ISSN-print: 2073-0764

ISSN-online: 2959-4340

TuJNaS

# Thamar University Journal of Natural & Applied Sciences

A peer-reviewed Open-Access Journal

Volume 11

Issue (1)

June 2026

Thamar University Publications

# Thamar University Journal of
# Natural and Applied Sciences (*TUJNAS*)

Thamar University Journal of Natural & Applied Sciences (TUJNAS) is a peer-reviewed journal. It is an open-access journal published by Thamar University, Dhamar, Yemen twice a year. The aim of the journal is to publish original and review articles in the fields of science, agriculture, engineering, medicine, environment, and computer science. The journal is published in English only.

The journal has the following international standard codes:

***ISSN-print: 2073-0764***

***ISSN-online: 2959-4340***

For more information on editorial policy, manuscript submission, ethics, etc., please visit the TUJNAS Journal website:

https://www.tu.edu.ye/journals/index.php/TUJNAS

# Vision, Mission, and Aims

## Vision

Through several procedures, TUJNAS aspires to be a leading regional and international natural and applied sciences journal. TUJNAS strives for inclusion in major indexing services such as Scopus and Web of Science to boost scholarly visibility and citation impact. It publishes open-access research aligned with the UN SDGs. The journal continuously refines its editorial policies, ethical guidelines, and production workflows in line with COPE and ICMJE recommendations to ensure the timely, transparent, and reliable publication of high-quality scientific content.

## Mission

TUJNAS exists to publish rigorous, high-impact research in natural and applied sciences. It upholds academic integrity through a strict double-masked peer review process, safeguarding quality and originality. The journal fosters interdisciplinary collaboration to bridge theoretical science with practical applications, driving innovative solutions to global and local challenges. Committed to open science, it provides free, immediate access to research, enhancing engagement among researchers, practitioners, and policymakers. The journal aims to address pressing societal and scientific issues worldwide by prioritizing actionable insights.

## Aims

**1. Broad disciplinary coverage: by publishing contributions in:**

- Pure Sciences: Biology, Chemistry, Physics, Mathematics, Earth and Environmental Sciences.
- Applied Sciences & Engineering: Civil, Mechanical, Electrical, Chemical, Environmental, Mechatronics, and Renewable Energy Engineering.
- Life and Health Sciences: Agriculture, Veterinary Medicine, Pharmacology, Biotechnology, Clinical Medicine, Public Health, Biomedical Engineering, and Health Informatics.
- Computer & Information Technology: Data Science, Artificial Intelligence, Software Engineering, Network Systems, Cybersecurity, and Health Informatics.

**2. Diverse article types: by encouraging submission of:**

- Original Research Articles presenting novel empirical or theoretical insights.
- Review Articles synthesizing and critically evaluating current knowledge.
- Case Studies highlighting unique applications or phenomena in real-world contexts.
- Editorials offering expert perspectives on emerging trends or policy issues.
- Short communications deliver concise, timely findings of high relevance.

**3. Regional engagement and global outreach:**

- Support Early-Career Researchers by offering clear author guidelines, mentorship through the review process, and rapid turnarounds.
- Highlight Local Contexts to address Yemen-specific environmental, agricultural, and health challenges while drawing lessons of broader relevance.
- Promote international collaboration by soliciting contributions from global experts and forming editorial alliances with related journals and societies.

**4. Ethics and Transparency:**

- Enforce strict adherence to authorship criteria, conflict-of-interest declarations, and research ethics (including human/animal welfare and data integrity).
- Embrace transparent archival of data and methods to enable reproducibility and long-term accessibility.

# Editorial Team

## General Supervisor

**Prof. Dr. Mohammed Mohammed Al-Haifi**

**Position:** Rector of the University, Thamar University, Dhamar, Yemen.

**Address:** University Presidency, Thamar University, P O Box 87246 Dhamar, Yemen.

**E-Mail:** dralhaifi@tu.edu.ye

## Editor-in-chief

**Prof. Dr. Adulkarem Esmail Zabiba**

**Position:** Vice-Rector of Postgraduate and Scientific Research, Thamar University, Dhamar, Yemen.

**Address:** Vice Presidency of the University for Postgraduate Studies and Scientific Research, Thamar University, P O Box 87246 Dhamar, Yemen.

**E-Mail:** karimzabiba@tu.edu.ye

## Editorial Director

**Prof. Dr. Abdullah Ahmed Ali Ahmed**

**Position:** Professor of Nanoscience, Physics Department, Faculty of Applied Sciences, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Applied Sciences, Thamar University, P O Box 87246 Dhamar, Yemen.

**E-Mail:** abdullah2803@tu.edu.ye

## Editorial Director Assistant

**Assoc. Prof. Annas Saeed Ahmad Al-Sharabi**

**Position:** Professor of Superconductivity, Physics Department, Faculty of Applied Sciences, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Applied Sciences, Thamar University, P O Box 87246 Dhamar, Yemen.

**E-Mail:** annas.AlSharabi@tu.edu.ye

## ✉ All correspondence should be sent to: ✉

Editorial Director,

Thamar University Journal of

Natural and Applied Sciences (TUJNAS)

**Thamar University, P O Box: 87246 Dhamar, Republic of Yemen**
**E-Mail**: tujnas@tu.edu.ye

# Advisory Board

## Assoc. Prof. Abdul Ghani Ali Mohammed

**Position:** Dean of Faculty of Agriculture & Veterinary Medicine, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Agriculture & Veterinary Medicine, Thamar University, Dhamar, Yemen.

**E-Mail:** abdulghani.ali@tu.edu.ye

## Assist. Prof. Fouad Mohammed Y. Al-Jarmouzi

**Position:** Dean of Faculty of Engineering, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Engineering, Thamar University, Dhamar, Yemen.

**E-Mail:** aljarmouzi@tu.edu.ye

## Assist. Prof. Nashwan Hamid Saleh Al-Tairi

**Position:** Dean of Faculty of Dentistry, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Dentistry, Thamar University, Dhamar, Yemen.

**E-Mail:** nashwanh9@tu.edu.ye

## Assoc. Prof. Adel Ali Ahmed Amran

**Position:** Dean of Faculty of Medical Science, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Medical Science, Thamar University, Dhamar, Yemen.

**E-Mail:** adelamran@tu.edu.ye

## Assist. Dr. Abdullah Al-Murtadha

**Position:** Dean of Faculty of Medicine, Thamar University, Dhamar, Yemen.

**Address:** Faculty of Medicine, Thamar University, Dhamar, Yemen.

**E-Mail:** abdullah.almurtadha@tu.edu.ye

# Editorial Board

## Prof. Dr. Levan Chkhartishvili (Georgia)

**Research field:** Semiconducting and Powder Composite Materials.

**Position:** Professor at the Department of Engineering Physics, Faculty of Informatics and Control Systems, Georgian Technical University, and researcher at F. Tavadze Metallurgy and Materials Science Institute, Semiconducting and Powder Composite Materials Laboratory.

**Affiliation:** Department of Engineering Physics, Faculty of Informatics and Control Systems, Georgian Technical University, 77 M. Kostava Ave., GTU Campus 4, Room 307, Tbilisi, 0160, Georgia.

**E-Mail:** levanchkhartishvili@gtu.ge , chkharti2003@yahoo.com

## Prof. Dr. Saeed M. Al-Ghalibi (Yemen)

**Research field:** Bacteriology, Fungi, Medical microbiology, and Food Microbiology.

**Position:** Deputy Dean of Faculty of Science for Academic Affairs and Graduate Studies, Faculty of Science, Sana'a University, Sana'a, Yemen

**Affiliation:** Department of Biology, Faculty of Science, Sana'a University, Sana'a, Yemen.

**E-Mail**: s.alghalabi@su.edu.ye , Alghalibi@gmail.com

## Prof. Dr. Abdulkarim A. Amad (Yemen)

**Research field:** Animal nutrition and production feed and feeding.

**Affiliation:** Faculty of Agriculture, Thamar University, Dhamar, Yemen and Institute of Animal Nutrition, Department of Veterinary Medicine, Frei Universität Berlin, Berlin, Germany.

**E-Mail**: abdulkarim.Amad@tu.edu.ye , abeerobeid@yahoo.com

## Prof. Dr. Nabil M. Al-Areeq (Yemen)

**Research field:** Petroleum Geology, Hydrology, Sedimentology, Integrated Water Resources Management, Data analysis and Resolving of Water Related Conflicts.

**Position:** Centre Director of Water Resources and Environment, Thamar University, Dhamar, Yemen.

**Affiliation:** Department of Geology and Environment, Faculty of Applied Science, Thamar University, Dhamar, Yemen.

**E-Mail**: alareeqnabil@tu.edu.ye , nabilalareeq@yahoo.com

## Assoc. Prof. AbdulSalam M. Al-Makdad (Yemen)

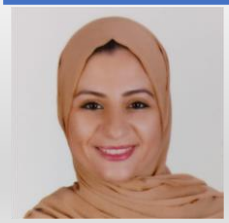
**Research field:** Diagnosis, management, and care of acute and chronic liver disease and GI diseases. Diagnostic and interventional GI endoscopy.

**Position:** President of the internal medicine department in AL-Wahda Teaching Hospital Maabar, Maabar City, Dhamar, Yemen.

**Affiliation:** Department of Internal Medicine, Faculty of Medicine, Thamar University, Dhamar, Yemen.

**E-Mail**: aalmakdad@tu.edu.ye

## Assoc. Prof. Shaimaa A. A. Momen (Egypt)



**Research field:** Entomology.

**Affiliation:** Department of Entomology, Faculty of Science, Ain Shams University, Khalifa El-Maamon Street, 11566, Cairo, Egypt.

