cases, and the associated decline in SSPE incidence represents an important, though indirect, benefit of effective measles control.

While KDCA surveillance data offer valuable insights into overall measles trends, our analysis primarily relied on HIRA and KOSIS data. Differences in data collection methods and case ascertainment among these sources limit direct comparisons. The inclusion of KDCA data provides broader historical context regarding the measles outbreak, facilitating interpretation of SSPE incidence trends in light of overall measles control efforts. The findings hold significant implications for global public health, particularly for countries aiming to eliminate measles and prevent SSPE. The Republic of Korea's success in controlling measles and subsequently reducing SSPE incidence serves as a compelling example of the long-term benefits derived from effective vaccination programs and public health interventions. By highlighting the effectiveness of measles control efforts in preventing SSPE, this study contributes valuable evidence supporting global measles elimination initiatives and vaccination campaigns.

In conclusion, our study demonstrates a remarkably low incidence of SSPE in the Republic of Korea following a significant measles outbreak, attributable primarily to a successful nationwide measles elimination program. Although reduced SSPE incidence is an encouraging outcome, it remains an indirect benefit of measles control efforts, which primarily aim to reduce measles cases. By highlighting this indirect impact, our findings provide valuable support for global initiatives promoting measles elimination and robust vaccination campaigns.

Notes

Ethics Approval

This study was approved by the Institutional Review Board of Korea University Anam Hospital (IRB No. 2023AN0395).

Conflicts of Interest

Jong-Koo Lee has been the editor-in-chief of *Osong Public Health and Research Perspectives* since October 2021. Young June Choe was not involved in the editorial process or review of this manuscript, despite being a member of the Editorial Board of the *Osong Public Health and Research Perspectives*. The other authors have no conflicts of interest to declare.

Funding

This study was supported by Korea University (Research No. K2327151).

Availability of Data

All data generated or analyzed during this study are included in this published article. For other data, these may be requested through the corresponding author.

Authors' Contributions

Conceptualization: YHL, JHB, YJC; Data curation: YHL, JHB; Formal analysis: YHL, CRK, JKL, YJC; Investigation: YHL, YJC; Methodology: YHL, YJC; Project administration: CRK, YJC; Resources: YJC; Supervision: JKL; Validation: YJC; Visualization: YJC; Writing–original draft: YHL, JHB, YJC; Writing–review & editing: all authors. All authors read and approved the final manuscript.

Acknowledgements

We gratefully acknowledge the dedication of public health workers across the Republic of Korea, whose commitment to measles prevention and control has significantly contributed to the low incidence of subacute sclerosing panencephalitis observed in this study.

References

1. Krumova S, Andonova I, Stefanova R, et al. Measles virus and subacute sclerosing panencephalitis. Clin Lab 2022;68.
2. Khetsuriani N, Sanadze K, Abuladze M, et al. High risk of subacute sclerosing panencephalitis following measles outbreaks in Georgia. Clin Microbiol Infect 2020;26:737–42.
3. Zilber N, Rannon L, Alter M, et al. Measles, measles vaccination, and risk of subacute sclerosing panencephalitis (SSPE). Neurology 1983;33:1558–64.
4. Choe YJ, Jee Y, Oh MD, et al. Measles elimination activities in the Western Pacific Region: experience from the Republic of Korea. J Korean Med Sci 2015;30(Suppl 2):S115–21.
5. Centers for Disease Control and Prevention (CDC). Elimination of measles: South Korea, 2001-2006. MMWR Morb Mortal Wkly Rep 2007;56:304–7.
6. Miller C, Andrews N, Rush M, et al. The epidemiology of subacute

https://doi.org/10.24171/j.phrp.2025.0067

sclerosing panencephalitis in England and Wales 1990-2002. Arch Dis Child 2004;89:1145–8.
7. Gadoth N. Subacute sclerosing panencephalitis (SSPE) the story of a vanishing disease. Brain Dev 2012;34:705–11.
8. Davidkin I, Kontio M, Paunio M, et al. MMR vaccination and disease elimination: the Finnish experience. Expert Rev Vaccines 2010;9:1045–53.
9. Schonberger K, Ludwig MS, Wildner M, et al. Epidemiology of subacute sclerosing panencephalitis (SSPE) in Germany from 2003 to 2009: a risk estimation. PLoS One 2013;8:e68909.
10. Aaby P, Bukh J, Lisse IM, et al. Risk factors in subacute sclerosing panencephalitis: age- and sex-dependent host reactions or intensive exposure? Rev Infect Dis 1984;6:239–50.
11. Okabe H, Hashimoto K, Norito S, et al. Patients with subacute sclerosing panencephalitis in Japan: a 2022 nationwide survey. Pediatr Infect Dis J 2024;43:313–9.
12. Okuno Y, Nakao T, Ishida N, et al. Incidence of subacute sclerosing panencephalitis following measles and measles vaccination in Japan. Int J Epidemiol 1989;18:684–9.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2024.0343
Osong Public Health Res Perspect 2025;16(3):292-299

Short Communication

First report of human *Mycobacterium bovis* infection in a veterinary laboratory worker in the Republic of Korea

Ji-Yeon Lee[1,#], Seong Wook Pyo[2], Jieun Kim[1], Young-Joon Park[1]

[1]Division of Tuberculosis Policy, Department of Infectious Disease Policy, Korea Disease Control and Prevention Agency, Cheongju, Republic of Korea
[2]Division of Bacterial Diseases, Department of Laboratory Diagnosis and analysis, Korea Disease Control and Prevention Agency, Cheongju, Republic of Korea

ABSTRACT

Objectives: In the Republic of Korea, the previous surveillance system for zoonotic tuberculosis (TB) involved the X-ray testing of humans in contact with *Mycobacterium bovis*-infected livestock. In contrast, the updated surveillance system incorporates the genotyping of cultured *Mycobacterium* isolates for high-risk occupational groups. This study aimed to systematically document the detection, diagnosis, assessment, and response in the epidemic investigation of zoonotic TB in a laboratory worker in the Republic of Korea.

Methods: *M. bovis* was confirmed using spoligotyping and whole genome sequencing. Clinical characteristics were reviewed through epidemiological investigation and interviews with the affected individual. Transmission routes and secondary spread were assessed via field epidemiological investigations and contact evaluations using chest X-ray and interferon gamma release assay for latent TB infection.

Results: A 56-year-old laboratory worker presented with chest X-ray findings compatible with TB and subsequently tested positive for *M. bovis*. She had no clinical or family history of TB and remained asymptomatic. She completed a 6 month treatment regimen of isoniazid, rifampin, ethambutol, and pyrazinamide without hospitalization. Although no direct transmission pathways for zoonotic TB were identified, her work in a laboratory, processing specimens for zoonotic TB, indicated potential laboratory related exposure.

Conclusion: This case underscores the importance of stringent use of personal protective equipment among high-risk occupational groups and the implementation of an enhanced surveillance system to report zoonotic TB. These findings highlight the need for a One Health approach and proactive surveillance, emphasizing the necessity of refining and strengthening surveillance systems for precise monitoring and an effective response.

Keywords: High-risk group; Laboratory worker; *Mycobacterium bovis*; Occupational exposure; One Health; Zoonoses

Received: December 24, 2024
Revised: March 4, 2025
Accepted: March 9, 2025

Corresponding author:
Young-Joon Park
Division of Tuberculosis Policy, Department of Infectious Disease Policy, Korea Disease Control and Prevention Agency, 187 Osongsaengmyeong 2-ro, Osong-eup, Heungdeok-gu, Cheongju 28159, Republic of Korea
E-mail: pahmun@korea.kr

[#]Current affiliation: Division of Emerging Infectious Disease Response, Department of Infectious Disease Emergency Preparedness and Response, Cheongju, Republic of Korea

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

In January 2024, zoonotic tuberculosis (TB) was confirmed in a laboratory worker, representing the first detection under the Republic of Korea's updated zoonotic TB surveillance system. The previous surveillance system focused on TB screening of humans who had been in contact with *Mycobacterium bovis*-infected cattle, such as livestock farm owners. In contrast, the enhanced surveillance system extended testing to high-risk occupational groups to detect *M. bovis* using spoligotyping and whole genome sequencing (WGS) of culture-positive *Mycobacterium tuberculosis* complex (MTBC) strains.

Zoonotic TB refers to TB transmitted from animals to humans, primarily through *M. bovis* [1]. The main hosts of *M. bovis* are cattle [2]. However, other animals, including livestock (e.g., deer and goats), pets (dogs and cats), and wildlife (wild boars and water deer), can also serve as reservoirs [3], complicating control and management efforts due to the complexity of transmission pathways.

The definitive diagnosis of zoonotic TB requires the isolation and identification of TB strains from human sputum cultures. However, differentiating between *M. tuberculosis* and *M. bovis* is challenging because of their similar clinical and radiological characteristics [4]. Both species belong to the MTBC and have genomes that are 99.9% identical at the nucleotide level [5]; thus, their differentiation using conventional laboratory tests is difficult. To confirm the presence of *M. bovis*, the Korea Disease Control and Prevention Agency (KDCA) relies on identifying its characteristic pyrazinamide resistance and employs spoligotyping and WGS for definitive analysis.

M. bovis is commonly transmitted through the ingestion of contaminated food, direct contact, or airborne exposure. Risk factors include consuming products from infected animals that are not fully cooked (e.g., dairy products such as cheese and milk) and percutaneous exposure via skin injuries during necropsy, slaughter, or laboratory work performed without appropriate personal protective equipment (PPE).

According to the European Union One Health 2023 Zoonoses Report published by the European Food Safety Authority and the European Center for Disease Prevention and Control, 138 confirmed human cases of zoonotic TB caused by *M. bovis* or *Mycobacterium caprae* were reported in 2023, corresponding to an incidence rate of 0.04 cases per 100,000 in the European Union and representing a 6.1% decrease in the notification rate from 2022 [6]. The Global Tuberculosis Report 2020, released by the World Health Organization, indicated that approximately 140,000 cases (1.4%) of zoonotic TB were confirmed among new TB cases in 2019, with about 11,400 (8.1%) deaths [7]. However, due to the lack of a zoonotic TB surveillance system in the Republic of Korea, the exact incidence of zoonotic TB in the country remains unknown [3].

HIGHLIGHTS

- We report the first case of zoonotic tuberculosis (TB) in a laboratory worker following an update to the Republic of Korea's zoonotic TB surveillance system.
- Under the previous system, TB testing was conducted on humans in contact with *Mycobacterium bovis*-infected livestock. With the improved system, testing was extended to high-risk occupational groups using genetic analysis to identify *M. bovis* in culture-positive cases.
- Epidemiological investigation suggested that this infection likely resulted from occupational exposure.
- Effective prevention, early detection, and appropriate response to zoonotic TB in high-risk groups require collaboration among relevant agencies and the implementation of a One Health approach.

To address this issue, the Division of Tuberculosis Policy of the KDCA has strengthened the surveillance framework for high-risk groups as part of its national TB control strategy. In collaboration with the Animal and Plant Quarantine Agency and the Ministry of Environment, the KDCA developed the Joint Epidemiological Investigation Manual for Zoonotic Tuberculosis in May 2023. This resource expands the genetic testing of TB strains to include workers in livestock and zoo settings, aiming to quantify zoonotic TB cases, improve contact management, and analyze transmission pathways.

Although zoonotic TB is relatively rare and may not appear to represent a major public health concern, the expansion of the livestock industry and increasing human-animal interactions heighten the risk of transmission [8]. Therefore, analyzing the surveillance system, transmission routes, and clinical characteristics of this first confirmed case under the enhanced system may provide key insights for future prevention and management.

Materials and Methods

Study Design and Population

In this retrospective study, we analyzed data from a 56-year-old female laboratory worker with confirmed zoonotic TB in January 2024, including the infection route and contacts. Fifteen individuals were identified as contacts, comprising 8 close contacts and 7 casual contacts.

This study systematically outlines the detection process and transmission route of zoonotic TB for laboratory workers through the Republic of Korea's national notifiable infectious disease surveillance system.

Case Definition

According to the National Tuberculosis Management Guidelines, zoonotic TB is defined as the identification of *M. bovis* through genetic analysis of a culture-positive TB strain [3]. A contact refers to anyone who shared living spaces or had regular, prolonged, and direct contact with a patient with TB in an indoor environment within 3 months prior to TB diagnosis [1].

Data Collection

Following laboratory confirmation of zoonotic TB in the patient on January 10, 2024, an in-depth investigation was conducted in accordance with the Joint Response Manual to trace the infection source and transmission route. In cooperation with the KDCA Division of Tuberculosis Policy, local health authorities coordinated with the institution where the patient worked, conducting site visits, interviews with the patient, data verification, and comprehensive investigations to hypothesize potential transmission pathways and shared exposure sources. In line with the National Tuberculosis Management Guidelines, the 15 identified contacts underwent chest X rays, sputum tests, and interferon gamma release assay (IGRA) testing.

Spoligotyping and WGS

Spoligotyping was performed using a spoligotyping kit (Ocimum Biosolutions) in accordance with the manufacturer's protocol, and the results were analyzed using the SITVIT2 and TB-Insight databases. WGS was performed using the Illumina Miseq system (Illumina Inc.) from a purified library prepared with the Illumina DNA Prep kit (Illumina Inc.). Subsequently, core genome multilocus sequence typing (cgMLST) based on 2,621 loci was analyzed using Ridom SeqSphere+ ver. 8.5.0 (Ridom GmbH).