**E-Mail**: Shaimaa_momen@sci.asu.edu.eg , Shaimaa_momen@hotmail.com

## Assoc. Prof. Essam A. Al-Moraissi (Yemen)

**Research field:** Oral and maxillofacial surgery, craniomaxillofacial trauma, temporomandibular joint disorders, orthognathic surgery, surgical pathology, cleft lip and palate, implant dentistry, lower third molar surgery, regenerative medicine, and adult mesenchymal stem cells.

**Affiliation:** Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Thamar University, Dhamar, Yemen.

**E-Mail**: dressamalmoraissi@tu.edu.ye

## Assoc. Prof. Salah Abdul-Jabbar Jassim (Iraq)

**Research field:** Thin Films, Semiconductor Devices, and Solid-State Physics.

**Affiliation:** Department of Dentistry, AL Kunooze University College, Basrah, Iraq.

**E-Mail**: salah.abdul.jabbar@kunoozu.edu.iq , salahjassim200@yahoo.com , salah.jassim@alayen.edu.iq

## Assoc. Prof. Amin Saif Ahmed (Yemen)

**Research field:** Energy and control systems engineering.

**Affiliation:** Department Mechatronics, Al-Saeed College of Engineering and Information Technology, Taiz University, Taiz, Yemen.

**E-Mail**: sameeralromima@yahoo.com , sameeralromima@gmail.com

## Assoc. Prof. Dina Salah Eldin M. Abdelrhman (Egypt)

**Research field:** Gold Nanoparticles, Photochemistry, Nanotechnology, and Nanomedicine.

**Affiliation:** Biophysics, Physics Department, Faculty of Science, Ain Shams University, Khalifa El-Maamon Street, 11566, Cairo, Egypt.

**E-Mail**: dinasalah@sci.asu.edu.eg , dandy741@hotmail.com , dandy741@gmail.com

## Assoc. Prof. Fawaz M. A. Al-Badaii (Yemen)

**Research field:** Microbiology, Antimicrobial resistance, Environmental Science, Heavy metals, and Adsorption Water quality.

**Affiliation:** Department of Biology, Faculty of Applied Sciences, Thamar University, Dhamar, Yemen.

**E-Mail**: fawaz.AlBadai@tu.edu.ye , abdualwhab1974@gmail.com

## Assoc. Prof. Abdulwahab B. Alwany (Yemen)

**Research field:** Solid State Physics, Thein Films, Materials Science, and Nanoscience.

**Affiliation:** Department of Physics, Faculty of Science, Ibb University, Ibb, Yemen.

**E-Mail**: abdualwhab@yahoo.com , abdualwhab1974@gmail.com

## Assoc. Prof. Ali Abdullah A. Al-Mehdar (Yemen)

**Research field:** Pharmacology & Therapeutics.

**Affiliation:** Department of Pharmacology and Toxicology, Faculty of Faculty of Medicine, Thamar University, Dhamar, Yemen.

**E-Mail**: ali.almehdar@tu.edu.ye , alialmehdar2006@yahoo.com

## Assoc. Prof. Sameer A. M. Abdulrahman (Yemen)

**Research field:** Pharmaceutical Analytical Chemistry and Water Treatment.

**Affiliation:** Department of Chemistry, Faculty of Education and Sciences-Rada'a, Albaydha University, Albaydha 14517, Yemen.

**E-Mail**: sameeralromima@yahoo.com , sameeralromima@gmail.com

**Assoc. Prof. Nada M. Al-Hamdani (Yemen)**

**Research field:** Histology, and Physiology, specializing in Endocrinology.

**Affiliation:** Department of Biology, Faculty of Science, Sana'a University, Sana'a, Yemen.

**E-Mail**: n.alhamdni@su.edu.ye , hamdaninadam@gmail.com

**Assoc. Prof. Abdulbari A. A. Saeed (UK)**

**Research field:** Preparation and characterization of mesoporous from solid waste as catalysis for water purification, Separation technology using an adsorption process, Water and wastewater treatment, and Biofuel production from organic solid waste.

**Affiliation:** School of Engineering, Institute for Infrastructure and Environment (IIE), University of Edinburgh, Edinburgh EH9 3JL, UK.

**E-Mail**: alborani_75@yahoo.co.uk , Abdulbari.Saeed@ed.ac.uk

**Assoc. Prof. Yahya Qaid Hasan Ali (Yemen)**

**Research field:** Differential Equations, Numerical Analysis, and Adomian Decomposition Method.

**Affiliation:** Department of Mathematics, Faculty of Applied Sciences, Thamar University, Dhamar, Yemen.

**E-Mail**: qaid.Yahya@tu.edu.ye , yahya217@yahoo.com

**Assoc. Prof. Abdullah Alwarafi (Yemen)**

**Research field:** Social Pharmacy.

**Position**: Vice Dean for Student Affairs, Faculty of Dentistry, Ibb University, Ibb, Yemen

**Affiliation:** Pharmacy Department, Faculty of Dentistry, Ibb University, Ibb, Yemen.

**E-Mail**: abdullahalwarafi@gmail.com , dentistry@ibbuniv.edu.ye

**Prof. Dr. Ahmed A. M. Alakwa (Yemen)**

**Research field:** Agricultural Economics.

**Position:** Head of Scientific Research, Vice Presidency of the University for Postgraduate Studies and Scientific Research, Thamar University, P O Box 87246 Dhamar, Yemen

**Affiliation:** Faculty of Agriculture, Thamar University, Dhamar, Yemen.

**E-Mail**: Hawali.ahmed@tu.edu.ye , alakwaahmed55@gmail.com

**Prof. Dr. Ahmed Ali Saleh Obayeha (Yemen)**

**Research field:** Orthodontics, Pediatric Dentistry, and Preventive Medicine.

**Position:** Dean of the Faculty of Dentistry, Al-Razi University, Sana'a, Yemen

**Affiliation:** Faculty of Dentistry Sana'a University, Sana'a, Yemen.

**E-Mail**: a.Obaya@su.edu.ye , Ahmedobeyah@yahoo.com

**Assoc. Prof. Fathi Ahmed ELShawish (Yemen)**

**Research field:** Identification and characterization of genetic sources of indigenous and introduced fruits, propagation and breeding of fruit crops, and design and layout of gardens.

**Position:** Deputy Dean for Postgraduate Studies and Scientific Research, Faculty of Agriculture, Thamar University, Dhamar, Yemen

**Affiliation:** Faculty of Agriculture, Thamar University, Dhamar, Yemen.

**E-Mail**: Fathi.ELShawish@tu.edu.ye ,

**Assoc. Prof. Khalid Al-Hussaini (Yemen)**

**Research field:** Information & Communication Technology (ICT), Computer Communications (Networks), Communication Engineering, and Computer Engineering.

**Position**: University Rector's Advisor for Academic Development & Automation and Vice Dean for Student Affairs, Faculty of Computer Science & Information Systems, Thamar University, Dhamar, Yemen.

**Affiliation:** Department of Information Technology, Faculty of Computer Science & Information System, Thamar University, Dhamar, Yemen.

**E-Mail**: khalid.alhussaini@tu.edu.ye

**Assoc. Prof. Rasheed M. Alsanafi (Yemen)**

**Research field:** Surveying & Urban Engineering and Planning, Civil Engineering.

**Position**: Postgraduate Studies and Scientific Research, Faculty of Engineering, Thamar University, Dhamar, Yemen.

**Affiliation:** Department of Civil Engineering Faculty of Engineering, Thamar University, Dhamar, Yemen.

**E-Mail**: alsanafy@tu.edu.ye , alsanafy@hotmail.com

**Assoc. Prof. Kamal O. I. Ba'hakem (Yemen)**

**Research field:** Surgical Gastroenterology, Upper & Lower GIT Endoscopy (diagnostic and therapeutic), Bleeding emergency GIT, Major conventional GIT surgery including biliary tree reconstructions, ERCP specialist, and General Surgeon.

**Position**: Vice director of AL-Wahda Teaching Hospital Maabar, Thamar University, Yemen.

**Affiliation:** Department of Surgery, Faculty of Medicine, Thamar University, Dhamar, Yemen.

**E-Mail**: mkamel1970@yahoo.com , khemo1970@gmail.com

**Prof. Dr. Mohammad Ghaffar Faraj (Iraq)**

**Research field:** Polymer Solar Cells and Solar Energy.

**Affiliation:** Department of Physics, Faculty of Science and Health, Koya University, Koya, Kurdistan Region, Iraq.

**E-Mail**: mohammad.ghaffar@koyauniversity.org

**Prof. Dr. Saad S. AL-Tobaili (Yemen)**

**Research field:** Graph Theory and Mathematical Modeling relationships.

**Affiliation:** Department of Mathematics, Faculty of Science, Hadhramout University, Mukalla, Hadhramout, Yemen.

**E-Mail**: saadaltabili1@yahoo.com

**Prof. Dr. Majid Raissan Al Bahrani (Iraq)**

**Research field:** Nano-Energy Materials & Devices, Solid-State Physics, Graphene, Carbon Nanotube, and Organic Photovoltaics.

**Affiliation:** Physics Department, College of Sciences, University of Thi-Qar, Nasiriyah, Iraq.

**E-Mail**: majidphy2016@utq.edu.iq or m20111973@yahoo.com

**Assoc. Prof. Hassan A. M. Al-Khawlani (Yemen)**

**Research field:** Horticulture and Biotechnology.

**Affiliation:** Agriculture research & extension authority (AREA), Dhamar, Yemen.

**E-Mail**: alkholanihassaan@gmail.com

**Assoc. Prof. Mohammed Jassim Mohammed Ali (Iraq)**

**Research field:** Integrated Optics & Photonic Devices, Optical Sensors & Biosensors, Solid-State Physics & Materials Characterization, Nanotechnology, Atomic & Molecular Spectroscopy, Computational Physics and Simulation.

**Affiliation:** Physics Department, College of Sciences, Mustansiriyah University, Baghdad, Iraq.

**E-Mail**: mohammedjassim@uomustansiriyah.edu.iq or mohammedjassim78@gmail.com

**Assoc. Prof. Amin M. A. AlWaseai (Yemen)**

**Research field:** Biotechnology and Food Technology.

**Position**: Head of Biotechnology & Food Technology Department, Faculty of Agriculture, Thamar University, Dhamar, Yemen.

**Affiliation:** Department of Biotechnology & Food Technology, Faculty of Agriculture, Thamar University, Dhamar, Yemen.

**E-Mail**: amin.alwaseai@tu.edu.ye , amin_alwaseai2000@yahoo.com

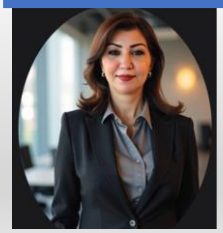
**Prof. Dr. Adil N. Ayyash (Iraq)**

**Research field:** Theoretical Physics, Molecular Physics and Spectroscopy, Spectral Analysis of Materials, and Computational Physics.

**Affiliation:** Department of Physics, College of Science, University of Anbar, Anbar, Iraq.

**E-Mail**: sc.adil_nameh78@uoanbar.edu.iq

**Assoc. Prof. Rana Ismael Khaleel (Iraq)**



**Research field:** Medical Physics.

**Affiliation:** Physics Department, College of Sciences, Mustansiriyah University, Baghdad, Iraq.

**E-Mail**: khleelrana79@uomustansiriyah.edu.iq

## Technical Team

Technical Editing by:

**Mr. Mohammed Subaie**

# Volume 11, Issue 1, June 2026

DOI: https://doi.org/10.59167/n5yahh03

# Articles

TUJNAS

Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 1– 7

DOI: https://doi.org/10.59167/131g4523

Original Article

# The Impact of Specific Microorganisms Causing the Abortion of a Random Sample of Pregnant Women in Dhamar City, Yemen

**Mobarak M. N. A. Rassam[1], Fuad Ahmed Ali Saad[1], Bakeel A. Radman[2,3]***

[1] Department of Biology, Faculty of Applied Science, Thamar University, Dhamar 87246, Yemen.

[2] Department of Biology, College of Science and Education, Albaydha University, Albaydha, Yemen.

[3] Department of Laboratory Medicine, The Third Affiliated Hospital of Southern Medical University, Guangzhou, China.

***Corresponding Author:** Bakeel A. Radman, Department of Biology, College of Science and Education, Albaydha University, Albaydha, Yemen, and Department of Laboratory Medicine, The Third Affiliated Hospital of Southern Medical University, Guangzhou, China. E-mail: b-radman@qq.com

Received: 2 November 2025. Revised: 5 March 2026. Accepted: 15 March 2026. Published: 29 June 2026.

## Abstract

This study aimed to determine the prevalence of Cytomegalovirus (CMV), Rubella, *Toxoplasma gondii*, and anti-phospholipid (APL) antibodies among women who have experienced abortion in Dhamar City, Yemen. Additionally, the research sought to identify risk factors associated with seropositivity for these infections and to compare these factors between pregnant women residing in urban and rural areas. The cross-sectional study involved 200 pregnant women, aged 15 to 38 years, who had a history of one or more unexplained recurrent abortions. These participants were recruited from various healthcare facilities, including Dhamar Hospital, local dispensaries, the Reproductive Health Center, private clinics, and diagnostic laboratories. Participants were screened for IgM and IgG antibodies against CMV, Rubella, *Toxoplasma*, and anti-phospholipids using Enzyme-Linked Immunosorbent Assay (ELISA). Demographic, socioeconomic, obstetric, and behavioral data were collected via face-to-face interviews utilizing a pretested questionnaire. The seroprevalence of CMV-specific IgG and IgM among the participants was 92.5% and 7.0%, respectively. Rubella-specific IgG and IgM antibodies were detected in 87.0% and 3.5% of the women, respectively. Furthermore, the prevalence of *Toxoplasma gondii* IgG and IgM was 62.5% and 8.0%, respectively. Seropositivity for anti-phospholipid IgG and IgM were 10.0% and 20.0%, respectively. In conclusion, the presence of IgM antibodies against Rubella, CMV, and *Toxoplasma gondii* correlated with recent or primary infections. Conversely, positive IgG results for these pathogens indicated prior exposure or convalescent infection.

**Keywords**: Serological Study; Abortion in Pregnant Women; Dhamar; Yemen

## 1. Introduction

Recurrent spontaneous abortion (miscarriage) is one of the most frequent reproductive complications of early pregnancy. It is clinically defined as three or more consecutive pregnancy losses prior to the 20$^{th}$ week of gestation, occurring before the fetus reaches a viable gestational age [1, 2]. Widely accepted etiologic causes include genetic anomalies, immunologic factors, placental abnormalities, endocrine disorders, nutritional deficiencies, environmental factors, and maternal systemic conditions such as diabetes mellitus and thyroid disease.

Additionally, maternal infections caused by *Toxoplasma gondii*, Cytomegalovirus (CMV), syphilis, rubella, and herpes, alongside the presence of anti-phospholipid (APL) antibodies, are associated with a significantly elevated risk of congenital complications [3]. Contracting these infections during pregnancy can lead to congenital anomalies and abortion, making them a leading cause of perinatal morbidity and mortality, particularly in developing nations [4]. Pathogen transmission can occur prenatally via transplacental passage, or postnatally through contact with infected blood, vaginal secretions, or breast milk (particularly for CMV and Rubella). Clinical evidence of these infections may manifest at birth, during infancy, or even years later [5].

The adverse outcomes produced by these pathogens generally mimic those of abortions, infertility, intrauterine fetal deaths, stillbirths, and congenital malformations. The prevalence of *Toxoplasma gondii*, Rubella, CMV, and APL infections varies significantly by geographic region. However, countries in Southeast Asia and Sub-Saharan Africa consistently report the highest rates of stillbirths associated with these infections [6, 7]. Rubella, *Toxoplasma*, and CMV are common causes of infection across all age groups and are generally asymptomatic; however, primary infection in pregnant women during the first trimester can cause severe fetal congenital malformations and abortion [4]. Because these maternal infections often present without symptoms and clinical diagnoses are inconsistent, it is paramount to identify susceptible women, especially those with acute maternal infections, and to recognize prevalent and recurrent pathogens [7].

Due to the absence of a national screening program, there is limited data available regarding the seroprevalence of specific IgM and IgG antibodies to *Toxoplasma gondii*, Rubella, CMV, and anti-phospholipids among pregnant women in Dhamar City. Therefore, this study aimed to evaluate the serological evidence of CMV and Rubella infections during the

first trimester among this demographic. Ultimately, the research seeks to outline the epidemiological and serological landscape of several leading infectious causes of abortion in Dhamar City, Yemen.

## 2. Materials and Methods

### 2.1. Study Period and Design

This descriptive, cross-sectional study was conducted over a 14-month period, from September 2019 to October 2020. This duration was specifically selected to ensure the collection of a robust sample size, thereby achieving statistical accuracy and enhancing the overall reliability of the results. The study cohort consisted of 200 women, aged 15 to 38 years, who had been infected by one or more abortion-causing agents. These participants were recruited while attending antenatal clinics in various hospitals and health centers across Dhamar City, Yemen.

### 2.2. Study Area

The research was carried out in Dhamar City. Geographically, the Dhamar governorate is located between latitudes 43.30 and 44.50, longitudes 14 and 15, and sits at an altitude of 2400 meters above sea level. The mean annual temperature ranges from 20 to 28 °C in the summer, while winter temperatures can drop to between 18 °C and -1 °C during the night and early morning. The relative humidity in the region averages 49% [8].

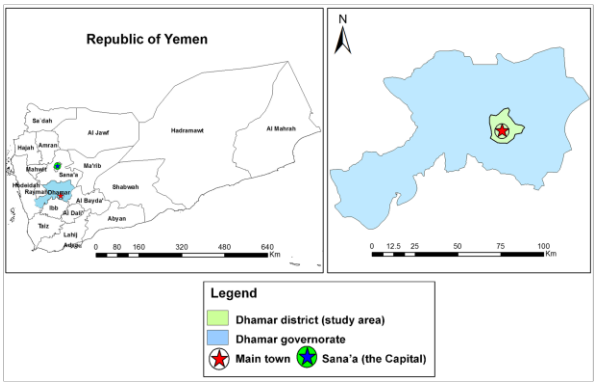



**Figure 1:** Map showing the location of Dhamar governorate and Dhamar district, Yemen (study area). Reprinted with permission from Abdulelah H. Al-Adhroey *et al.* [8] Copyright 2019 Springer Nature.

### 2.2 Study Population and Sampling

The study population comprised 200 pregnant women who exhibited infections from one or more targeted pathogens (*Toxoplasma gondii*, Rubella, CMV, and anti-phospholipid antibodies). Participants were between 15 and 38 years old and had a documented history of one or more consecutive, unexplained abortions. Participants were recruited from several facilities, including Dhamar General Hospital, Labanan Hospital, Dar AL-Shifa Hospital, and AL-Hayah Hospital, alongside local dispensaries (Mustawsaf Miahres), medical centers (AL-Amwma and Al-Tafula), and diagnostic laboratories (Al-Dubai Laboratory, New Med Laboratory, and Alpha Laboratory). Obstetric and behavioral data were gathered through face-to-face interviews using a pretested questionnaire.