Ethics Statement

This study was approved by the Institutional Review Board of the KDCA (No. KDCA-2024-10-02-P-01) and was performed in accordance with the principles of the Declaration of Helsinki. It involved an epidemiological investigation conducted in compliance with the Tuberculosis Prevention Act and the Korean National Tuberculosis Management Guidelines for TB control. Written informed consent was obtained from the patient.

Results

Effectiveness of the Enhanced Zoonotic TB Surveillance System

The previous surveillance system focused on conducting TB tests (X-ray and sputum culture when necessary) among individuals in contact with *M. bovis*-infected livestock, such as farmworkers. The enhanced system was established to proactively address zoonotic TB and to more effectively manage high-risk occupational groups. In May 2023, a Joint Response Manual was developed in collaboration with relevant agencies, and in June 2023, epidemiological survey forms were updated to capture the occupations of patients with TB, such as those working in livestock-related fields [9,10]. Additionally, protocols were updated to mandate that culture-positive strains be referred to the Bacterial Analysis Division via the Comprehensive Disease and Health Management System. In 2016, genetic testing of TB strains covered only patients with TB in institutional settings; this was gradually expanded to include children under 5 years of age in 2019. The 2023 revisions further expanded testing to include patients with multidrug resistant TB and groups at high risk of zoonotic TB (Table 1). Furthermore, upon detecting TB in livestock, human contacts such as farmworkers are now screened using chest X-rays, sputum tests, and, if necessary, IGRA for latent TB infection (LTBI).

Detection, Assessment, Confirmation, and Reporting

The patient in this case had been undergoing regular follow-up visits at the university hospital for rheumatic disease,

Table 1. Expansion of target for genotyping for TB detection

	Previous	Expanded (from 2023)
Target for genetic TB testing	· Patients with TB in institutional settings · Patients with TB < 5 years old	· Patients with TB in institutional settings · Patients with TB < 5 years old + · Patients with multidrug-resistant TB · High-risk groups for zoonotic tuberculosis (such as livestock workers)

TB, tuberculosis.

https://doi.org/10.24171/j.phrp.2024.0343

and on January 31, 2023, a chest X-ray revealed findings suggestive of TB (no cavity observed). Subsequent sputum smear tests performed using self-collected samples on February 1, 2023, returned negative results. Bronchoscopy was performed on February 22, 2023. Although an acid fast bacillus (AFB) smear and TB polymerase chain reaction (a nucleic acid amplification test for TB) were negative, her AFB culture was positive. Consequently, she was diagnosed with pulmonary TB on March 24, 2023.

According to the National Tuberculosis Management Guidelines, rapid drug susceptibility testing for isoniazid and rifampin, along with conventional drug susceptibility testing, is performed on the first culture for all patients with TB [1]. Rapid drug susceptibility testing indicated no resistance to isoniazid or rifampin. While conventional drug susceptibility results were inconclusive, WGS confirmed pyrazinamide resistance.

Spoligotyping results confirmed the BOV-3 clade and the *M. bovis* lineage; additional analysis using WGS revealed that the Rd4 region was deleted and identified a sequence mutation (C→G) at base 169 of the *pncA* gene, a characteristic of *M. bovis* associated with intrinsic pyrazinamide resistance. In the cgMLST analysis, the strain isolated from the patient (E2306) was closely related to strains isolated from animals in the Republic of Korea (Figure 1).

The patient remained asymptomatic and exhibited no specific TB related symptoms apart from those associated with her preexisting rheumatic disease, for which she was not on immunosuppressive therapy. She began a 6 month regimen of isoniazid, rifampin, ethambutol, and pyrazinamide on March 24, 2023, and completed treatment without hospitalization. From March 24 to April 7, 2023, the patient was restricted from working to prevent transmission. On April 10, 2023, a physician confirmed that she was no longer infectious, allowing her to resume work and her regular activities.

Upon reporting the case on March 23, 2023, the local health

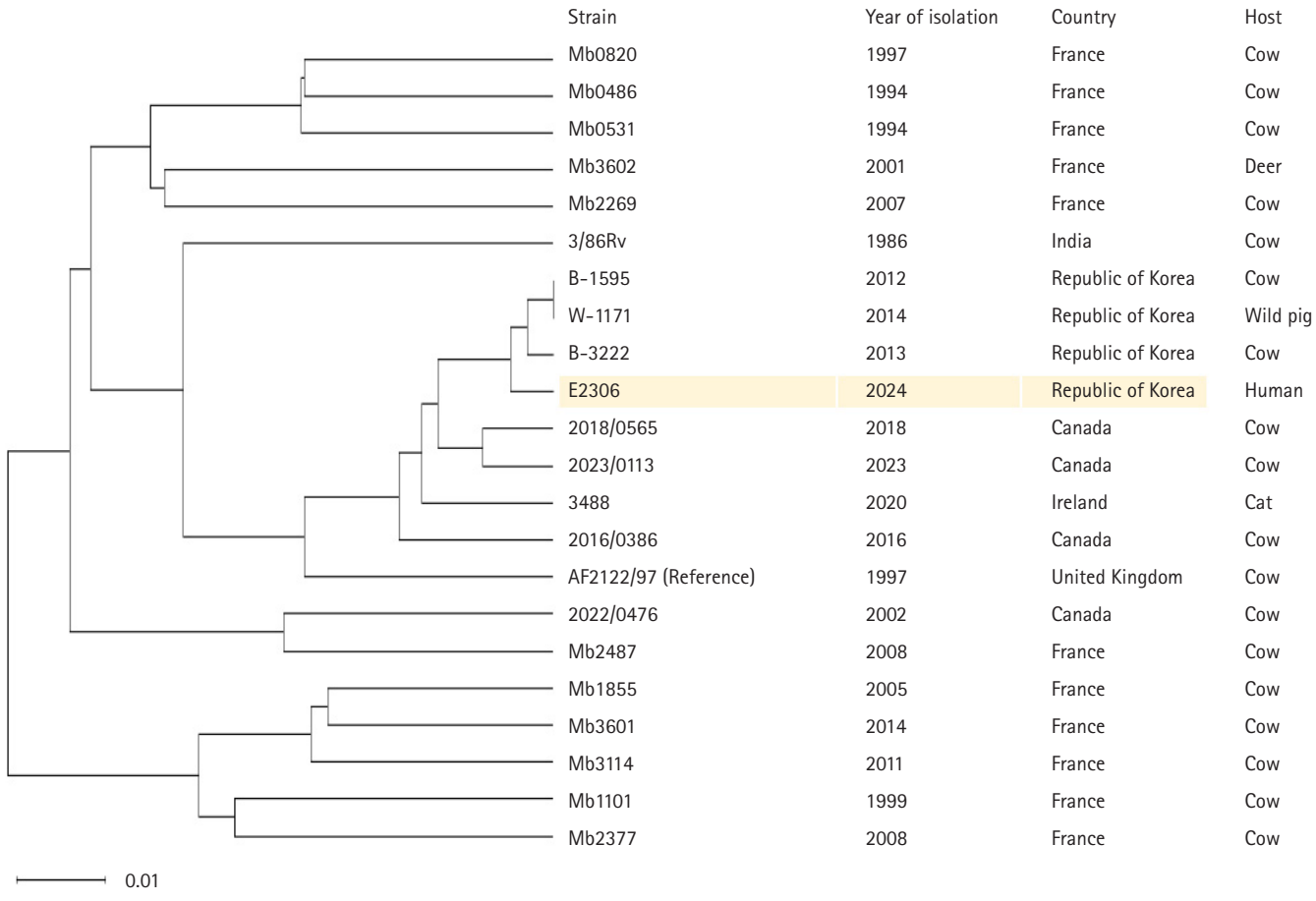


Figure 1. Phylogenetic analysis of *Mycobacterium bovis* strains. The core genome multilocus sequence typing of *M. bovis* was analyzed using Ridom SeqSphere+ ver. 8.5.0 (Ridom GmbH). A phylogenetic tree was constructed using the *M. bovis* whole genome sequences registered in the National Center for Biotechnology Information database with the unweighted pair-group method with arithmetic mean method based on 2,612 loci for the *Mycobacterium tuberculosis* complex.

authority identified the patient as belonging to a group at high risk of zoonotic TB. Subsequently, in June 2023, her TB-positive culture was submitted for genetic analysis—including spoligotyping and WGS—via the Comprehensive Disease and Health Management System to the KDCA Division of Bacterial Disease. On January 10, 2024, genetic testing revealed *M. bovis* infection, thereby confirming zoonotic TB (Figure 2).

Response and Investigation

The contact investigation period covered the 4 weeks (28 days) preceding the earliest diagnostic test (either sputum smear or chest X ray), spanning a total of 11 weeks from January 3, 2023, to March 24, 2023. One household contact was asymptomatic and tested negative for TB (via chest X ray) and LTBI (via IGRA).

In the workplace, 15 of the patient's colleagues were identified as contacts. Among these, 8 who worked closely with the index patient in the same laboratory were classified as close contacts, while the remaining 7—who interacted casually during lunch breaks or incidental encounters within the same building—were categorized as casual contacts.

Close contacts underwent both TB and LTBI testing, along with follow up TB testing, whereas casual contacts were tested only for TB. Initial TB testing was conducted on April 15, 2023, and follow up testing on August 2, 2023. All 15 workplace contacts tested negative for TB. Among the 8 close contacts, 1 had a history of TB treatment and was not subjected to LTBI testing; of the remaining 7 individuals, 2 received a positive IGRA result, yielding a positivity rate of 28.6%. Both of these contacts completed LTBI treatment between May and August 2023 (Table 2).

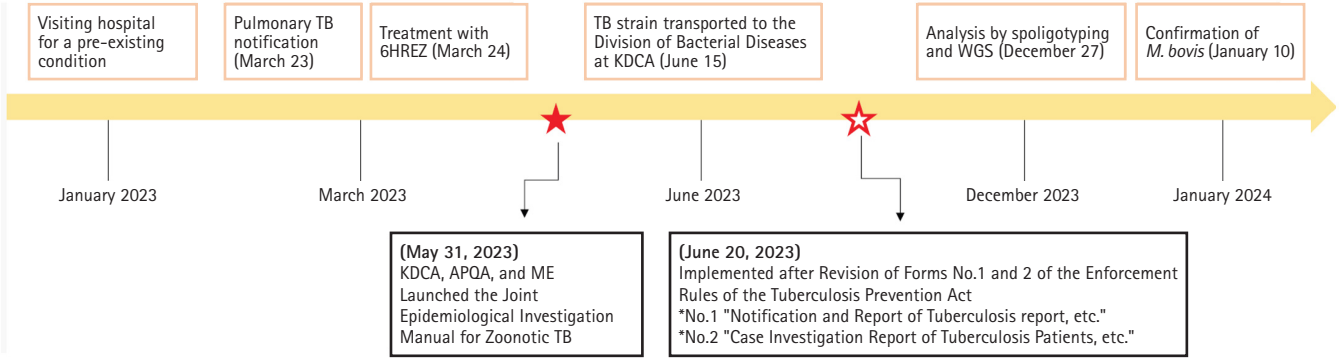


Figure 2. Timeline of the detection of zoonotic tuberculosis.
TB, tuberculosis; KDCA, Korea Disease Control and Prevention Agency; WGS, whole genome sequencing; *M. bovis*, *Mycobacterium bovis*; APQA, Animal and Plant Quarantine Agency; ME, Ministry of Environment.

Table 2. Characteristics and test results of exposed contacts

Characteristic	Total	Chest X-ray result		IGRA test result	
		Positive	Negative	Positive	Negative
Total	15 (100.0)	0 (0.0)	15 (100.0)	2 (100.0)	5 (100.0)
Sex					
Male	7 (46.7)	0 (0.0)	7 (46.7)	1 (50.0)	3 (60.0)
Female	8 (53.3)	0 (0.0)	8 (53.3)	1 (50.0)	2 (40.0)
Age group (y)	42.9±10.5		42.9±10.5	49.0±5.0	35.8±10.5
20–29	3 (20.0)	0 (0.0)	3 (20.0)	0 (0.0)	2 (40.0)
30–39	3 (20.0)	0 (0.0)	3 (20.0)	0 (0.0)	2 (40.0)
40–49	4 (26.7)	0 (0.0)	4 (26.7)	1 (50.0)	0 (0.0)
50–59	5 (33.3)	0 (0.0)	5 (33.3)	1 (50.0)	1 (20.0)
Contacts					
Close	8 (53.3)	0 (0.0)	8 (53.3)	2 (100.0)	5 (100.0)
Casual	7 (46.7)	0 (0.0)	7 (46.7)	N/A	N/A

Data are presented as *n* (%) or mean±standard deviation.
IGRA, interferon gamma release assay; N/A, not applicable.

https://doi.org/10.24171/j.phrp.2024.0343

Transmission Pathway

The patient had no personal or family history of TB. Genetic analysis of the TB strain using spoligotyping and WGS confirmed the presence of the *M. bovis* genotype, suggesting a negligible likelihood of association with Bacillus Calmette–Guérin vaccination. Moreover, the patient reported that she dislikes milk and had no history of consuming unpasteurized milk, raw meat, or animal by products. She also had no history of handling animals, visiting farms with open skin wounds, or direct contact with pets, wild animals, or *M. bovis* infected livestock. Consequently, her likelihood of exposure through previous zoonotic TB cases, direct contact with infected livestock, or exposure to livestock by-products was very low.