### 2.3 Sample Collection

Five milliliters of blood were collected from each participant via venipuncture using sterile disposable syringes. The blood was transferred into a plain tube containing a gel clot activator (without anticoagulant) and allowed to stand for one hour at room temperature to facilitate clot formation. Subsequently, the tubes were centrifuged at 3000 rpm for 10 minutes [9]. The serum was then carefully aspirated using a Pasteur pipette, dispensed into sterile Eppendorf tubes, immediately transported in an ice-box to the laboratory, and stored at -20°C until analysis [9, 10]. Personal patient information was recorded based on standardized questionnaire responses, which captured data on age, region, residence, education status, pregnancy history, and previous abortions [9]. Each sample was labeled with a specific study code corresponding to the participant's questionnaire and informed consent form [10].

### 2.4 Laboratory Analysis

Samples were analyzed to detect the presence of specific IgG and IgM antibodies against CMV, Rubella, *Toxoplasma*, and APL. The ELISA technique was utilized in strict accordance with the manufacturer's instructions to determine both positive and negative results. ELISA was selected due to its high sensitivity for detecting specific immunoglobulin markers. A test value greater than 1.5 was classified as a positive sample, whereas a value less than 1.0 was considered negative. Values falling between 1.0 and 1.5 were deemed equivocal. All serological screenings were conducted at Al-Dubai Labs, Alfa Labs, and New Med Lab in Dhamar City using the cobas 411 analyzer (Roche Diagnostics, Mannheim, Germany) [10].

### 2.5 Statistical Analysis

Data analysis was performed using the Statistical Package for the Social Sciences (SPSS version 22.0, Chicago, IL). Results were presented using descriptive statistics, relying on frequencies and percentages to determine the statistical significance of the distribution of CMV, Rubella, *Toxoplasma*, and APL infections. Additional variables, including age, geographic region, residence, education status, pregnancy timelines, and abortion history, were also factored into the analysis.

## 3 Results

This study aimed to determine the epidemiological and serological prevalence of several infectious agents causing abortions among pregnant women in Dhamar City, Yemen. The findings are detailed in the subsequent sections, tables, and figures.

### 3.1 Places of Sample Collection

A total of 200 samples were collected from various hospitals, dispensaries, and laboratories. The highest proportion of samples was obtained from Lebanon Hospital, contributing 58 samples (29.0%), while the lowest number was collected from Al-Amwma and Al-Tafula, with 9 samples (4.5%). Additional samples were sourced as follows: Al-Dubai Labs provided 42 samples (21.0%), Al-Hayyah Hospital contributed 21 (10.5%), and Dhamar Hospital Dispensary and Alpha Labs each supplied 18 samples (9.0%). Furthermore, Mustawsaf Mihras accounted for 13 samples (6.5%), New Med Lab for 11 (5.5%), and Dar Al-Shifa Hospital for 10 (5.0%). These can be summarized as shown in Table 1.

**Table 1.** Number and percentage of samples collected from Any Hospital and Clinical in Dhamar.

| Place of the collected samples | No. | % |
|---|---|---|
| Lebanon Hospital | 58 | 29.0 |
| Al-Dubai Labs | 42 | 21.0 |
| Al-Hayyah Hospital | 21 | 10.5 |
| Alpha Labs | 18 | 9.0 |
| Dhamar Hospital Dispensary | 18 | 9.0 |
| Mustawsaf Mihras | 13 | 6.5 |
| New Med Lab | 11 | 5.5 |
| Dar Al-Shifa Hospital | 10 | 5.0 |
| Al-Amwma and Al-tafula | 9 | 4.5 |
| **Total** | **200** | **100.0** |

### 3.2 Distribution of Pregnant Women According to Causative IgG and IgM Antibodies:

Among the 200 participants, all exhibited infections by one or more of the targeted agents associated with abortions (*Toxoplasma gondii*, Rubella, CMV, and anti-phospholipid antibodies). The results revealed that positive IgM antibody tests for APL were the most frequent (20.0%), followed by *Toxoplasma gondii* (7.5%) and CMV (7.0%). Rubella demonstrated the lowest IgM seropositivity at 3.5%. Conversely, for IgG antibodies, CMV showed the highest prevalence (92.5%), followed by Rubella (87.0%) and *Toxoplasma gondii* (62.5%), while APL had the lowest IgG positivity rate at 10.0%. These results were presented in Figure 2 and Tables 2 and 3.

**Table 2.** Showing seroprevalence of CMV IgM and IgG and Rubella IgM and IgG in pregnant women.

| Tests | CMV IgM | | CMV IgG | | Rubella IgM | | Rubella IgG | |
|---|---|---|---|---|---|---|---|---|
| **Result** | No. | % | No. | % | No. | % | No. | % |
| Positive | 14 | 7.0 | 185 | 92.5 | 7 | 3.5 | 174 | 87.0 |
| Negative | 186 | 93.0 | 15 | 7.5 | 139 | 96.5 | 26 | 13.0 |
| **Total** | **200** | **100** | **200** | **100** | **200** | **100** | **200** | **100** |

**Table 3.** Showing seroprevalence of *Toxoplasma* IgM and IgG and Anti-phospholipid IgM and IgG in pregnant women.

| Tests | *Toxoplasma* IgM | | *Toxoplasma* IgG | | Anti-phospholipid IgM | | Anti-phospholipid IgG | |
|---|---|---|---|---|---|---|---|---|
| **Result** | No. | % | No. | % | No. | % | No. | % |
| Positive | 15 | 7.5 | 125 | 62.5 | 40 | 20.0 | 20 | 10.0 |
| Negative | 185 | 92.5 | 75 | 37.5 | 160 | 80.0 | 180 | 90.0 |
| **Total** | **200** | **100** | **200** | **100** | **200** | **100** | **200** | **100** |

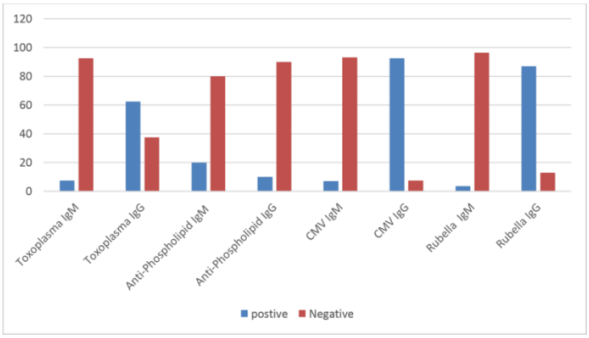



**Figure 2:** Epidemiology of pregnant women according to the type of IgG and IgM Antibodies.

### 3.3 Distribution by Geographic Location

The majority of the pregnant women resided in the rural districts of Dhamar City, accounting for 106 participants (53.0%), compared to 94 participants (47.0%) from urban areas. In the rural district, the highest IgM prevalence was observed for APL (23.6%), while the lowest was for Rubella (3.8%). Similarly, in urban areas, APL showed the highest IgM positivity (15.9%), whereas Rubella and *Toxoplasma* tied for the lowest at 3.2%. Regarding IgG antibodies, CMV was the most prevalent in both rural (93.4%) and urban (91.4%) districts. APL showed the lowest IgG prevalence in both regions, at 11.4% and 8.5%, respectively. These results were shown in Tables 4 and 5.

### 3.4 Times of the Abortion

When analyzing the number of abortions relative to IgM positivity, the highest percentage occurred among women with a single abortion (40.5%), followed by those with two (32.5%), and those with more than two (27.0%). For women with a single abortion, APL IgM was the most prevalent (27.1%), and Rubella was the least prevalent (1.2%). Among those with two abortions, APL again had the highest IgM ratio (23.1%), with Rubella being the lowest (4.6%). For women experiencing more than two abortions, *Toxoplasma* showed the highest IgM prevalence (7.4%), while CMV and Rubella both presented at 5.5%, as shown in Tables 6 and 7.

In terms of IgG seropositivity, women with a single abortion showed the highest rates for CMV and Rubella (90.2%), and the lowest for APL (14.8%). For those with two abortions, CMV IgG was highest (92.3%), and APL was lowest (7.7%). This trend continued for women with more than two abortions, where CMV IgG prevalence was 96.3% and APL was 5.6%, as shown in Tables 6 and 7.

### 3.5 Relationship Between Education and Abortion

Educational status significantly impacted the observed trends, with educated women representing 64.5% of the cases compared to 35.5% for uneducated women. Among educated women, the highest IgM positivity was for APL (18.6%), and the lowest was for Rubella (3.2%).