The investigation additionally indicated a high likelihood of exposure through laboratory work. The patient had worked in a laboratory setting for approximately 20 years, since 2002, and was involved in specimen processing. She primarily assisted with IGRA and serological testing (e.g., separating blood samples drawn with syringes) and performed histopathological examinations (processing lung nodules).

Although the patient consistently wore disposable gloves and gowns to prevent accidents involving needles or scalpels and to minimize exposure to animal fluids, she reported having experienced needle-stick incidents in the past. Individuals engaged in laboratory work face a risk of exposure to *M. bovis*-infected body fluids from contaminated needles, accidental splashes of biological material into the eyes, or unnoticed skin abrasions, all of which could serve as entry points for the pathogen.

Given the variable latency period of TB infection, which depends on an individual's immune status, the exact timing of infection remains unclear. However, the patient likely acquired the infection through direct or percutaneous exposure to contaminated biological materials in a confined laboratory environment, with the infection remaining latent until immunosuppression triggered its progression to active TB.

Discussion

Between 1993 and 2004, studies in the United Kingdom reported the incidence of zoonotic TB among farmers and farm workers exposed to *M. bovis*-infected cattle; however, it remains unclear whether LTBI was diagnosed or whether LTBI testing was performed in those cases [11]. In contrast, an outbreak of *M. bovis* among zoo animals in the Republic of Korea from July 2021 to September 2022 presented a different scenario. Of the 21 exposed zoo workers tested, 6 (20.7%) were positive for LTBI on the initial IGRA screening, and an additional case (raising the positivity rate to 24.1%) was detected during follow-up [12]. Although the prolonged latency of TB complicates the determination of the exact timing of infection, these findings underscore the need for vigilant monitoring of occupational groups at high risk of exposure to zoonotic TB.

Based on the investigation of the infection source in this study, our patient likely acquired the infection through occupational exposure. Further studies are needed to investigate the risks associated with co exposure to livestock and their by products, in order to better define transmission routes and assess the risk of infection in individuals exposed to *M. bovis* through multiple pathways. In addition, individuals at high risk of exposure to zoonotic TB—such as laboratory workers and farm workers—should be educated about the importance of wearing personal protective equipment, including gowns and mask, gloves etc., to support the prevention, detection, and response to zoonotic infections.

Internationally, sporadic cases of zoonotic TB associated with *M. bovis* have been reported among individuals with occupational exposure. Examples include a case from New Zealand in 2015 involving a woman in her 50s who worked in a facility processing animal organs (mainly beef) for 7 years [13], a case from the United Kingdom in 2009 involving a 42-year-old veterinary nurse who conducted TB testing on cattle and handled an injured badger [14], and a case involving a 25-year-old veterinarian in the United Kingdom in 2011 who contracted *M. bovis* while performing invasive procedures (e.g., testing, intravenous injection, and autopsy) on infected alpacas without proper PPE [15]. These cases illustrate the varied contexts in which occupational exposure can lead to zoonotic TB transmission.

This study encountered several limitations. The slow growth and challenging culture conditions of TB bacteria, combined with the difficulty of differentiating *M. bovis* infection from human *M. tuberculosis* using standard laboratory tests, posed significant challenges. First, the substantial delay between the initial TB diagnosis and genotyping hindered prompt epidemiological investigation and response. Second, this delay resulted in the confirmation of zoonotic TB after the patient's treatment had concluded. This case emphasizes the need for faster diagnostic processes, as early identification of *M. bovis* is crucial for managing drug-resistant TB. If the causative pathogen is not promptly identified as *M. bovis* and treatment is initiated with pyrazinamide, there is a risk of treatment failure and progression to drug-resistant TB. This was evidenced by a case of zoonotic TB in Peru in 2009, in which initial symptom improvement was followed by relapse during second-line treatment, ultimately resulting in

respiratory failure and death in 2012 [16]. Additionally, although this study systematically reviewed the epidemiological investigation process, it only covers a single case and limits the generalizability of the findings. Further research, including large-scale cohort studies, is needed to identify various risk factors related to zoonotic TB transmission in high-risk occupational settings.

Nevertheless, this study provides meaningful implications. First, it is the initial response to zoonotic TB using the Joint Epidemiological Investigation Manual launched in May 2023, and the updated legal investigation form announced in June 2023. Second, it systematically organizes the process from detection, assessment, confirmation, response and investigation of the first zoonotic tuberculosis case identified through the enhanced national notifiable infectious diseases surveillance system. Third, it emphasizes the need for an enhanced surveillance system and the importance of further strengthening the surveillance system.

Conclusion

From a public health perspective, this study underscores the need for sophisticated surveillance and appropriate interventions to control zoonotic TB. It also requires education of people who are occupationally exposed to zoonotic diseases about the importance of wearing appropriate PPE. Integrating these strategies can reduce the risk of zoonotic diseases and improve the preemptive response to emerging threats.

Notes

Ethics Approval

This study was approved by the Institutional Review Board of the KDCA (No. KDCA-2024-10-02-P-01) and was performed in accordance with the principles of the Declaration of Helsinki. The epidemiological investigation was conducted in compliance with the Tuberculosis Prevention Act and the Korean National Tuberculosis Management Guidelines for tuberculosis control. Written informed consent was obtained from the patient.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Funding

None.

Availability of Data

All data generated or analyzed during this study are included in the published article. Additional data may be requested from the corresponding author.

Authors' Contributions

Conceptualization: JYL, YJP; Data curation: JYL, JK, SWP; Formal analysis: JYL, SWP; Investigation: JYL, JK; Methodology: JYL, YJP, SWP; Project administration: JL, JK; Resources: JYL, SWP, JK, YJP; Software: JYL; Supervision: JK, YJP; Validation: JK, YJP; Visualization: JYL, SWP; Writing–original draft: JYL; Writing–review & editing: all authors. All authors read and approved the final manuscript.

Additional Contributions

We extend our gratitude to the laboratory staff and officials at the local public health clinics for their valuable cooperation during the investigation.

References

1. Korea Disease Control and Prevention Agency (KDCA) (KR). 2024 National tuberculosis management guidelines [Internet]. KDCA; 2024 [cited 2024 Jul 22]. Available from: https://url.kr/xmxrla. Korean.
2. Centers for Disease Control and Prevention (CDC) (US). About bovine tuberculosis in humans [Internet]. CDC; 2024 [cited 2024 May 20]. Available from: https://www.cdc.gov/tb/about/m-bovis.html.
3. Korea Disease Control and Prevention Agency (KDCA), Animal and Plant Quarantine Agency (APQA), Ministry of Environment (ME) (KR). Zoonotic disease (3. Zoonotic tuberculosis) Joint Epidemiological Investigation manual [Internet]. KDCA; 2023 [cited 2024 May 31]. Available from: https://url.kr/lqdqvu. Korean.
4. Grange JM. Mycobacterium bovis infection in human beings. Tuberculosis (Edinb) 2001;81:71–7.
5. Garnier T, Eiglmeier K, Camus JC, et al. The complete genome sequence of Mycobacterium bovis. Proc Natl Acad Sci U S A 2003;100:7877–82.
6. European Food Safety Authority (EFSA); European Centre for Disease Prevention and Control (ECDC). The European Union One Health 2023 Zoonoses report. EFSA J 2024;22:e9106.
7. World Health Organization (WHO) (CH). Global tuberculosis report 2020 [Internet]. WHO; 2019 [cited 2024 Jul 22]. Available from: https://www.who.int/publications/i/item/9789240013131.
8. Vayr F, Martin-Blondel G, Savall F, et al. Occupational exposure to human Mycobacterium bovis infection: a systematic review. PLoS Negl Trop Dis 2018;12:e0006208.
9. Korean Law Information Center (KLIC) (KR). Enforcement Rules of the Tuberculosis Prevention Act, Forms No. 1 "Notification and Report of Tuberculosis report, etc." [Internet]. KLIC; 2024 [cited 2024 Oct 7]. Available from: https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EA%B2%B0%ED%95%B5%EC%98%88%EB%B0%A9%EB%B2%95%20%EC%8B%9C%ED%96%89%EA%B7%9C%EC%B9%99/(00979,20231201). Korean.
10. Korean Law Information Center (KLIC) (KR). Enforcement Rules of the Tuberculosis Prevention Act, Forms No. 2 "Case Investigation Report of Tuberculosis Patients, etc." [Internet]. KLIC; 2024 [cited 2024 Sep 10]. Available from: https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EA%B2%B0%ED%95%B5%EC%98%88%EB%B0%A9%EB%B2%95%20%EC%8B%9C%ED%96%89%EA%B7%9C%EC%B9%99/(00979,20231201). Korean.
11. UK Health Security Agency (UKHSA) (GB). Prevalence of latent tuberculosis (TB) infection in occupational groups who have prolonged indoor contact with cows: a rapid review [Internet]. UKHSA; 2023

https://doi.org/10.24171/j.phrp.2024.0343

[cited 2024 Sep 10]. Available from: https://assets.publishing.service.gov.uk/media/650acd48fbd7bc0013cb51dd/prevalence-of-TB-in-occupational-groups-indoor-contact-with-cows-rapid-review.pdf.

12. Lee HY, Kwon Y, Lee SE, et al. A Mycobacterium bovis outbreak among exhibition animals at a zoo in the Republic of Korea: the first contact investigation of zoonotic tuberculosis. Osong Public Health Res Perspect 2024;15:248–59.
13. Chan HH, Mpe J. A rare cause of pulmonary tuberculosis. N Z Med J 2015;128:81–3.
14. Shrikrishna D, de la Rua-Domenech R, Smith NH, et al. Human and canine pulmonary Mycobacterium bovis infection in the same household: re-emergence of an old zoonotic threat? Thorax 2009;64:89–91.
15. Twomey DF, Higgins RJ, Worth DR, et al. Cutaneous TB caused by Mycobacterium bovis in a veterinary surgeon following exposure to a tuberculous alpaca (Vicugna pacos). Vet Rec 2010;166:175–7.
16. Shrestha A, Picoy J, Torres A, et al. A case report of transmission and disease caused by Mycobacterium caprae and Mycobacterium bovis in Lima, Peru. BMC Infect Dis 2021;21:1265.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

pISSN 2210-9099 eISSN 2233-6052
https://doi.org/10.24171/j.phrp.2024.0319
Osong Public Health Res Perspect 2025;16(3):300-307

Brief Report

Standardizing the approach to clinical-based human microbiome research: from clinical information collection to microbiome profiling and human resource utilization

Jung Wook Kim, Eun Chae Choi, Kwang Jun Lee

Division of Zoonotic and Vector Borne Diseases Research, Center for Infectious Disease Research, National Institute of Health, Korea Disease Control and Prevention Agency, Cheongju, Republic of Korea

ABSTRACT

Objectives: This study presents the standardized protocols developed by the Clinical-Based Human Microbiome Research and Development Project (cHMP) in the Republic of Korea.
Methods: It addresses clinical metadata collection, specimen handling, DNA extraction, sequencing methods, and quality control measures for microbiome research.
Results: The cHMP involves collecting samples from healthy individuals and patients across various body sites, including the gastrointestinal tract, oral cavity, respiratory system, urogenital tract, and skin. These standardized procedures ensure consistent data quality through controlled specimen collection, storage, transportation, DNA extraction, and sequencing. Sequencing encompasses both amplicon and whole metagenome methods, followed by stringent quality checks. The protocols conform to international guidelines, ensuring that the data generated are both reliable and comparable across microbiome studies.
Conclusion: The cHMP underscores the importance of methodological standardization in enhancing data integrity, reproducibility, and advancing microbiome-based research with potential applications for improving human health outcomes.

Keywords: Metagenomics; Microbiota; Standardization

Received: November 14, 2024
Revised: February 4, 2025
Accepted: March 6, 2025

Corresponding author:
Kwang Jun Lee
Division of Zoonotic and Vector Borne Diseases Research, Center for Infectious Disease Research, National Institute of Health, Korea Disease Control and Prevention Agency, 220 Osongsaengmyeong 2-ro, Osong-eup, Heungdeok-gu, Cheongju 28160, Republic of Korea
E-mail: kwangjun@korea.kr

Jung Wook Kim and Eun Chae Choi contributed equally to the work as the co-first authors.

Introduction

The human microbiome comprises all microbes inhabiting various organs and their associated ecosystems [1]. Advancements in high-throughput sequencing and bioinformatics have made microbiome research more feasible, revealing links between microbiomes and diseases [2]. In this field, standardization is crucial to ensure comparable results across studies and accelerate

The opinions expressed by authors contributing to this journal do not necessarily reflect the opinions of the Korea Disease Control and Prevention Agency and Korea National Institute of Health.

© 2025 Korea Disease Control and Prevention Agency.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

research progress. International initiatives, including the European Union Metagenomics of the Human Intestinal Tract project (http://www.gutmicrobiotaforhealth.com/metahit) and the National Institutes of Health (NIH) Human Microbiome Project (http://commonfund.nih.gov/human-microbiome-project-hmp), have sought to standardize microbiome research methods. The International Human Microbiome Consortium (http://www.human-microbiome.org/) was organized to coordinate these efforts. However, it offers limited guidelines, as it neither covers the full scope of human microbiome research nor includes considerations for other study types.

Since 2023, the Korea Disease Control and Prevention Agency (KDCA) and the Korea National Institute of Health (KNIH), in collaboration with the Ministry of Health and Welfare, have led a Clinical-Based Human Microbiome Research and Development Project (cHMP). The project collects clinical samples from both healthy individuals and patients, including specimens from the gastrointestinal tract, oral and respiratory systems, urogenital system, and skin. These samples are analyzed by a sequencing consortium employing standardized methods to ensure data harmonization. All data are stored, managed, and made publicly available to researchers through an integrated database within the KNIH (Figure 1). Data generated by the cHMP adhere to principles of standardization, digitization,

HIGHLIGHTS

- Standardized clinical microbiome research protocols were developed to achieve consistent data quality.
- Samples were collected from gastrointestinal, oral, respiratory, urogenital, and skin sites.
- The reliability of data was ensured through strict quality control in amplicon and metagenome sequencing procedures.
- Protocols were established for data storage, management, and disclosure in the microbiome data core.
- Resources have been deposited in National Bank of Korea, Clinical & Omics Data Archive, and National Culture Collection for Pathogens.

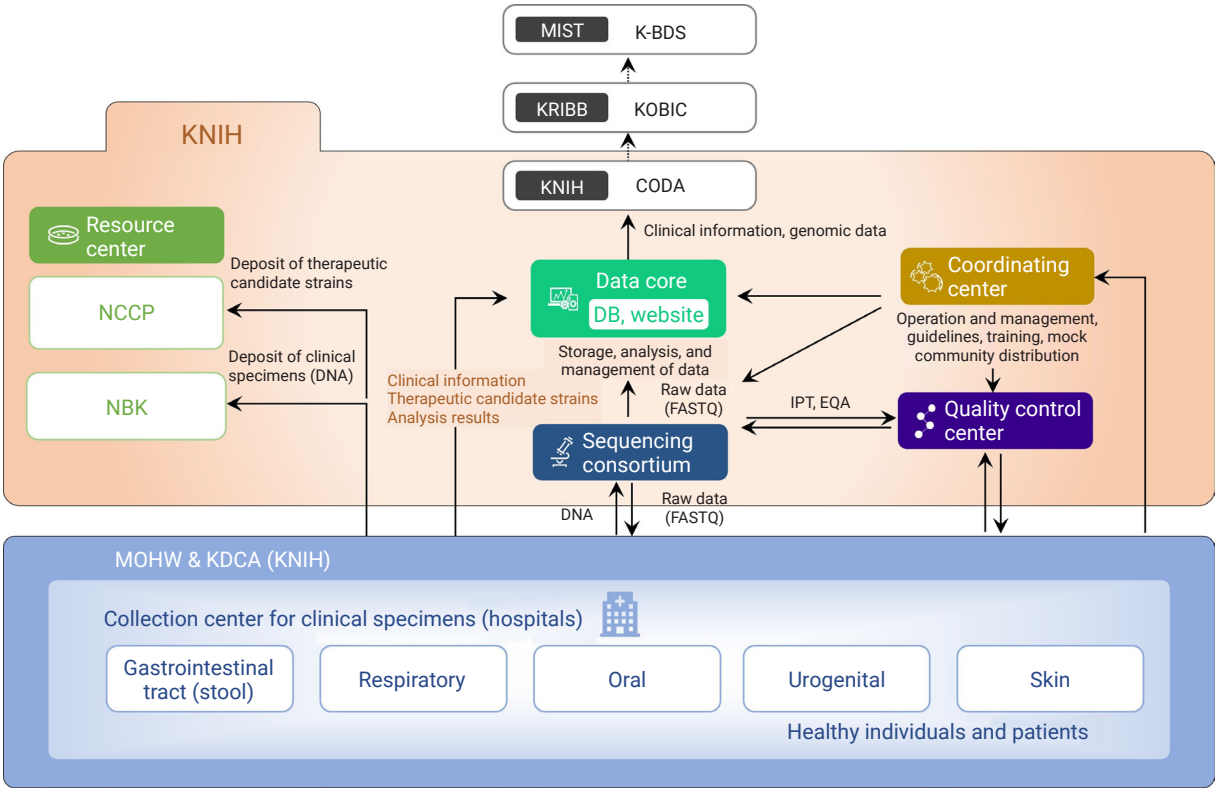


Figure 1. Structure and governance of Clinical-Based Human Microbiome Research and Development Project, Republic of Korea.
KNIH, Korea National Institute of Health; MIST, Ministry of Science and ICT; K-BDS, Korea BioData Station; KRIBB, Korea Research Institute of Bioscience & Biotechnology; KOBIC, Korea Bioinformation Center; CODA, Clinical & Omics Data Archive; MOHW, Ministry of Health and Welfare; KDCA, Korea Disease Control and Prevention Agency.

centralization, and national-level standard data sharing. Here, we summarize the principles and processes established for standardizing the cHMP.

Materials and Methods

Clinical Metadata Collection

Accurate microbiome data collection necessitates corresponding clinical metadata, which is essential for interpreting metagenome and multiomics data in clinical settings. Essential patient information comprises details on both antibiotic and non-antibiotic medication use, dietary habits, and health history recorded within 6 months of specimen collection. Clinical data are collected via case report forms and anonymized by assigning unique participant codes. The case report form excludes identifying information such as names or registration numbers, and the rate of missing clinical data should be less than 10%. Participants are categorized into disease, healthy, and disease control groups; the disease control group comprises individuals without the disease under study. Demographic, comorbidity, and medication data are collected for all groups, while additional blood tests are performed for disease groups and controls, with the exception of oral and skin specimens (Table 1). Essential information items have been defined for gastrointestinal, respiratory, urogenital, and oral specimens. For gastrointestinal specimens, information regarding bowel habits, daily activities, and dietary habits is mandatory (Table 2).

Sample Collection, Storage, and Transportation

Sample collection

For gut microbiota analysis, specimens such as feces, colonic biopsies, and rectal swabs are utilized [1]. Colonic biopsies are invasive and challenging to obtain from healthy individuals, as they necessitate a colonoscopy. The condition of the stool specimen is recorded according to the Bristol stool chart. A minimum of 1 g of solid stool and 5 mL of liquid stool is required [3]. Rectal swabs are selectively employed for gut microbiota analysis due to the high risk of human DNA contamination. Urogenital specimens mainly include vaginal swabs [4–6] and urine samples. In addition, cervical and

Table 1. Essential and additional clinical information for healthy individuals, disease groups, and disease control groups

Categories of information	Details	Disease and control group	Healthy group
Demographic information	Institution, subject registration number, date of birth, gender, height, weight, blood pressure, pulse, body temperature, smoking history, alcohol consumption history, pet ownership in the last 2 years, presence of family members, birth by cesarean section, right-handed or left-handed, highest education level, hospitalization in the last 6 months, ICU admission in the last 6 months, surgical history	Required	Required
Underlying diseases	Hypertension, diabetes, myocardial infarction, heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic obstructive pulmonary disease, connective tissue disease, peptic ulcer disease, liver disease, hemiplegia, chronic kidney disease, localized solid tumor, metastatic solid tumor, leukemia, lymphoma, human immune deficiency virus infection, organ transplantation, inflammatory bowel disease, irritable bowel syndrome, history of gastrointestinal surgery, *Clostridioides difficile* infection, atopy, allergic rhinitis, asthma, food/drug allergy, depression, schizophrenia, other psychiatric diagnoses, thyroid Disease	Required	Required
History of antibiotic and medication use in the last 6 months	Antibiotics, systemic steroids, immunosuppressants, probiotics, acid suppressants, diabetes medications, non-steroidal anti-inflammatory drugs, psychiatric medications (medication start and end dates, ingredients)	Required	Required
Results of blood test[a)]	White blood cell (mm^3), hemoglobin (g/dL), platelets (mm^3), C-reactive protein (mg/L), blood urea nitrogen (mg/dL), creatinine (mg/dL), aspartate transaminase (IU/L), alanine transferase (IU/L), total bilirubin (mg/dL), glucose (mg/dL), albumin (g/dL)	Required	
Disease-related information	Disease name, time of disease diagnosis, stage or severity of disease, specific test findings for the disease	Required	
Prognosis-related information	Mortality status, time of death, last follow-up time	Additional	

ICU, intensive care unit.
[a)]Optionally collected oral and skin samples.

https://doi.org/10.24171/j.phrp.2024.0319

Table 2. Essential clinical information required for disease groups and disease control groups according to the research field

Research field	Categories of information	Details
Gastrointestinal tract	Bowel habits and daily lifestyle	Average daily bowel movements, characteristics of bowel movements, frequency of overall exercise, average daily brushing frequency
	Dietary habits	Breakfast consumption, frequency of 3 meals a day, western dietary habits (high-fat, high-calorie, high-protein diet focusing on various meats, flour, butter), Mediterranean dietary habits (diet emphasizing vegetables, fish, olive oil while minimizing dairy, red meat, alcohol consumption), gluten-free dietary habits (avoiding wheat or rye), daily consumption of dairy products (milk and yogurt, etc.), daily consumption of fruits or vegetables, frequent consumption of kimchi with almost every meal
	Dietary preferences	Vegan (no consumption of animal products including meat, fish, dairy, eggs), pescatarian (no meat but consumption of fish and seafood), flexitarian (consumption of poultry such as chicken), ovo-lacto vegetarian (no meat or fish but consumption of eggs and dairy), ovo-vegetarian (no dairy but consumption of eggs), lacto-vegetarian (no meat but consumption of dairy products like milk, yogurt, butter), ketogenic (reduction of carbohydrates and increased fat intake), calorie restriction (for over a month), average frequency of eating out in the past year (every meal, workplace, catering, school meal), foods completely not eaten, allergies to specific foods
Respiratory	Upper respiratory	Endoscopic findings, allergic history
	Lower respiratory	Allergic history, FEV1, FVC, FEF25%–75% (forced expiratory flow between 25% and 75% of FVC), PC20
Genitourinary	Medical and treatment history (common for both sexes)	History of urinary tract infections, history of sexually transmitted infections within the past 2 years (condyloma, human papillomavirus, HIV, herpes simplex virus type 2, syphilis, gonorrhea, chlamydia, etc.), history of kidney transplantation, urinary incontinence
		Genitourinary cancer, bladder disorders (overactive bladder syndrome, neurogenic bladder), urinary stones, use of sex hormone preparations (such as estrogen, progesterone, testosterone), genitourinary surgery history, use of urinary catheter within the past year
	Sexual history (common for both sexes)	Date of recent sexual activity, history of anal intercourse, number of sexual partners within the past year, use of medications/devices within the past week (emergency contraceptive pills, male/female condoms, lubricants, other contraceptive devices)
	Specific to females	Pregnancy, number of pregnancies, number of deliveries, history of cesarean section, breastfeeding status, menopausal age, last menstrual period, average menstrual cycle, average menstrual duration, menopause status/age, use of chlorinated pools or hot tubs within 48 hours, vaginal cleansing within 48 hours, use of vaginal suppositories within the past week, pelvic congestion syndrome, history of hysterectomy
	Specific to males	Chronic prostatitis, benign prostatic hyperplasia, history of circumcision
Oral	Investigation items when collecting oral samples from patients with oral diseases	Daily tongue cleaning, use of interdental brushes, use of dental floss, use of mouthwash, use of oral irrigators (water-pick), dental treatment within the last 3 months, scaling treatment within the last 3 months, conditions of oral soft tissues (gums, oral mucosa, tongue, tonsils, palate), number of teeth, number of untreated dental caries, presence and severity of periodontal disease (evaluated based on the community periodontal index of treatment needs)

FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; HIV, human immune deficiency virus.

urethral swabs can be used selectively for specific research purposes. Urine samples can be collected via various methods, including clean-catch midstream urine [7,8], catheterized urine, and suprapubic aspiration. However, suprapubic aspiration has been excluded as a practical collection method due to its invasive nature. Respiratory specimens are collected from both the upper airways (e.g., nasopharyngeal and oropharyngeal swabs) [9,10] and the lower airways (e.g., sputum, bronchial washing, and bronchoalveolar lavage [BAL]) [11]. For oral microbiome analysis, saliva is the preferred specimen, collected either by non-stimulated methods [12] or through rinsing. Subgingival plaque is collected using a curette-based [13] or paper strip-based method [14]. Skin microbiome sampling primarily relies on swabbing and taping, with instructions to refrain from washing the area shortly before collection. For lesion sampling, both the lesion and adjacent non-lesion areas are sampled.

Preventing contamination during sample collection is vital, and gloves and sterilized tools are required. Different preprocessing techniques are used depending on the analysis type—metagenome, metatranscriptome, metabolome, or culturome—to maintain sample integrity and support accurate microbiome research.