**Table 4.** Showing seroprevalence of CMV IgM and IgG and Rubella IgM and IgG in pregnant women according to location.

| Tests | Total | CMV IgM | | | | CMV IgG | | | | Rubella IgM | | | | Rubella IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No. | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **Urban** | 94 | 4 | 4.3 | 90 | 95.7 | 86 | 91.4 | 8 | 8.5 | 3 | 3.2 | 91 | 96.8 | 84 | 89.6 | 10 | 10.6 |
| **Rural** | 106 | 10 | 9.4 | 96 | 90.6 | 99 | 93.4 | 7 | 6.6 | 4 | 3.8 | 102 | 96.2 | 90 | 85.0 | 16 | 15.0 |

**+ = Positive – = Negative**

**Table 5.** Showing seroprevalence of *Toxoplasma* IgM and IgG and anti-phospholipid IgM and IgG in pregnant women according to location.

| Test | Total | *Toxoplasma* IgM | | | | *Toxoplasma* IgG | | | | Anti-phospholipid IgM | | | | Anti-phospholipid IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No. | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **Urban** | 94 | 3 | 3.2 | 91 | 96.8 | 70 | 74.4 | 24 | 25.5 | 15 | 15.9 | 79 | 84.0 | 8 | 8.5 | 86 | 91.5 |
| **Rural** | 106 | 12 | 11.3 | 94 | 88.7 | 55 | 51.8 | 51 | 48.1 | 25 | 23.6 | 81 | 76.4 | 12 | 11.3 | 94 | 88.7 |

**+ = Positive – = Negative**

**Table 6.** Showing seroprevalence of CMV IgM and IgG and Rubella IgM and IgG in pregnant women according to enumerated abortions.

| Tests | Total | CMV IgM | | | | CMV IgG | | | | Rubella IgM | | | | Rubella IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **Once** | 81 | 6 | 7.4 | 75 | 92.6 | 73 | 90.2 | 8 | 9.8 | 1 | 1.2 | 80 | 98.8 | 73 | 90.2 | 8 | 9.8 |
| **Twice** | 65 | 5 | 7.7 | 60 | 92.3 | 60 | 92.3 | 5 | 7.7 | 3 | 4.6 | 62 | 95.4 | 59 | 90.7 | 8 | 12.3 |
| **More than** | 54 | 3 | 5.5 | 51 | 94.4 | 52 | 96.3 | 2 | 3.8 | 3 | 5.5 | 51 | 94.4 | 44 | 81.5 | 10 | 18.5 |
| **Total** | **200** | **14** | | **186** | | **185** | | **15** | | **7** | | **193** | | **174** | | **26** | |

**+ = Positive – = Negative**

For uneducated women, APL IgM was also the highest (22.5%), while Rubella was the lowest (4.2%). A possible explanation is that educated women may possess greater awareness regarding their reproductive health, leading them to seek clinical care more frequently. For IgG antibodies, CMV was the most prevalent among both educated (91.5%) and uneducated (94.0%) women. APL showed the lowest IgG prevalence in both groups, at 10.0% and 9.8%, respectively. These results were tabulated as shown in Tables 8 and 9.

### 3.6 Relationship Between Age and Abortion

The results indicated that pregnant women aged 21–26 years experienced the highest rate of abortions (41.0%), followed by the 27–32 age group (33.0%) and the 15–20 age group (18.0%). The lowest incidence (8.0%) was recorded among women aged 33–38 years (Table 10).

Across all age groups, APL showed the highest IgM seroprevalence: 13.8% for ages 15–20; 22.0% for ages 21–26; 21.2% for ages 27–32; and 18.7% for ages 33–38. Rubella generally exhibited the lowest IgM rates across these cohorts (Tables 10 and 11). Regarding IgG tests, CMV consistently demonstrated the highest seropositivity across all age categories, peaking at 95.5% in the 27–32 age group. Conversely, APL showed the lowest IgG prevalence across all age brackets (Tables 10 and 11).

**Table 7.** Showing seroprevalence of *Toxoplasma* IgM and IgG and Anti-phospholipid IgM and IgG in pregnant women according to enumerated abortions.

| Test | Total | *Toxoplasma* IgM | | | | *Toxoplasma* IgG | | | | Anti-phospholipid IgM | | | | Anti-phospholipid IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **Once** | 81 | 6 | 7.4 | 75 | 92.6 | 57 | 70.4 | 24 | 29.6 | 22 | 27.1 | 59 | 72.9 | 12 | 14.8 | 69 | 85.1 |
| **Twice** | 65 | 5 | 7.7 | 60 | 92.3 | 37 | 57.0 | 28 | 43.0 | 15 | 23.1 | 50 | 76.9 | 5 | 7.7 | 60 | 92.3 |
| **More than** | 54 | 4 | 7.4 | 50 | 92.6 | 31 | 57.4 | 23 | 42.6 | 3 | 5.6 | 51 | 94.4 | 3 | 5.6 | 51 | 94.4 |
| **Total** | **200** | **15** | | **185** | | **125** | | **75** | | **40** | | **160** | | **20** | | **180** | |

**+ = Positive – = Negative**

**Table 8.** Showing seroprevalence of CMV IgM and IgG and Rubella IgM and IgG in pregnant women according to education.

| Test | Total | CMV IgM | | | | CMV IgG | | | | Rubella IgM | | | | Rubella IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **Educated** | 129 | 7 | 5.4 | 122 | 94.6 | 118 | 91.5 | 11 | 8.5 | 4 | 3.2 | 125 | 96.8 | 116 | 90.0 | 13 | 10.0 |
| **Uneducated** | 71 | 7 | 9.8 | 64 | 90.2 | 67 | 94.0 | 4 | 5.6 | 3 | 4.2 | 68 | 95.7 | 58 | 81.7 | 13 | 18.3 |
| **Total** | **200** | **14** | | **186** | | **185** | | **15** | | **7** | | **193** | | **174** | | **26** | |

**+ = Positive – = Negative**

**Table 9.** Showing seroprevalence of *Toxoplasma* IgM and IgG and Anti-phospholipid IgM and IgG in pregnant women according to education.

| Test | Total | *Toxoplasma* IgM | | | | *Toxoplasma* IgG | | | | Anti-phospholipid IgM | | | | Anti-phospholipid IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **Educated** | 129 | 9 | 7.0 | 120 | 93.0 | 74 | 57.4 | 55 | 42.6 | 24 | 18.6 | 105 | 81.4 | 13 | 10.0 | 116 | 90.0 |
| **Uneducated** | 71 | 6 | 8.4 | 65 | 91.6 | 51 | 71.8 | 20 | 28.2 | 16 | 22.5 | 55 | 77.5 | 7 | 9.8 | 64 | 90.1 |
| **Total** | **200** | **15** | | **185** | | **125** | | **75** | | **40** | | **160** | | **20** | | **180** | |

**+ = Positive – = Negative**

**Table 10.** Showing seroprevalence of CMV IgM and IgG and Rubella IgM and IgG in pregnant women according to age group.

| Test | Total | CMV IgM | | | | CMV IgG | | | | Rubella IgM | | | | Rubella IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **15-20** | 36 | 3 | 8.3 | 33 | 91.6 | 34 | 94.4 | 2 | 5.6 | 2 | 5.6 | 34 | 94.4 | 30 | 83.3 | 6 | 16.6 |
| **21-26** | 82 | 5 | 6.1 | 77 | 93.9 | 73 | 89.0 | 9 | 10.9 | 2 | 2.4 | 80 | 97.5 | 71 | 86.5 | 11 | 13.4 |
| **27-32** | 66 | 5 | 7.5 | 61 | 92.5 | 63 | 95.5 | 3 | 4.5 | 3 | 4.5 | 63 | 95.4 | 60 | 90.9 | 6 | 9.1 |
| **33-38** | 16 | 1 | 6.3 | 15 | 93.7 | 15 | 93.7 | 1 | 6.2 | 0 | 0 | 16 | 100 | 13 | 81.2 | 3 | 27.2 |
| **Total** | **200** | **14** | | **186** | | **185** | | **15** | | **7** | | **193** | | **174** | | **26** | |

**+ = Positive – = Negative**

**Table 11.** Showing seroprevalence of Toxoplasma IgM and IgG and anti-phospholipid IgM and IgG in pregnant women according to age group.

| Tests | Total | *Toxoplasma* IgM | | | | *Toxoplasma* IgG | | | | Anti-phospholipid IgM | | | | Anti-phospholipid IgG | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | + | | – | | + | | – | | + | | – | | + | | – | |
| | | No | % | No | % | No | % | No | % | No | % | No | % | No | % | No | % |
| **15-20** | 36 | 2 | 5.6 | 34 | 94.4 | 22 | 61.1 | 14 | 38.8 | 5 | 13.8 | 31 | 86.1 | 5 | 13.8 | 31 | 86.1 |
| **21-26** | 82 | 8 | 9.7 | 74 | 90.2 | 51 | 62.2 | 31 | 37.8 | 18 | 22.0 | 64 | 78.0 | 8 | 9.7 | 74 | 90.2 |
| **27-32** | 66 | 4 | 6.0 | 62 | 94.0 | 41 | 62.1 | 25 | 37.8 | 14 | 21.2 | 52 | 78.7 | 6 | 9.1 | 60 | 90.9 |
| **33-38** | 16 | 1 | 6.3 | 15 | 93.7 | 11 | 68.7 | 5 | 31.2 | 3 | 18.7 | 13 | 81.3 | 1 | 6.2 | 15 | 93.7 |
| **Total** | **200** | **15** | | **185** | | **125** | | **75** | | **40** | | **160** | | **20** | | **80** | |

**+ = Positive** **– = Negative**

## 4. Discussion

During pregnancy, the maternal immune system undergoes significant adaptations to tolerate the developing fetus; however, this natural immunosuppression, particularly during the first trimester, increases the mother's susceptibility to various infections [11]. Exposure to pathogens such as Rubella, CMV, and other viruses can further compromise immune defenses. Consequently, the weakened immune system struggles to protect the fetus, significantly elevating the risk of abortion [12].

In this study, comparisons of positive IgM antibodies among the targeted pathogens revealed that APL was the most frequent, followed by *Toxoplasma gondii* and CMV, with Rubella showing the lowest recent infection rate. For IgG antibodies, CMV was the most prevalent, followed by Rubella and *Toxoplasma gondii*, while APL was the least common. These findings align with reports from other regions globally, including other areas of Yemen and neighboring countries [13]. Although the CMV IgG prevalence in the present study was slightly lower than rates reported in some developed nations, it is consistent with previous findings in Hodeidah City, Yemen (98.7%) [14], as well as in Iraq (98.3%) [15], Turkey (97.2%) (Uysal *et al.*, 2012) [16], Qatar (96.5%) [17], Bahrain (100%) [18], Tunisia (96.3%) [19], Palestine (99.6%) [20], Egypt (100%) [21], Sudan (97.5%) [22], Nigeria (93.2%) [23], and Ethiopia (88.5%) [24].