Sample storage and transportation

The storage method depends on the time interval between specimen collection at the hospital and delivery to the analytical institution (consortium). If sample delivery to the research institution can be achieved within 2 hours of collection, the specimen should be immediately placed in an icebox for transport. For delivery times between 2 to 4 hours, specimens should be refrigerated at 4 °C until delivery, after which they are placed in an icebox for transportation. For delivery times exceeding 4 hours, specimens must be stored at –20 °C and transported in a frozen state. All specimens should reach the analytical institution within 72 hours of collection, with frozen specimens transported within 24 hours under a maintained cold chain. Upon receipt, specimens must be stored at –70 to –80 °C, as appropriate, to minimize freeze–thaw cycles. The analytical institution must conduct nucleic acid extraction from the received specimens within 72 hours.

Nucleic Acid Extraction, Storage, and Transportation

For stool specimens, frozen samples are thawed at room temperature, homogenized with a spatula, and an appropriate volume is aliquoted for DNA extraction. Urogenital specimens may undergo additional preprocessing steps, such as homogenization and centrifugation, or the sediment may be used after centrifugation, depending on specimen conditions and research objectives. Preliminary validation must be conducted before analysis. Frozen urine specimens are thawed at room temperature and centrifuged at speeds exceeding 3,000 × *g* for at least 10 minutes at 4 °C. For upper respiratory swab specimens, the swabs are aseptically cut with a sterile scalpel and vortexed with a transport medium; DNA extraction is performed after removing liquid and swab debris. For induced sputum, mucus removal is essential. Bronchial washings and BAL fluids may be concentrated as needed before nucleic acid extraction. Oral specimens often contain substantial human DNA; therefore, selective removal of human DNA can enhance microbiome sequencing efficiency. Several methods may be employed to reduce host genomic DNA contamination and enrich microbial DNA from oral samples, including differential lysis, enzymatic depletion, commercial host DNA removal kits, human-specific DNA blocking during library preparation, and density-based ultracentrifugation. DNA extraction is carried out according to IHMS SOP 01 ver. 2. Each experiment must include either a commercially available mock community or a custom-made mock community designed for the experiment's specific purpose. DNA should be stored at 4 °C for up to 1 week and at –70 to –80 °C for longer storage periods [15]. The DNA must be transported to the sequencing consortium while maintaining the cold chain.

Sequencing Analysis

DNA samples provided to the sequencing consortium were used to prepare libraries for both amplicon sequencing and whole metagenome sequencing. Sequencing was performed using equipment specifically designated and validated for microbiome analysis by the consortium. The consortium periodically conducted parallel tests to verify sequencing data reliability across instruments and runs; data were deemed reliable if the Bray-Curtis dissimilarity between results was below 0.3. Raw metagenomic reads underwent preprocessing through the following steps: (1) trimming low-quality bases (Phred score >20) and adapter sequences using tools like Trimmomatic; (2) removing duplicate reads (e.g., using Trf); and (3) filtering out human-derived reads by aligning against the human reference genome GRCh37 (hg19) or GRCh38 (hg38) with read mapping programs such as Bowtie or BWA.

Amplicon sequencing

DNA libraries were prepared targeting the hypervariable V3–V4 region of the 16S rRNA gene. The V3–V4 region was amplified using the 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3') primers. The amplification program comprised: initial denaturation at 95 °C for 5 minutes; 25 cycles of denaturation at 95 °C for 30 seconds, primer annealing at 55 °C for 30 seconds, and extension at 72 °C for 30 seconds; followed by a final elongation at 72 °C for 5 minutes. Post-sequencing, a minimum of 20,000 quality-controlled reads is required for fecal specimens and 5,000 for other human tissue specimens. The amplicon size should be at least 1,200 bp; if shorter amplicon sequences are used, their use must be justified. Databases such as the National Center for Biotechnology Information (NCBI), SILVA, and Greengene are recommended. Software performance should be validated using virtual data for quality control (QC); recommended programs include Divisive Amplicon Denoising Algorithm 2, K-mer-based Rapid Classification and Identification of Metagenomes, Ribosomal Database Project classifier, and Quantitative Insights Into Microbial Ecology [16–19]. The output should be saved as a CSV table. The table's first row should list sample names (columns 2 to the last column), while the first column (from the second row onward) should list taxon names; subsequent columns should indicate the proportion of each taxon for each sample, displayed to 4 decimal places. The table should consist of 6 levels: phylum, class, order, family, genus, and species. Bacterial names should adhere to the NCBI Taxonomy version released after January 1, 2023 (Figure 2).

https://doi.org/10.24171/j.phrp.2024.0319

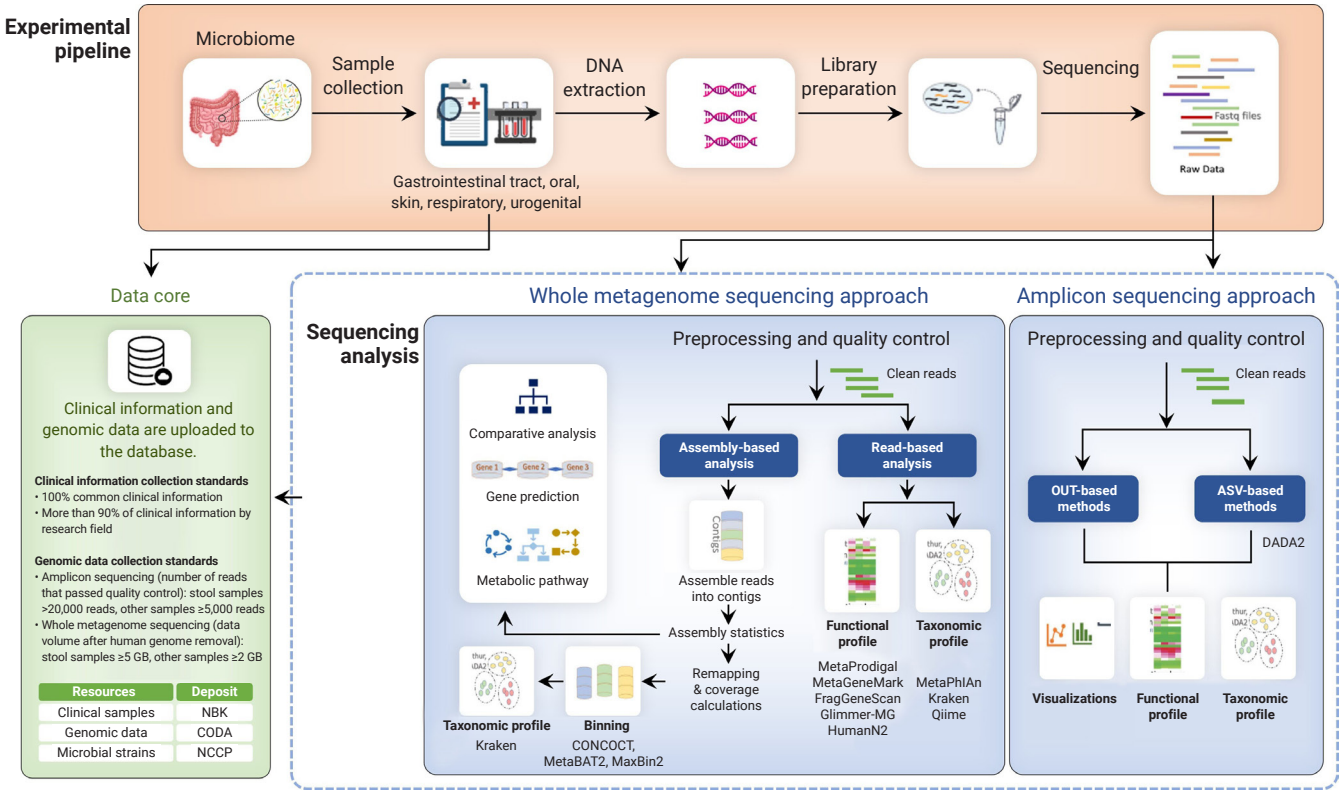


Figure 2. Schematic flow of clinical sample collection, transfer, analysis and storage processes in the Clinical-Based Human Microbiome Research and Development Project, Republic of Korea.
NBK, National Biobank of Korea; CODA, Clinical & Omics Data Archive; NCCP, National Culture Collection for Pathogens.

Whole metagenome sequencing

After sequencing, the quantity of data that passes QC and human genome removal should be at least 5 GB for fecal samples and 2 GB for other human tissue specimens. Before use, software performance should be validated with virtual data for QC; MetaPhlAn [20] and Kraken [17] are recommended for this purpose. The output format should match that used for amplicon sequencing. The assembly process is necessary to generate bacterial genomes from the microbiome. Co-assembly should be performed using MEGAHIT, MetaSPAdes, and MetaVelvet [21–23], followed by contig binning using CONCOCT, MetaBAT2, and MaxBin2 [24–26]. The quality of the generated metagenome-assembled genome (MAG) should be assessed based on its completeness and contamination. Completeness is evaluated using the ratio of single-copy genes, while contamination is determined by the frequency of single-copy genes present more than once. MAGs that meet the criteria for medium or higher quality are selected according to MAG evaluation standards. For gene function prediction, tools such as MetaProdigal, MetaGeneMark, FragGeneScan, and Glimmer-MG [27–30] are recommended. Annotation must include at least 1 of the following: Enzyme Commission number, Kyoto Encyclopedia of Genes and Genomes Orthology, Pfam number, or Gene Ontology terms (Figure 2).

Quality Control

Within the cHMP, external quality assessment is performed using mock community standard materials produced by a QC center. Inter-laboratory proficiency tests (IPT) are carried out using a small amount of positive control material. A comparison study was performed on reagents, and target values were assigned to QC materials at each stage. IPT is conducted on 1% of the analyzed samples, randomly selected. The QC center receives these samples to verify the consistency of test results obtained through nucleic acid extraction and sequencing analysis.

Data Collection, Disclosure, and Deposit of Human Resources

All clinical and genomic data collected through this project are stored in the data core, where authorized researchers can access and analyze them (Figure 1). The data sets from cHMP deemed publicly shareable are made available to the public.

Human resources collected in this project (referring to DNA extracted from samples) are generally deposited in the National Bank of Korea (NBK) (Figure 2). The deposit procedure complies with the "Regulations on the Operation and Management of the NBK (KDCA Directive No. 42)" and the "Guidelines for the Management of Human Resources of the NBK" under Section 2.2, Deposit of Human Resources.

Results

Standardizing experimental protocols is essential to minimize biases in microbiome research, especially in sample collection, DNA extraction, sequencing, and analysis. These guidelines provide recommended practices for the initial stages of microbiome studies across various fields, including gastrointestinal, respiratory, oral, skin, urogenital, and clinical research.

Discussion

The cHMP can substantially advance microbiome-based research and clinical applications by establishing rigorous protocols for sample collection, processing, and analysis, which may lead to improved human health outcomes. However, the current Korean microbiome initiative faces several challenges: sporadic government initiatives affecting research and development (R&D) efficiency and national competitiveness, insufficient establishment and management of microbiome banks and databases, inadequate comprehensive R&D support for various disease targets, a lack of commercialization-based technologies for microbiome applications, and an absence of strategies to achieve outcomes distinct from previous initiatives. To overcome these challenges, sustained efforts are required to integrate, manage, and standardize data through large-scale, inter-agency national R&D projects involving multiple government ministries.

Notes

Ethics Approval
Not applicable.

Conflicts of Interest
The authors have no conflicts of interest to declare.

Funding
This research was supported by the National Institute of Health (NIH) Research Project (Project No. 2023-NI-019-01).

Availability of Data
All data generated or analyzed during this study are included in this published article. For other data, these may be requested through the corresponding author.

Authors' Contributions
Conceptualization: ECC, KJL; Data curation: JWK; Funding acquisition: KJL; Project administration: JWK, ECC; Investigation: JWK, ECC; Visualization: JWK; Writing–original draft: JWK; Writing–review & editing: all authors. All authors read and approved the final manuscript.

References

1. Berg G, Rybakova D, Fischer D, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome 2020;8:103.
2. Aggarwal N, Kitano S, Puah GR, et al. Microbiome and human health: current understanding, engineering, and enabling technologies. Chem Rev 2023;123:31–72.
3. Santiago A, Panda S, Mengels G, et al. Processing faecal samples: a step forward for standards in microbial community analysis. BMC Microbiol 2014;14:112.
4. Wise NM, Wagner SJ, Worst TJ, et al. Comparison of swab types for collection and analysis of microorganisms. Microbiologyopen 2021; 10:e1244.
5. National Center for Biotechnology Information (NCBI), National Institutes of Health (NIH) (US). Manual of Procedures for Human Microbiome Project; Core Microbiome Sampling Protocol A (HMP-A), ver. 12.0. NCBI. NIH; 2010.
6. Forney LJ, Gajer P, Williams CJ, et al. Comparison of self-collected and physician-collected vaginal swabs for microbiome analysis. J Clin Microbiol 2010;48:1741–8.
7. Leber AL, Burnham CA. Clinical microbiology procedures. 4th ed. ASM Press; 2016.
8. Karstens L, Asquith M, Caruso V, et al. Community profiling of the urinary microbiota: considerations for low-biomass samples. Nat Rev Urol 2018;15:735–49.
9. Marty FM, Chen K, Verrill KA. How to obtain a nasopharyngeal swab specimen. N Engl J Med 2020;382:e76.
10. Mancabelli L, Ciociola T, Lugli GA, et al. Guideline for the analysis of the microbial communities of the human upper airways. J Oral Microbiol 2022;14:2103282.
11. Carney SM, Clemente JC, Cox MJ, et al. Methods in lung microbiome research. Am J Respir Cell Mol Biol 2020;62:283–99.
12. Lim Y, Totsika M, Morrison M, et al. The saliva microbiome profiles are minimally affected by collection method or DNA extraction protocols. Sci Rep 2017;7:8523.
13. Hussein H. Oral sampling techniques. Methods Mol Biol 2021;2327:17–29.
14. Pei J, Li F, Xie Y, et al. Microbial and metabolomic analysis of gingival crevicular fluid in general chronic periodontitis patients: lessons for a predictive, preventive, and personalized medical approach. EPMA J 2020;11:197–215.
15. Cardona S, Eck A, Cassellas M, et al. Storage conditions of intestinal microbiota matter in metagenomic analysis. BMC Microbiol 2012; 12:158.
16. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: high-resolution

https://doi.org/10.24171/j.phrp.2024.0319

sample inference from Illumina amplicon data. Nat Methods 2016;13:581–3.
17. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol 2019;20:257.
18. Lan Y, Wang Q, Cole JR, et al. Using the RDP classifier to predict taxonomic novelty and reduce the search space for finding novel organisms. PLoS One 2012;7:e32491.
19. Hall M, Beiko RG. 16S rRNA gene analysis with QIIME2. Methods Mol Biol 2018;1849:113–29.
20. Segata N, Waldron L, Ballarini A, et al. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods 2012;9:811–4.
21. Haider B, Ahn TH, Bushnell B, et al. Omega: an overlap-graph de novo assembler for metagenomics. Bioinformatics 2014;30:2717–22.
22. Luo R, Liu B, Xie Y, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 2012;1:18.
23. Namiki T, Hachiya T, Tanaka H, et al. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 2012;40:e155.
24. Alneberg J, Bjarnason BS, de Bruijn I, et al. Binning metagenomic contigs by coverage and composition. Nat Methods 2014;11:1144–6.
25. Kang DD, Li F, Kirton E, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 2019;7:e7359.
26. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 2016;32:605–7.
27. Hyatt D, LoCascio PF, Hauser LJ, et al. Gene and translation initiation site prediction in metagenomic sequences. Bioinformatics 2012;28:2223–30.
28. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res 2010;38:e132.
29. Rho M, Tang H, Ye Y. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res 2010;38:e191.
30. Kelley DR, Liu B, Delcher AL, et al. Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. Nucleic Acids Res 2012;40:e9.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

Instruction for authors

Enacted: January 1, 2010
Last revision: June 1, 2024
Partial revision: December 4, 2024

Osong Public Health and Research Perspectives (PHRP) is the international bimonthly (published at the end of February, April, June, August, October, and December) journal founded in 2010 by the Korea Disease Control and Prevention Agency (KDCA). With the mission of the KDCA, to create a disease-free world, PHRP encourages sharing medical information and knowledge in the areas of public health. PHRP publishes original articles, review articles, guidelines, data profiles (including cohort profiles), special articles, short communications, viewpoints, editorials and correspondence, with a focus on the following areas of expertise: emerging infectious diseases, vaccinology, zoonotic diseases, non-communicable diseases, intractable and rare diseases, and human genomics.

Before submitting a manuscript, authors should carefully read and follow the instructions for writing an article for PHRP. For issues not addressed in these instructions, authors should refer to the Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals (http://www.icmje.org/recommendations/) from the International Committee of Medical Journal Editors (ICMJE). Manuscripts submitted to PHRP that do not follow these instructions will be returned to the authors without further review.

Contact Us

Editorial Office: Korea Disease Control and Prevention Agency National Center for Medical Information and Knowledge, 202 Osongsengmyung 2nd street, Osong-eup, Heungdeok-gu, Cheongju 28159, Korea
E-mail: ophrp@korea.kr

Table of Contents

ARTICLE PROCESSING CHARGES
RESEARCH AND PUBLICATION ETHICS
EDITORIAL POLICY
SUBMISSION & PEER REVIEW PROCESS
MANUSCRIPT PREPARATION
FINAL PREPARATION FOR PUBLICATION

ARTICLE PROCESSING CHARGES

The author does not have pay publication charges for open access. The KDCA will pay to make the article open access.

RESEARCH AND PUBLICATION ETHICS

The journal adheres to the guidelines and best practices published by professional organizations, including the ICMJE Recommendations and the Principles of Transparency and Best Practice in Scholarly Publishing (joint statement by the Committee on Publication Ethics [COPE], Directory of Open Access Journals [DOAJ], World Association of Medical Editors [WAME], and Open Access Scholarly Publishers Association [OASPA]; https://doaj.org/bestpractice). Furthermore, all processes of handling research and publication misconduct shall follow the applicable COPE guidances (https://publicationethics.org/)

Human and Animal Rights

Clinical research should be conducted in accordance with the World Medical Association's Declaration of Helsinki (https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/) and approved by the Institutional Review Board (IRB) of the institution where the experiment was performed. Animal experiments should also be reviewed by an appropriate committee (Institutional Animal Care and Use Committee [IACUC]) for the care and use of animals. Studies involving pathogens requiring a high degree of biosafety should pass review of a relevant committee (Institutional Biosafety Committee [IBC]). Clinical studies that do not meet the Helsinki Declaration will not be considered for publication.

Statement of Informed Consent and Institutional Approval

The editor of PHRP may request submission of copies of informed consent forms from human subjects in all studies and IRB approval documents. For articles involving human subjects who can be identified through descriptions, photographs, or pedigrees, a signed informed consent statement must be provided. This consent must be obtained from each identifiable participant, or from a parent or legal guardian if the participant is unable to provide consent. The statement should explicitly permit the publication of the relevant descriptions, photographs, and pedigrees in both print and online formats. Articles describing the use of animals in experiments must be approved by the relevant authorities.

© 2025 Korea Disease Control and Prevention Agency

Protection of privacy and confidentiality

Patients have a right to privacy that must not be violated without informed consent. Identifying information, such as names, initials, or hospital numbers, should not be published unless essential for scientific purposes and with written informed consent from the patient (or parent or guardian). Nonessential identifying details should be omitted. If there is any doubt about maintaining anonymity, informed consent is necessary, as masking the eye region in photographs is inadequate. If identifying characteristics are deidentified, authors must assure, and editors must confirm, that these changes do not distort scientific meaning.

Originality

Manuscripts are considered with the understanding that no part of the work has been published previously in print or electronic format and the paper is not under consideration by another publication or electronic medium.

Secondary Publication

It is possible to republish manuscripts if the manuscripts satisfy the conditions for secondary publication as described in the ICMJE Recommendations (http://www.icmje.org/icmje-recommendations.pdf).

Plagiarism and Duplicate Publication

Attempting to publish substantially similar work more than once without attribution of the original source(s) is considered a redundant publication. The definition of "substantially similar" is as follows: (1) at least one of the authors is common to all reports (it is likely to be plagiarism if there are no common authors); (2) the subject or study populations are the same or similar; (3) the methodology is typically identical or nearly so and; (4) the results and interpretation vary little or not at all.

If all or part of the subject population has been reported previously, it should be declared in the Materials and Methods and must be appropriately referenced. In cases where authors are concerned with any potential overlap with published manuscripts or manuscripts being reviewed, the authors must include a letter explaining how the manuscript submitted to PHRP significantly differs from other materials. For more information, please refer to ICMJE Recommendation (available at: http://www.icmje.org/recommendations/).

Authorship and the Author's Responsibilities

Authorship credit must be based on (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; (3) final approval of the version to be published; and (4) agreeing to be accountable for all aspects of the work in ensuring that the questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The authors should meet these 4 conditions. If the number of authors exceeds 3, the specific role(s) of authors should be described at the end of the main text.

- Correction of authorship: Any requests for changes in authorship (adding author(s), removing author(s), or re-arranging the order of authors) after the initial manuscript submission and before publication should be explained in writing to the editor in a letter or e-mail from all authors. This letter must be signed by all authors of the paper. A copyright assignment must be completed by every author.
- Role of the corresponding author: The corresponding author takes primary responsibility for communication with the journal during the manuscript submission, peer review, and publication process. The corresponding author typically ensures that all of the journal's administrative requirements, such as providing the details of authorship, ethics committee approval, clinical trial registration documentation, and conflict of interest forms and statements, are properly completed, although these duties may be delegated to one or more coauthors. The corresponding author should be available throughout the submission and peer review process to respond to editorial queries in a timely manner, and after publication, should be available to respond to critiques of the work and cooperate with any requests from the journal for data or additional information or questions about the article.
- Contributors: Any researcher who does not meet all 4 ICMJE criteria for authorship discussed above but contributes substantively to the study in terms of idea development, manuscript writing, conducting research, data analysis, and financial support should have their contributions listed in the Notes section of the article.
- Recommendations for working with people with personal connections: Authors who intend to include minors (under the age of 19) or their family members (such as spouse, children, and relatives) in their research, including when publishing or presenting papers jointly with them, should clearly indicate this in the cover letter. For further information, please refer to the "Guidelines

for Preventing Illegitimate Authorship" by the National Research Foundation of Korea (https://www.nrf.re.kr/eng/).

Conflict of Interest Statement

The corresponding author must inform the editor of any potential conflicts of interest that could influence the authors' interpretation of the data. Potential conflicts should be reported even if the authors believe they were not influenced in preparing the manuscript. All authors must disclose i.e., (1) financial relationships (such as employment, consultancies, stock ownership, honoraria, paid expert testimony), (2) personal relationship, (3) academic competition, and (4) intellectual passion. These disclosures should be included on the title page.

Role of the Funding Source

The author is requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor (s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement, then this should be stated.

Inclusive Language

Authors are encouraged to use inclusive language that acknowledges diversity, respects all individuals, and promotes equal opportunities. Avoid assumptions about the reader's beliefs or identities, and refrain from language that implies superiority based on factors such as age, gender, race, ethnicity, culture, sexual orientation, disability, or health condition. Writing should be free of bias, stereotypes, and cultural assumptions. Gender neutrality is encouraged by using plural nouns (e.g., "patients") instead of gendered pronouns ("he/she"). Descriptors of personal attributes should only be used if relevant and valid.

Process for Managing Research and Publication Misconduct

When the journal faces suspected cases of research and publication misconduct such as redundant (duplicate) publication, plagiarism, fraudulent or fabricated data, changes in authorship, an undisclosed conflict of interest, ethical problems with a submitted manuscript, a reviewer who has appropriated an author's idea or data, complaints against editors, and so on, the resolution process will follow the guidances provided by the COPE (https://publicationethics.org/guidance). The editorial boards of PHRP will carry out the discussion and decision for suspected cases. We will not hesitate to publish errata, corrigenda, clarifications, retractions, and apologies when needed.

Complaints and Appeals Policy

The policies of PHRP are principally intended to protect the authors, reviewers, editors, and the publisher of the journal. The process of handling complaints and appeals follows the guidances of COPE (https://publicationethics.org/guidance).

Editorial Responsibilities

The Editorial Board will continuously work to monitor and safeguard publication ethics, including guidelines for retracting articles; maintenance of the integrity of the academic record; preclusion of business needs from compromising intellectual and ethical standards; publishing corrections, clarifications, retractions, and apologies when needed; and excluding plagiarism and fraudulent data. The editors maintain the following responsibilities: responsibility and authority to reject and accept articles; avoiding any conflict of interest with respect to articles they reject or accept; promoting publication of corrections or retractions when errors are found; and preservation of the anonymity of reviewers.

EDITORIAL POLICY

Copyright

The Korea Disease Control and Prevention Agency (publisher) holds the copyright on all submitted materials and the right to publish, transmit, sell, and distribute them in the journal or other media. The publisher applies the Creative Commons Attribution license to works it publishes. Under this license, although the publisher retains ownership of the copyright for content, it allows anyone to download, reuse, reprint, distribute, and/or copy the content for non-commercial purposes.

Open Access License

Every article appearing in this journal will be published as open-access. Articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derives (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. Author(s) do not need to

permission to use tables or figures published in PHRP in other journals, books, or media for scholarly and educational purposes.

Article Sharing (Author Self-Archiving) Policy

PHRP is an open access journal, which means that authors who publish with us are able to freely share their research in various ways, including preprint servers, social media platforms, conferences, and educational materials, in accordance with our open access policy.

• Accepted version
The accepted version, which incorporates all peer review amendments prior to final publication, may be shared on non-commercial websites and repositories. Authors who previously uploaded their article to a preprint server should update the preprint with the accepted manuscript following peer review.

• Published version
The final, published version can be shared immediately upon publication. Under the CC BY-NC 4.0 license, this version permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data Sharing Policy

To foster transparency, we encourage authors to state the availability of their data in your submission.
Subject to ethical and legal considerations, authors are encouraged to:

- Upload research data during the submission process; otherwise, share research data in a relevant public data repository with DOI for the data location.
- Include a data availability statement linking to the data. If it is not possible to share the data, use the statement to confirm why it cannot be shared.

Clinical Trials: PHRP accepts the ICMJE Recommendations for data sharing statement policies. Authors may refer to the editorial, "Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors," in the Journal of Korean Medical Science (https://dx.doi.org/10.3346/jkms.2017.32.7.1051).