Our results are also in agreement with Tamer *et al.* [25], who investigated the seroprevalence of these agents in western Turkey, finding similar trends in IgG and IgM distributions. Furthermore, Aynioglu *et al.* [4] reported comparable findings among pregnant women in northwestern Turkey. Additionally, our observation that *Toxoplasma gondii* IgG positivity (62.5%) significantly outpaced IgM positivity (7.5%) mirrors findings from Hadi *et al.* [26] in Qadisiyah province, Iraq, where IgG and IgM rates were 44% and 4%, respectively. Similar seroepidemiological patterns have been documented in Sana'a [27] and Taiz [28], Yemen, as well as in western Sudan [29]. Consistent findings were also reported by Hajipour *et al.* [30] in their cross-sectional assessment of TORCH infections.

Infections by *Toxoplasma gondii* are primarily linked to the environmental distribution of oocysts shed by cats, which can persist for extended periods under normal conditions [31]. Waterborne transmission via contaminated sources is a well-documented global issue [32]. The consumption of bottled water and improvements in the quality of tap water [33] could reduce oocyst ingestion. Regional variations in seroprevalence may stem from differences in meat consumption, cooking practices, and the underlying *Toxoplasma* burden in local livestock [8].

Viruses like CMV have evolved sophisticated mechanisms to evade the host's immune system, such as capturing host genes encoding cytokines, because these viral proteins have cellular counterparts [34], or down-regulating MHC class I and II synthesis. Additionally, some express secreted proteins that suppress inflammatory responses, such as murine IL-10 [35] and ebvIL-10 [10]. When comparing CMV and Rubella in our cohort, CMV consistently demonstrated higher rates of both IgM and IgG seropositivity. This predominance of CMV is a recognized pattern, supported by studies in Ghana [10], western Turkey [25], Ethiopia [36], Iraq , and among Syrian refugees [37]. Similar distributions were reported in Hodeidah [14] and Iran [38].

Demographically, a higher proportion of cases originated from rural districts (53.0%) compared to urban areas (47.0%). This trend is supported by Alghalibi *et al.* [14], as well as research in China [39] and Tanzania [40]. This geographic disparity may be influenced by varying social backgrounds, living environments, education levels, access to healthcare [39], alongside a greater lack of reproductive health information and differing cultural factors in rural settings [41-43].

The abortion frequencies observed here align with findings from Iraq [44] and other regional studies in Yemen [45], as well as China [39], confirming that multiple factors are associated with the acquisition of CMV [29].

Notably, education appeared to play a critical role; educated women represented a larger portion of our study sample (64.5%), likely due to higher health literacy and more frequent interaction with wider social environments [46]. This agrees with prior research in Yemen [27], though it contrasts with findings from China [39]. Studies in Taiz [28] and other international cohorts [10, 47] further highlight these disparities. Uneducated women may not visit clinics as regularly, meaning educated pregnant women are represented at a higher average in our data compared to uneducated women [48]. Educated women may also be more proactive in using protection, preventing infection transmission [49].

Finally, age distribution analysis revealed that women aged 21–26 years were the most affected (41.0%), a finding consistent with findings from Sana'a [45] and Hodeidah [14]. Similar age trends were recorded in Iran [38] and Ghana [10]. This peak in abortion rates among young women may be associated with hormonal fluctuations, particularly involving progesterone and estrogen, alongside specific immune factors present during pregnancy [46].

## 5. Conclusions

CMV emerged as the predominant causative agent among the IgG-seropositive pregnant women in Dhamar City. Overall, recent infections (indicated by positive IgM) were most frequently associated with APL, followed by *Toxoplasma gondii* and CMV, while Rubella showed the lowest incidence. Conversely, prior exposure or chronic infection (indicated by positive IgG) was highest for CMV, followed by Rubella and *Toxoplasma gondii*, with APL being the least common. The highest burden of these infections was observed among pregnant women aged 21 to 26 years. The significant differences between IgM and IgG seropositivity rates highlight varying levels of past and recent exposures among the participants. Based on these findings, we strongly recommend further investigation into the spread of these infectious agents to prevent stillbirths and abortions. Enhancing patient awareness regarding the critical importance of regular antenatal clinical visits, especially during the first trimester, is essential. Furthermore, evidence-based recommendations, such as administering the Rubella vaccine to women of childbearing age, should be actively promoted, alongside continued research into maternal infectious diseases.

## Ethical Approval

Ethical approval was obtained from Thamar University, Faculty of Applied Sciences, Biology Department, No. 1, 25-01-2020. Ethical approval has been granted to the hospitals.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Conflict of Interest

The authors declare no conflicts of interest.

## References


[1] Matthiesen, L., Kalkunte, S., Sharma, S. (2012) Multiple pregnancy failures: an immunological paradigm, *American Journal of Reproductive Immunology* **67**: 334-340.

[2] Pandey, M.K., Rani, R., Agrawal, S. (2005) An update in recurrent spontaneous abortion, *Archives of Gynecology and Obstetrics* **272**: 95-108.

[3] Ahmed, E.A.A.M. (2017) Seroprevelance of Torch among pregnant women attending maternity hospitals in Khartoum state. Ph.D. Thesis, National Ribat University, Khartoum, Sudan.

[4] Aynioglu, A., Aynioglu, O., Altunok, E.S. (2015) Seroprevalence of Toxoplasma gondii, rubella and Cytomegalovirus among pregnant females in north-western Turkey, *Acta Clinica Belgica* **70**: 321-324.

[5] Blaszkowska, J., Góralska, K. (2014) Parasites and fungi as a threat for prenatal and postnatal human development, *Annals of Parasitology* **60**: 225-234.

[6] Aminu, M., Unkels, R., Mdegela, M., Utz, B., Adaji, S., Van Den Broek, N. (2014) Causes of and factors associated with stillbirth in low-and middle-income countries: a systematic literature review, *BJOG: An International Journal of Obstetrics & Gynaecology* **121**: 141-153.

[7] Josheghani, S.B., Moniri, R., Taheri, F.B., Sadat, S., Heidarzadeh, Z. (2015) The prevalence of serum antibodies in TORCH infections during the first trimester of pregnancy in Kashan, Iran, *Iranian Journal of Neonatology* **6**: 8-12.

[8] Al-Adhroey, A.H., Mehrass, A.A.-K.O., Al-Shammakh, A.A., Ali, A.D., Akabat, M.Y., Al-Mekhlafi, H.M. (2019) Prevalence and predictors of Toxoplasma gondii infection in pregnant women from Dhamar, Yemen, *BMC Infectious Diseases* **19**: 1089.

[9] Abdulkhaliq, R.J., Mohammed, S., Abbas, A.A.-H. (2017) The role of some cytokines inwomen withRecurrent Abortion in Iraqi Women, *Pakistan Journal of Medical and Health Sciences* **11**: 496-502.

[10] Ayensu, F. (2014) Cytomegalovirus, Rubella virus and Herpes simplex-2 virus infections in pregnant women attending the Komfo Anokye Teaching Hospital for antenatal care (ANC) services. *Clinical Microbiology*, Ph.D. Thesis, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana.

[11] Mieszkin, S., Caprais, M.-P., Le Mennec, C., Le Goff, M., Edge, T., Gourmelon, M. (2013) Identification of the origin of faecal contamination in estuarine oysters using Bacteroidales and F-specific RNA bacteriophage markers, *Journal of Applied Microbiology* **115**: 897-907.

[12] Silasi, M., Cardenas, I., Kwon, J.Y., Racicot, K., Aldo, P., Mor, G. (2015) Viral infections during pregnancy, *American Journal of Reproductive Immunology* **73**: 199-213.

[13] Hama, S.A., Abdurahman, K.J. (2013) Human cytomegalovirus IgG and IgM seropositivity among pregnant women in Sulaimani city and their relations to the abortion rates, *Current Research Journal of Biological Sciences* **5**: 161-167.

[14] Alghalibi, S.M., Abdullah, Q.Y., Al-Arnoot, S., Al-Thobhani, A. (2016) Seroprevalence of cytomegalovirus among pregnant women in Hodeidah city, Yemen, *Journal of Human Virology & Retrovirology* **3**: 00106.

[15] Al-musawi, M.H.J. (2018) Cytomegalovirus Antibodies Among Pregnant Ladies at Kamal Al Samarrai Hospital in Baghdad City/Iraq, *Pakistan Journal of Biotechnology* **15**: 83-87.

[16] Uysal, A., Taner, C.E., Cüce, M., Atalay, S., Göl, B., Köse, S., Uysal, F. (2012) Cytomegalovirus and rubella seroprevalence in pregnant women in Izmir/Turkey: follow-up and results of pregnancy outcome, *Archives of Gynecology and Obstetrics* **286**: 605-608.

[17] Abu-Madi, M.A., Behnke, J.M., Dabritz, H.A. (2010) Toxoplasma gondii seropositivity and co-infection with TORCH pathogens in high-risk patients from Qatar, *The American journal of tropical medicine and hygiene* **82**: 626.

[18] AlKhawaja, S., Ismaeel, A., Botta, G., Senok, A.C. (2012) The prevalence of congenital and perinatal cytomegalovirus infections among newborns of seropositive mothers, *The Journal of Infection in Developing Countries* **6**: 410-415.

[19] Hannachi, N., Marzouk, M., Harrabi, I., Ferjani, A., Ksouri, Z., Ghannem, H., Khairi, H., Hidar, S., Boukadida, J. (2011) Seroprevalence of rubella virus, varicella zoster virus, cytomegalovirus and parvovirus B19 among pregnant women in the Sousse region, Tunisia, *Bulletin de la Societe de Pathologie Exotique (1990)* **104**: 62-67.