Archiving Policy

PHRP provides electronic backup and preservation of access to the journal content in the event the journal is no longer published by archiving the journal content in National Library of Korea and PubMed Central (https://www.ncbi.nlm.nih.gov/pmc/journals/2151/) from the first volume, 2010.

Preprint Policy

A preprint can be defined as a version of a scholarly paper that precedes formal peer review and publication in a peer-reviewed scholarly journal. PHRP allows authors to submit the preprint to the journal. It is not treated as duplicate submission or duplicate publication. PHRP recommend authors to disclose it with DOI in the letter to the editor during the submission process. Otherwise, it may be screened from the plagiarism check program—Similarity Check (Crosscheck) or Copy Killer. Preprint submission will be processed through the same peer-review process with a usual submission. If the preprint is accepted for publication, authors are recommended to update the info at the preprint with a link to the published article in PHRP, including DOI at PHRP. It is strongly recommended that authors cite the article in PHRP instead of the preprint at their next submission to journals.

Use of Artificial Intelligence (AI)-Assisted Technologies

The journal has adopted policies, as specified by the ICMJE, regarding the use of AI in the preparation of materials intended for publication in the journal. Generative AI, including language models, chatbots, image creators, machine learning, or similar technologies, may be employed to enhance readability and language accuracy in scientific writing. However, chatbots or other AI-assisted technologies cannot be listed as authors.

Generative AI images: We do not permit the use of generative AI or AI-assisted tools for creating or altering images in submitted manuscripts. Adjustments like brightness, contrast, or color balance are acceptable if they don't obscure original information. Exceptions apply if AI is integral to the research design or methods, in which case details must be provided in the methods section. Authors should adhere to AI software policies and may be asked for pre-AI-adjusted versions of images for editorial review.

Declaration of AI and AI-assisted technologies in the writing process: Authors are required to disclose whether AI-assisted technologies were used in the production of the submitted work at the time of manuscript submission.

It must be clearly reported in a dedicated section of the Methods, or in the Acknowledgements section for article types lacking a Methods section. This disclosure should provide details about the specific tools used, including the model name, version, and manufacturer, along with an explanation of the capacity in which they were utilized. Authors should affirm that there is no plagiarism of text or images in materials produced by AI. It is not acceptable to cite AI-generated material as a primary source.

Peer Review Policy

All papers, including those invited by the editor, are subject to peer review. PHRP has adopted a double-blind peer review policy, where the author identities remain anonymous to the reviewers, and vice versa, and the identities of the reviewers and authors are visible to (decision-making) the editor throughout the peer review process. The Editorial Board selects reviewers based on expertise, publication history, and past reviews. During the peer review process, reviewers can interact directly or exchange information (e.g., via submission systems or email) with only an editor, which is known as "independent review." An initial decision will normally be made within 4–6 weeks after the reviewers agree to review a manuscript. No information about the review process or editorial decision process is published on the article page.

- AI use by peer reviewers: Generative AI tools can lack up-to-date knowledge and may produce nonsensical, biased or false information. Manuscripts may also include sensitive or proprietary information that should not be shared outside the peer review process. For these reasons we ask that, while the journal explores providing our peer reviewers with access to safe AI tools, peer reviewers do not upload manuscripts into generative AI tools. If any part of the evaluation of the claims made in the manuscript was in any way supported by an AI tool, we ask peer reviewers to declare the use of such tools transparently in the peer review report.

SUBMISSION & PEER REVIEW PROCESS

Online Submission

All manuscripts should be submitted online at https://mc04.manuscriptcentral.com/osongphrp (PHRP online submission system: ScholarOne). The entire process of manuscript submission, peer-review, and resubmission to PHRP is done through the online system.

Manuscripts submitted to PHRP will be preliminarily reviewed by the Editorial Office. Manuscripts not conforming to the instructions will be returned to the corresponding authors without being considered for publication. Submitted manuscripts are also screened for possible plagiarism or duplicate publication using Crossref Similarity Check. If a paper that might be regarded as duplicate or redundant had already been published in another journal or submitted for publication, the author should notify the fact in advance at the time of submission.

Any inquiry concerning manuscript submission should be directed to the editorial office at ophrp@korea.kr.

Peer Review Process

This journal operates a **double-blind** review process. All contributions will be initially assessed by the editor for suitability for the journal. Papers deemed suitable are then typically sent to a minimum of 2 independent expert reviewers to assess the scientific quality of the paper. The Editor is responsible for the final decision regarding acceptance or rejection of articles. The Editor's decision is final. The detailed review process is as follows.

- The Editorial Office of PHRP receives and reviews all submitted manuscripts, and all submitted manuscripts are considered confidential. The submitted manuscripts are initially screened for formatting. Once the manuscript is provisionally accepted, it is sent to the 2 most relevant reviewers for review.
- The reviewers are selected by the editor from the Editorial Board's database or the board members' recommendation. The reviewers are then requested to evaluate the manuscript based on originality, validity, presentation, and importance and interest, and, when considered necessary, statistics.
- Acceptance of a manuscript depends on the evaluation, critiques, and recommended decision made by the reviewers. A reviewer may recommend "accept," "minor revision," "major revision," and "reject." If there are conflicting decisions between reviewers, or between the author and reviewer(s), the Editor-in-Chief has the full right to decide whether the manuscript will be published in the journal. Three repeated decisions of "major revisions" are equivalent to rejection, and rejected papers will not be considered further.
- The reviewed manuscript with comments, recommendations, and revisions is returned to the corresponding author. The corresponding author is to submit the revised manuscript

accompanied by point-to-point replies to the comments given by the editor and how the revisions have been made. There should be a reasonable explanation for any noncompliance with the recommendations. In cases where references, tables, or figures are moved, added, or deleted during the revision process, renumbering must be done so that all references, tables, and figures are cited in numeric order. If the revised paper is not received within 2 months of decision, the manuscript is considered to have been withdrawn.

- When the final decision on the acceptance of the manuscript is made, the Editorial Office notifies the corresponding author. The peer-review process takes approximately 8–12 weeks.

Submission by Editors

Final decisions regarding manuscript publication are made by the editor-in-chief or a designated editor who does not have any relevant conflicts of interest. In the event that an editor has a conflict of interest with a submitted manuscript or with the authors, the manuscript will be handled by one of the other editors who does not have a conflict with the review and who is not at the same institution as the submitting editor. In such circumstances, full masking of the process will be ensured so that the anonymity of the peer reviewers is maintained.

MANUSCRIPT PREPARATION

General Requirements

- All manuscripts must be in grammatically correct English and should be created using MS Word. The manuscript must be double-spaced and written in an A4 page format. Do not leave a space between paragraphs. Only a single font (preferably Times New Roman) should be used in 11 point with margins of 2.5 cm.
- All pages should be paginated consecutively.
- All numbers should be written in Arabic numerals throughout the manuscript except for the first word of the sentence. Texts should be justified on both sides and not hyphenated and headings should be in bold letters, aligned in the center. If possible, avoid using abbreviated words at the beginning of sentences.
- Abbreviations: Where a term/definition is repeatedly referred to (i.e., 3 times in the text), it is written in full when it first appears, followed by the subsequent abbreviation in parentheses (even if it was previously defined in the abstract); thereafter, the abbreviation is used.
- Gene nomenclature: Current standard international nomenclature for genes should be adhered to. Genes should be typed in italic font and include the accession number. For human genes, use the genetic notation and symbols approved by the HUGO Gene Nomenclature Committee (http://www.genenames.org/) or refer to PubMed (http://www.ncbi.nlm.nih.gov/sites/entrez).
- Units: Système International (SI) units must be used, with the exception of blood pressure values, which are to be reported in mmHg. Please use the metric system for expressions of length, area, mass, and volume. There should be a space between the numerals and the unit symbol. When indicating time, the 24-hour system is to be used.
- Math formulae: Present simple formulae in the line of normal text where possible and use the solidus (/) instead of a horizontal line for small fractional terms, e.g., X/Y. In principle, variables are to be presented in italics. Powers of e are often more conveniently denoted by "exp." Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

Reporting Guidelines for Specific Study Designs

For specific study designs, such as randomized control studies, studies of diagnostic accuracy, meta-analyses, observational studies, and non-randomized studies, authors should adhere the reporting guidelines relevant to their specific research design and submit a checklist. A good source of reporting guidelines is the EQUATOR Network (https://www.equator-network.org/) and NLM (https://www.nlm.nih.gov/services/research_report_guide.html).

Manuscript Types

PHPR publishes editorials, original articles, review articles, guidelines, data profiles (including cohort profiles), special articles, short communications, viewpoints, editorials, commentaries, and correspondence, and book reviews.

- **Original articles** are papers containing results of basic and clinical investigations, which are sufficiently well documented to be acceptable to critical readers. These articles should be written in the following format: title page; abstract and keywords; main body (introduction, materials and methods, results, discussion, conclusion [if any]); references; and tables and figure legends. Manuscript limitations are 5,000 words, excluding the abstract,

references, and tables and figure legends.

- **Review articles** provide concise reviews of subjects important to medical researchers, and can be written by an invited medical expert. These have the same format as original articles, but the details may be more flexible depending on the content. Manuscript limitations are 6,500 words from introduction to conclusion, 100 references, 10 figures and 10 tables. The abstract should not exceed 200 words, and must be written as one unstructured paragraph.
- **Systematic reviews** should adhere to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The template for systematic reviews can be downloaded from https://ophrp.org/file/PHRP_Systematic%20Review_PRISMA.docx.
- **Guidelines** are similar to original articles, but provide evidence-based recommendations expected to impact clinical research and practice. This category can include consensus-based statements of reporting standards or clinical practice guidelines.
- **Data Profiles (including Cohort Profiles)** present large data sets from specific populations that could be analyzed in epidemiological studies. Data Profiles should be structured with the following headings in the main text: Introduction, Collection, Data Resource Use, Strengths and Weaknesses, and Access. Cohort Profiles present up-to-date information about large population-based cohorts for which long-term data collection is planned. Data Profiles should be structured with the following headings in the main text: Introduction, Study Participants, Measurements, Key Findings, Strengths and Weaknesses, and Access. The main text of Data and Cohort Profiles is limited to 4,000 words, with an unstructured abstract of up to 200 words, a maximum of 7 tables and figures, and no more than 40 references.
- **Special Articles** deal with topics or issues that are relevant to public health, but without following a traditional study format. For example, articles in this category may address scientific methodology, wide-ranging ethical and social issues, scientific methodology, or other scholarly topics. Reports from consensus committees and working groups can be published as Special Articles. This category has a main text limit of 3,500 words, with an unstructured abstract of no more than 200 words, a maximum of 7 tables and figures, and no more than 40 references.
- **Brief reports** deal with issues of importance to biomedical researchers and follow the general guidelines for original articles. The manuscript should not exceed 2,000 words, excluding references, tables and figures, and must include a structured abstract of no more than 250 words.
- **Short communications** follow the general guidelines for original articles but present original research and/or preliminary findings in a more concise format with fewer details. The manuscript should not exceed 3,000 words, excluding references, tables, and figures, and must include a structured abstract of no more than 250 words.
- **Viewpoints** may deal with almost any topic deemed to be important in the fields of public health, ethics, health law, prevention, or health policy, and are not typically written in response to a specific article. Viewpoints should have a clear focus and present material in a well-organized and scholarly manner, but should not contain novel research findings or previously unpublished data. Although we welcome unsolicited viewpoint contributions, we request that authors contact the Editorial Office (ophrp@korea.kr) prior to submission to confirm that the proposed topic is suitable for the journal. The main text of Viewpoints is limited to 3,000 words, with an unstructured abstract of up to 150 words, a maximum of 4 tables and figures, and no more than 30 references.
- **Editorials** provide invited perspective on an area of PHRP, dealing with very active fields of research, current interests, fresh insights, and debates. An abstract is not required and a brief unstructured text should be prepared. Although editorials are normally invited or written by an editor, unsolicited editorials may be submitted. Manuscript limitations are 1,000 words and 20 references.
- **Commentaries** are brief articles with a narrow focus. The journal commissions most commentaries, but unsolicited commentaries will also be considered. Commentaries may undergo peer review. The length of commentaries should be limited to 500 words, 10 references, and 2 figure or small table.
- **Correspondence** is a comment from readers regarding a published article with a reply from the authors of the article. Manuscript limitations are 500 words, 2 tables/figures, and 5 references.
- **Book reviews** may be published. Please dispatch a book to the editorial office if you think the book is essential to public health personnel.

Key features and limits of articles are summarized in Table 1 below.

Table 1. Key features and limits of articles

Type of article	Abstract (words)	Text (words)*	References	Tables and figures
Original article	Structured, 250	5,000	-	-
Review article	200	6,500	100	10 Tables 10 Figures
Data profile	200	4,000	40	7
Special article	200	3,500	40	7
Brief report	Structured, 250	2,000	-	-
Short communication	Structured, 250	3,000	-	-
Viewpoint	150	3,000	30	4
Editorial	Not required	1,000	20	-
Commentary	Not required	500	10	2
Correspondence	Not required	500	-	2

*Excluding abstract, references, tables, and figure legends.