[20] Neirukh, T., Qaisi, A., Saleh, N., Rmaileh, A.A., Zahriyeh, E.A., Qurei, L., Dajani, F., Nusseibeh, T., Khamash, H., Baraghithi, S. (2013) Seroprevalence of Cytomegalovirus among pregnant women and hospitalized children in Palestine, *BMC Infectious Diseases* **13**: 528.

[21] Kamel, N., Metwally, L., Gomaa, N., Sayed Ahmed, W., Lotfi, M., Younis, S. (2013) Primary cytomegalovirus infection in pregnant Egyptian women confirmed by cytomegalovirus IgG avidity testing, *Medical Principles and Practice* **23**: 29-33.

[22] Khairi, S., Intisar, K., Enan, K., Ishag, M., Baraa, A., Ali, Y. (2013) Seroprevalence of cytomegalovirus infection among pregnant women at Omdurman Maternity Hospital, *Journal of Medical Laboratory and Diagnosis* **4**: 45-49.

[23] Akinbami, A.A., Rabiu, K.A., Adewunmi, A.A., Wright, K.O., Dosunmu, A.O., Adeyemo, T.A., Adediran, A., Osunkalu, V.O. (2011) Seroprevalence of cytomegalovirus antibodies amongst normal pregnant women in Nigeria, *International Journal of Women's Health* **3**: 423-428.

[24] Mamuye, Y., Nigatu, B., Bekele, D., Challa, F., Desale, A. (2015) Seroprevalence and absence of cytomegalovirus infection risk factors among pregnant women in St. Paul's Hospital Millennium Medical College, *Gynecology & Obstetrics* **5**: 299.

[25] Tamer, G.S., Dundar, D., Caliskan, E. (2009) Seroprevalence of Toxoplasma gondii, rubella and cytomegalovirus among pregnant women in western region of Turkey, *Clinical and Investigative Medicine* **32**: E43-E47.

[26] Hadi, H.S., Kadhim, R.A., Al-Mammori, R. (2016) Seroepidemiological aspects for Toxoplasma gondii infection in women of Qadisiyah province, Iraq, *International Journal of Pharmtech Research* **9**: 252-259.

[27] Al-Eryani, S.M., Al-Mekhlafi, A.M., Al-Shibani, L.A., Mahdy, M.M., Azazy, A.A. (2016) Toxoplasma gondii infection among pregnant women in Yemen: Factors associated with high seroprevalence, *The Journal of Infection in Developing Countries* **10**: 667-672.

[28] Mahdy, M.A., Alareqi, L.M., Abdul-Ghani, R., Al-Eryani, S.M., Al-Mikhlafy, A.A., Al-Mekhlafi, A.M., Alkarshy, F., Mahmud, R. (2017) A community-based survey of Toxoplasma gondii infection among pregnant women in rural areas of Taiz governorate, Yemen: the risk of waterborne transmission, *Infectious Diseases of Poverty* **6**: 26.

[29] Hamdan, H.Z., Abdelbagi, I.E., Nasser, N.M., Adam, I. (2011) Seroprevalence of cytomegalovirus and rubella among pregnant women in western Sudan, *Virology Journal* **8**: 217.

[30] Hajipour, N., Mohammady, E., Barzegar, G. (2025) Prevalence of antibodies against rubella virus, cytomegalovirus, hepatitis b, and Toxoplasma gondii in women of reproductive age prior to conception in Iran, *BMC Infectious Diseases* **25**: 1112.

[31] Teutsch, S.M., Juranek, D.D., Sulzer, A., Dubey, J., Sikes, R.K. (1979) Epidemic toxoplasmosis associated with infected cats, *New England Journal of Medicine* **300**: 695-699.

[32] Guigue, N., Léon, L., Hamane, S., Gits-Muselli, M., Le Strat, Y., Alanio, A., Bretagne, S. (2018) Continuous decline of Toxoplasma gondii seroprevalence in hospital: a 1997–2014 longitudinal study in Paris, France, *Frontiers in Microbiology* **9**: 2369.

[33] Kotenko, S.V., Saccani, S., Izotova, L.S., Mirochnitchenko, O.V., Pestka, S. (2000) Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10), *Proceedings of the National Academy of Sciences* **97**: 1695-1700.

[34] Takahashi, C., Mittler, R.S., Vella, A.T. (1999) Cutting edge: 4-1BB is a bona fide CD8 T cell survival signal, *The Journal of Immunology* **162**: 5037-5040.

[35] Hsu, D.-H., Malefyt, R.d.W., Fiorentino, D.F., Dang, M.-N., Vieira, P., Devries, J., Spits, H., Mosmann, T.R., Moore, K.W. (1990) Expression of interleukin-10 activity by Epstein-Barr virus protein BCRF1, *Science* **250**: 830-832.

[36] Tamirat, B., Hussen, S., Shimelis, T. (2017) Rubella virus infection and associated factors among pregnant women attending the antenatal

care clinics of public hospitals in Hawassa City, Southern Ethiopia: a cross-sectional study, *BMJ Open* **7**: e016824.

[37] Gürses, G., Doni, N.Y., Şimşek, Z., Aksoy, M., Hilali, N.G., Özek, B. (2024) Evaluation of T. gondii, rubella, and cytomegalovirus seroprevalences among female Syrian refugees in Sanliurfa, Turkiye, *The Journal of Infection in Developing Countries* **18**: 964-971.

[38] Khodabandehloo, M., Sharifi, P. (2020) Seroprevalence of Cytomegalovirus Antibodies by Electrochemiluminescence Method in Young Women Referred to the Clinical Laboratory, Sanandaj, Iran, *Epidemiology and Health System Journal* **7**: 92-98.

[39] Zheng, D., Li, C., Wu, T., Tang, K. (2017) Factors associated with spontaneous abortion: a cross-sectional study of Chinese populations, *Reproductive Health* **14**: 33.

[40] Rasch, V., Kipingili, R. (2009) Unsafe abortion in urban and rural Tanzania: method, provider and consequences, *Tropical Medicine & International Health* **14**: 1128-1133.

[41] Hamzehgardeshi, Z., Ahmadian, M., Rezaei, M., Shahhosseini, Z., Hamzehgardeshi, L., Golchin, N.A.H. (2025) Comparison of Reproductive Health Literacy and Related Factors in Urban and Rural Women of Reproductive Age in Iran 2023, *Preprint*, **Research Square.** https://doi.org/10.21203/rs.3.rs-5763987/v1.

[42] Abebe, M., Tebeje, T.M., Yimer, N., Temesgen, T., Melaku, G., Hareru, H.E. (2025) Epidemiology of second trimester induced abortion in Ethiopia: a systematic review and meta-analysis, *Frontiers in Global Women's Health* **6**: 1452114.

[43] Kalaf, S.H., Jameel, Z.J. (2023) Evaluation the Sero-Prevalence of cytomegalovirus infection among abortion women in Baqubah City, *Academic Science Journal* **1**: 93-108.

[44] Kadhim, B.M., Abdullhusein, H.S. (2020) A Serological Study to Diagnose the Causes of Recurrent Viral and Immune Miscarriage in Aborted Women Who Attend the Shatrah General Hospital, *International Journal of Health and Medical Sciences* **3**: 42-47.

[45] Al-Sabri, A., Al-Arnoot, S., Al-Madhagi, A., Al-Shamahy, H. (2017) Seroprevalence of cytomegalovirus among healthy blood donors in Sana'a City, Yemen, *Infectious and Non Infectious Diseases* **3**: 016.

[46] Xu, L., Wei, Q., Wu, Q., Zhong, Y., Li, Y., Xu, J., Zhu, Y. (2019) Higher β-human chorionic gonadotropin and estrogen levels during the first 6 weeks of pregnancy are associated with threatened abortion, *BioScience Trends* **13**: 245-252.

[47] Street Jr, R.L., Gordon, H.S., Ward, M.M., Krupat, E., Kravitz, R.L. (2005) Patient participation in medical consultations: why some patients are more involved than others, *Medical Care* **43**: 960-969.

[48] Jonsson, M.K., Wahren, B. (2004) Sexually transmitted herpes simplex viruses, *Scandinavian Journal of Infectious Diseases* **36**: 93-101.

[49] Robinson, E., de Valk, H., Villena, I., Le Strat, Y., Tourdjman, M. (2021) National perinatal survey demonstrates a decreasing seroprevalence of Toxoplasma gondii infection among pregnant women in France, 1995 to 2016: impact for screening policy, *Eurosurveillance* **26**: 1900710.

TUJNAS

*Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 8– 14*

**DOI: https://doi.org/10.59167/x74z1870**

Original Article

# Management of Tibial Infected Nonunion with Segmental Bone Defect by the Ilizarov Technique

**Hefzulla M H. Abdulla, Abdulrakeeb Shojaa, Shehab Hazza, Bashar A.M. AL- Mujahed***

Orthopedics and Traumatology Unit, Department of Surgery, Faculty of Medicine, Thamar University, Dhamar 87246, Yemen.

***Corresponding Author:** Bashar A.M. AL- Mujahed, Orthopedics and Traumatology Unit, Department of Surgery, Faculty of Medicine, Thamar University, Dhamar 87246, Yemen. E-mail: mujahedye14@gmail.com

Received: 8 November 2025. Revised: 22 June 2026. Accepted: 23 June 2026. Published: 29 June 2026.