Title Page

Title page should include (1) the title of the article (less than 50 words); (2) name of the authors (first name, middle initial, last name in capitals) and institutional affiliation including the name of department(s) and institution(s) of each author; (3) name, full address (including the postal code) of the institutional affiliation, telephone and e-mail address of the corresponding author; (4) a running title of 50 characters or less including blank spaces; and (5) notes (disclaimers). Notes include ethics approval and consent to participate, conflict of interest, funding, availability of data, authors' contributions according to the CRediT taxonomy (https://credit.niso.org/), additional contributions, and ORCID of all authors. All contributors who do not meet the criteria for authorship as defined above should be listed in an additional contribution section. Examples of those who might be acknowledged include a person who provided purely technical help, writing assistance, or a department chair who provided only general support. Authors should disclose whether they had any writing assistance and identify the entity that paid for this assistance.

Abstract and Keywords

An abstract and 3–6 relevant keywords (in alphabetical order) are required. Abstracts should be no more than 250 words in length. Abstracts should be structured, with the following section headings: Objectives, Methods, Results, Conclusion. For selecting keywords, refer to the MeSH browser (http://www.ncbi.nlm.nih.gov/mesh).

Highlights

All papers must include 3–5 short sentences presenting short summary or findings in the next of title page. The highlight section should be no more than 100 words, including spaces.

Main Body

- **Introduction** should provide concise yet sufficient background information about the study to provide the readers with a better understanding of the study, avoiding a detailed literature survey or a summary of the results.
- **Materials and methods** should contain detailed procedures of the study or experiment including investigation period, methods of subject selection, and information on subjects such as age, sex or gender, and other significant features, in order to enable the experiment to be repeated. A procedure that has been already published or standardized should be described only briefly using literature citations. Clinical trials or experiments involving laboratory animals or pathogens must elaborate on the animal care and use and experimental protocols, in addition to mentioning approval from the relevant committees. The sources of special equipment and chemicals must be stated with the name of the manufacturer. All statistical procedures used in the study and criteria for determining significance levels must be described. Ensure correct use of the terms "sex" (when reporting biological factors) and "gender" (identity, psychosocial or cultural factors). Unless inappropriate, report the sex and/or gender of study participants, the sex of animals or cells, and describe the methods used to determine sex or gender. If the study involved an exclusive population (only one sex, for example), authors should justify why, except in obvious cases (e.g., prostate cancer). Authors should define how they determined race or ethnicity, and justify its relevance. Institutional Review Board approval and informed consent procedures can be described as follows: The study protocol was approved by the Institutional Review Board of OOO (IRB No: OO-OO-OO). Informed consent was obtained by all participants (or the participant's legal guardian) / Informed consent was waived by the IRB.
- **Results** should be presented in logical sequence. Only the most important observations should be emphasized or summarized, and the main or the most important findings should be mentioned first. Tables and figures must be numbered in the order they are cited in the text, kept to a minimum, and should not be repeated. Supplementary materials and other details can be separately presented in an appendix. The authors

should state the statistical method used to analyze the results (statistical significance of differences) with the probability values given in parentheses.
- **Discussion** should contain an interpretation and explanation of the results and important aspects of the study, followed by the conclusions drawn from them. Information already mentioned in the Introduction or Results sections should not be repeated and the main conclusions of the study may be presented in the discussion.
- **Conclusion** (if any) must be linked with the purpose of the study stated in the abstract, and clearly supported by the data produced in the study. New hypotheses may be stated when warranted, but must be clearly labeled.

References

Authors are responsible for the accuracy and completeness of their references and for correct text citations.

- References are presented with [ ] following a surname in the main text, such as Kim [1] and Kim et al. [2]. When a reference is cited within the content, it is shown as [3] or [4,5] at the end. References should be searchable online.
- The last names and initials of all the authors (up to 3) should be included. For articles with more than 3 authors, list the first 3 authors only followed by "et al."
- When citing organizations that are national bodies such as government agencies, if a nationality is not part of the name, place the country in parentheses after the name, using the two-letter ISO country code.
- References cited in tables or figure legends should be included in sequence at the point where the table or figure is first mentioned in the main text.
- Do not cite abstracts unless they are the only available reference to an important concept.
- Uncompleted work or work that has not yet been accepted for publication (i.e., an "unpublished observation" or "personal communication" should not be cited as a reference). In the references list, references should be limited to those cited in the text and listed in the order in which they appear in the text. The journals should be abbreviated according to the style used in the list of journals indexed in the NLM Journal Catalog (http://www.ncbi.nlm.nih.gov/nlmcatalog/journals).
- Use of DOI is highly encouraged. Note that missing data will be highlighted at the proof stage for the author to correct.
- Other types of references not described below should follow the ICMJE Recommendations (https://www.nlm.nih.gov/bsd/uniform_requirements.html).

Please refer to the following examples.

- **Journal articles**
  1. Park AK, Kim IH, Kim J, et al. Genomic surveillance of SARS-CoV-2: distribution of clades in the Republic of Korea in 2020. Osong Public Health Res Perspect 2021; 12:37-43.
  2. Hyun J, Lee JH, Park Y, et al. Interim epidemiological and clinical characteristic of COVID-19 28 cases in South Korea. Public Health Wkly Rep 2020;13:464-74. Korean.
  3. Gultekin V, Allmer J. Novel perspectives for SARS-CoV-2 genome browsing. J Integr Bioinform 2021 Mar 15 [Epub]. https://doi.org/10.1515/jib-2021-0001.
- **Books**
  1. Riffenburgh RH, Gillen DL. Statistics in medicine. 4th ed. Academic Press; 2020.
  2. Miller DD. Minerals. In: Damodaran S, Parkin KL, editors. Fennema's food chemistry. 5th ed. CRC Press; 2017. p. 627-80.
  3. Ministry of Employment and Labor (KR). Statistics on occupational injuries and illnesses, 2008. Ministry of Employment and Labor; 2009. Korean.
- **Websites**
  1. World Health Organization (WHO) (CH). COVID-19 vaccines [Internet]. WHO; 2021 [cited 2021 Mar 15]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines.
- **Conference papers**
  1. Christensen S, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3-5; Kinsdale, IE. Springer; 2002. p. 182-91.
- **Dissertation**
  1. Park HY. The role of the thrombomodulin gene in the development of myocardial infarction [dissertation]. Yonsei University; 2000.

Tables and Figures

Tables should be simple, self-explanatory, and supplemental, and should not duplicate the text or figures. Each table must be on a separate page, not exceeding 1 page when printed, and have a concise and informative title. The tables should be numbered with Arabic numerals in consecutive order. Each column should be appropriately headed with units in parentheses if numerical measures are given. All units of measurements and concentrations must be indicated. Footnotes are followed by the source notes, other general notes, abbreviation, notes on specific parts of the table ([a], [b], [c],

[d]...), and notes on level of probability (*, **, *** for *p*).

Figures should be numbered with Arabic numerals consecutively in figure legends. The figures must not be interfered and must be clearly seen. The legend for each light microscopic image should include name of the stain and magnification. Electron microscopic images should contain an internal scale marker. All figures may be altered in size by the editor. The legends should briefly describe the data shown, explain abbreviations or reference points, and identify all units, mathematical expressions, abscissas, ordinates, and symbols.

Figures that are drawn or photographed professionally should be sent as JPG or PPT files. However, if an article receives approval for publication, files must be submitted as .tiff or .pdf. Each figure must have a caption explaining the figure. The preferred size of the images is 8 × 8 cm but 16.5 cm in width × 8 cm in length is also acceptable. It is authors' full responsibility to submit images of sufficient quality for accurate reproduction and to approve the final color galley proof. All images must be correctly exposed, sharply focused, and prepared in files of 500 dpi or more.

When tables and figures are mentioned together in the text, they should be presented in parentheses as follows: (Table 1; Figure 1), (Tables 1, 2; Figures 1–3).

Appendix and Supplemental Data

If any materials are not enough to be included in the main text such as questionnaires, they can be listed in the Appendix. Any supplementary materials that help the understanding of readers or contain too great an amount of data to be included in the main text may be placed as supplementary data. Not only a recording of the abstract, text, audio or video files, but also data files should be added here.

FINAL PREPARATION FOR PUBLICATION

Final Version

After the paper has been accepted for publication, the author(s) should submit the final version of the manuscript. The names and affiliations of the authors should be double-checked, and if the originally submitted image files were of poor resolution, higher-resolution image files should be submitted at this time. Symbols (e.g., circles, triangles, squares), letters (e.g., words, abbreviations), and numbers should be large enough to be legible on reduction to the journal's column widths. All symbols must be defined in the figure caption. If references, tables, or figures are moved, added, or deleted during the revision process, renumber them to reflect such changes so that all tables, references, and figures are cited in numeric order.

Manuscript Corrections

Before publication, the manuscript editor will correct the manuscript such that it meets the standard publication format. The author (s) must respond within 48 hours when the manuscript editor contacts the corresponding author for revisions. If the response is delayed, the manuscript's publication may be postponed to the next issue.

Proofs and Reprints

The author(s) will receive the final version of the manuscript as a PDF file. Upon receipt, the author(s) must notify the editorial office of any errors found in the file within 48 hours. Any errors found after this time are the responsibility of the author(s) and will have to be corrected as an erratum.

Correction

To correct errors in published articles, the corresponding author should contact the journal's editorial office with a detailed description of the proposed correction. Corrections that profoundly affect the interpretation or conclusions of the article will be reviewed by the editors. Corrections will be published as author correction or publisher correction in a later issue of the journal.

Minor errors will be corrected directly in the online version of the article. An indication of the correction, along with the date it was made, will be added to the article information in both the HTML and PDF versions. A separate correction note will not be published.

NOTICE: These recently revised instructions for authors will be applied beginning with the January 2025 issue.

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

Author's checklist

General Requirements

- ▫ The corresponding author (or the representative author of the co-corresponding authors) is the submitter of this manuscript.
- ▫ All manuscripts should be written in English.
- ▫ The main document with manuscript text and tables should be prepared in an MS Word (docx) or RTF file format.
- ▫ Manuscripts should be double-spaced in A4-size pages.
- ▫ Manuscripts should include line numbers.
- ▫ All pages should be numbered consecutively, starting with the abstract.

Title Page

- ▫ The title page and the rest of the manuscript text are prepared separately in two files (not combined together).
- ▫ The title page is arranged in the following order: article title, authors' full name(s), affiliation(s), and corresponding author's information, running title (less than 50 characters), notes.
- ▫ The notes section including (1) ethics approval and consent to participate, (2) conflicts of interest, (3) funding, (4) availability of data, (5) author contributions, (6) additional contributions, and ORCID is in title page, not in the manuscript.

Abstract

- ▫ The abstract does not exceed 250 words (Objectives, Methods, Results, Conclusion) for original articles and 200 words for reviews. Up to 3–6 keywords are listed at the bottom of the abstract.

Main Text

- ▫ The manuscript is organized according to following sequence: Title page, Abstract and keywords, Main text, References, Tables, and Figure legends.

Tables and Figures

- ▫ All tables and figures are numbered in the order of their appearance in a main text.
- ▫ Tables are included at the end of the manuscript as editable text and not as images.
- ▫ Figures are as separate files, in jpg, ppt, tiff, or pdf format.

References

- ▫ References are listed in proper format.
- ▫ All references listed in the reference section are cited in the text and vice versa.

© 2025 Korea Disease Control and Prevention Agency

phrp
OSONG PUBLIC HEALTH AND RESEARCH PERSPECTIVES

Copyright Transfer and Conflict of Interest Disclosure Form

Title of the paper: _____

_____

I. Copyright Transfer Form

The authors hereby transfer all copyrights in and to the manuscript named above in all forms and media, now or hereafter known, to the Korea Disease Control and Prevention Agency effective if and when the paper is accepted for publication in the *Osong Public Health and Research Perspectives*. The authors reserve all proprietary right other than copyright, such as patent rights.
Everyone who is listed as an author in this article should have made a substantial, direct, intellectual contribution to the work and should take public responsibility for it.
This paper contains works that have not previously published or not under consideration for publication in other journals.

II. Conflict of Interest Disclosure Form

All authors are responsible for recognizing any conflict of interest that could bias their work in the acknowledgments, disclosing all financial support and any other personal connections.

Please check the appropriate box below:

- ▫ No author of this paper has a conflict of interest, including specific financial interests, relationships, and/or affiliations relevant to the subject matter or materials included in this manuscript.

OR

- ▫ The authors certify that all conflicts of interest, as applicable to each author, including specific financial interests, relationships, and/or affiliations relevant to the subject matter or materials are disclosed in the manuscript.
  (Please describe in detail about these interests.)

_____

_____

_____

These interests may include one or more of the following: employment; consultancy within the past two years; ownership interests – including stock options – in a start-up company, the stock of which is not publicly traded; ownership interest - including stock options but excluding indirect investments through mutual funds and the like - in a publicly traded company; research funding; honoraria directly received from an entity; paid expert testimony within the past two years; any other financial relationship (e.g., receiving royalties); membership on another entity's Board of Directors or its advisory committees (whether for profit or not for profit).

- ▫ All authors certify that the work followed the research ethics and have approved the submission of the manuscript for publication.

© 2025 Korea Disease Control and Prevention Agency

phrp
OSONG PUBLIC HEALTH
AND RESEARCH PERSPECTIVES

	Name	Signature	Date
Corresponding author:			
Authors:			

(This form must be signed by all authors in an order as appeared in the article, and should be returned to the editorial office.)

© 2025 Korea Disease Control and Prevention Agency