## Abstract

**Background and Objective:** Managing tibial fractures complicated by extensive bone loss, soft tissue defects, and infected nonunion remains a significant orthopedic challenge. While techniques like vascularized bone grafts or bone transport exist, they often fail to address infection and nonunion simultaneously. This study evaluates the effectiveness of the Ilizarov technique in simultaneously managing these complex tibial defects. **Patients and Methods:** A retrospective analysis was conducted on 79 patients (aged 12–60 years) with tibial diaphyseal defects > 5 cm treated using the Ilizarov technique between 2004 and 2016. The cohort comprised 67 open fractures (84.8%) and 12 closed fractures with postoperative infections (15.2%). Etiologies included gunshot injuries (n=53, 67.1%), traffic accidents (n=21, 26.6%), and falls or osteomyelitis (n=5, 6.3%). In 25 cases (31.6%) with adequate soft tissue and no active infection, 2.5–3 mm flexible intramedullary K-wires were used to guide the transported segment. The remaining 54 patients (69%) with active infection or poor soft tissue coverage were managed exclusively with external fixation. **Results:** Patients had undergone a mean of 2.8 prior failed surgical procedures (range: 1–16). The mean bone defect length was 9.3 cm (range: 5–18 cm). Infection eradication and bone union were successfully achieved in all 79 cases. The mean external fixation index was 1.3 months/cm. Bone results were excellent in 71 patients (89.9%), good in 5 (6.3%), fair in 2 (3.8%), and poor in 1 (1.3%). Functional results were excellent in 46 patients (58.2%), good in 28 (35.4%), and fair in 2 (3.8%), with no poor or failed outcomes reported. Skin invagination at the gap site occurred in 34 patients (43%), requiring surgical adjustment in 27 (34.2%); notably, such adjustments were unnecessary when intramedullary flexible K-nails were utilized. **Conclusion:** The Ilizarov technique is a highly effective solution for challenging tibial defects, allowing for the simultaneous resolution of bone loss, infection, nonunion, and soft tissue compromises. Furthermore, the adjunctive use of flexible intramedullary K-nails mitigates the need for subsequent docking site adjustments.

**Keywords:** Ilizarov Technique; Tibial Bone Defect; Nonunion; Tibial Bone Lengthening

## 1. Introduction

Tibial fractures complicated by extensive bone and soft tissue loss, or secondary bone defects resulting from infected nonunions, remain formidable challenges for orthopedic surgeons and are associated with significant long-term morbidity. While acute shortening of more than 4 cm is sometimes utilized, this approach can induce tortuous vasculature, resulting in a low-flow state with severe detrimental consequences. Furthermore, acute compression of open soft-tissue wounds can lead to tissue bunching, devascularization, significant edema, and an elevated risk of subsequent necrosis and infection [1-4].

Historically, the complexity of managing segmental long-bone defects often necessitated amputation; however, the clinical emphasis has progressively shifted toward limb salvage. As Keating *et al.* [5] noted, limb reconstruction in the presence of massive bone loss is technically demanding, time-intensive, and physically and psychologically exhausting for the patient, often yielding unpredictable outcomes. Successful reconstruction fundamentally requires an infection-free environment coupled with stable, reliable fixation.

Currently, a variety of surgical strategies are utilized to address these defects, including fibular grafts [6], free vascularized iliac crest grafts [7], tibiofibular synostosis [8], the Masquelet induced membrane technique [9], and distraction osteogenesis via bone transport [10-15]. The sheer diversity of these reconstructive methods underscores the inherent difficulty of achieving osseous union across extensive defect gaps. Consequently, the present study aims to evaluate the clinical efficacy of the Ilizarov technique in managing tibial nonunions complicated by segmental bone loss.

## 2. Patients and Method

This study utilizes a prospective cohort design reviewing patients in a multicenter study from 2004 to 2016, which was carried out at the orthopedic unit of Al Wahda Educational Hospital, Maabar, Thamar University, First Orthopedic and Plastic Hospital, Sana'a, and General Military Hospital, Sana'a, and informed consent was obtained from all patients. Out of 79 patients, 75 (94.9%) were male, and 4 (5.1%) were female.

The procedure is ethically sound when performed with appropriate medical and ethical standards. The institutional review board approval has been obtained from the ethics committee of each hospital included in

our research. A written inform consent has been gotten from patients or their families.

Patients from either gender were included with an age range from 12 to 60 years, if they had a tibial bone defect of more than 5cm due to bone loss after trauma or debridement. Patients lost to follow-up, and those demanding premature fixator removal were excluded. The defect developed in 53 (67.1%) patients as a result of gunshot injury and in 21 (26.6%) patients as a result of traffic accidents, and 5 (6.3%) as a result of falling from height and osteomyelitis.

We defined primary bone loss as bone missing immediately at the time of injury and secondary bone loss as the bone defect resulting from subsequent surgical debridement. In our cohort, primary bone loss was present in 67.09% of patients, while secondary bone loss occurred in 32.91% of patients. Fractures were located in the distal third of the tibia in 41, 8 in the proximal third, and 30 were located in the middle third; thirty were on the right, and forty-nine were on the left lower extremity.

At presentation, a detailed history was obtained for the initial injury and previous surgical interventions. The patients were examined for the condition of soft tissue, presence of shortening, neurovascular deficiency, and active infection, if any, and the function of relevant joints was documented. Radiological evaluation was done to determine the fracture pattern, plane of deformity, alignment, and to look for signs of osteomyelitis.

At the time of surgery, in cases presenting with infected nonunion, the definitive bone defect size was measured intraoperatively after radical debridement and resection of all infected and non-viable bone segments. 36 (45.6%) tibial nonunions with bone defect or shortening were diagnosed as infected. Resection of all devitalized or sclerotic bone and fibrous tissue was removed from the site, leaving healthy, viable bone ends. The medulla was opened, and the defect was measured. In those cases, with a big defect (more than 10 cm), we used 2 flexible intramedullary nails or 2.5 mm K wires to guide the transported segment, and we didn't apply an intramedullary nail in cases of infection or poor soft tissue coverage. In case of bad skin cover, muscle-cutaneous flab is used to cover the anterior aspect of the leg. Soft tissue coverage was achieved using local rotational flaps, free microvascular flaps, or split-thickness skin grafting, depending on the wound requirements.

Data were analyzed for mean age, mean bone defects, mean follow-up, bone union, bone results, functional results, complications per patient, external fixation time, and external fixation index were recorded and statistically analyzed using weighted means based on the sample size in the study by SPSS software. Continuous variables were analyzed using the Student's t-test or Mann-Whitney U test, depending on normality. Categorical variables were compared using the Chi-square test or Fisher's exact test. A statistically significant difference was set at $P < 0.05$.

## 3. Management Protocol & Operative Technique

After completing pre-operative investigations, we performed a careful preparation that included assessment of the size of the bone defect, identification of the infecting microorganisms, and evaluation of the condition of both the adjacent soft tissue and the neighboring joints, as shown in Figure 1.

In infected cases, according to microbiological sensitivity tests, an appropriate antibiotic was selected and used until there were no clinical signs of acute infection (redness, swelling, or pus discharge). The patient is placed in the supine position on a radiolucent operating table. The injured leg is in position. Anesthesia is given. The site of the infected nonunion was exposed through a standard approach. Samples for culture and sensitivity were taken promptly to the laboratory and processed immediately.

Resection of all devitalized or sclerotic bone and fibrous tissue was removed from the site, leaving healthy, viable bone ends; the medulla was opened (Figure 2). If the fibula was intact, a resection was carried out between the fracture and osteotomy site, so that the lengthening could be done.

To decrease the operation time, an Ilizarov frame was constructed preoperatively according to the fracture and cortico- or osteotomy sites. For bone transport, the simplest, most commonly used configuration consists of three rings, with one proximal to the corticotomy site, one distal to the zone of the original defect, and one in the transport segment. Most of the time, we used an image intensifier to achieve reduction and near-normal alignment of the fracture. For each ring, a minimum of 2 wires was used.

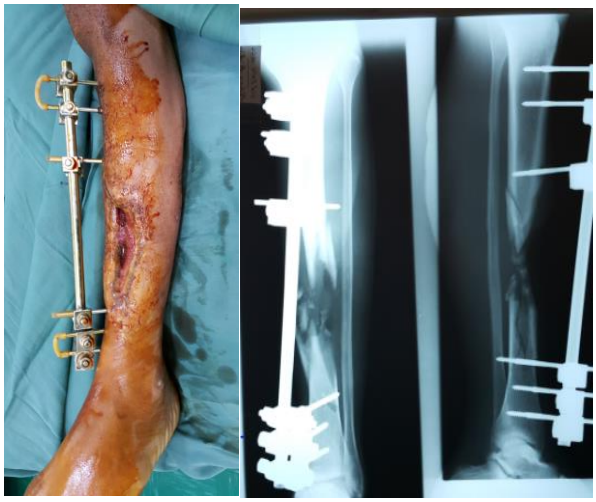


**Figure 1:** Preoperative clinical presentation and radiograph of a patient with a left tibial bone defect resulting from a gunshot injury, initially stabilized with a temporary external fixator.

**Figure 2:** Intraoperative clinical view following the removal of the temporary external fixator and the radical debridement of infected, devitalized bone and soft tissue.

While inserting the wires, they were first gently pushed up to the bone through the skin and then drilled with a power drill. As soon as they come out through the other cortex, they are gently hammered to get out to the other side. Muscles were at their maximum length while inserting the pins, and all the wires were passed through safe zones. All the wires were tensionized before fastening to rings, either with a wire tensioner or manually with spanners at both ends simultaneously on plain wires and only the opposite end on olive tip wires. A corticotomy is then performed using a Gigli saw from the posterior aspect and an osteotome from the medial and lateral aspects. At the end of the surgical procedure, the hemovac drain was routinely used. In cases with bad skin cover, muscle-cutaneous flab is used to cover the anterior aspect of the leg (was done in 2 cases).

In active infected cases, treatment was performed in two stages. In the first stage, an accurate *debridement* of the infected nonunion site, with bone end transverse resection until healthy bone was observed, and complete excision of the infected and necrotic soft tissues was performed, and cultures were taken. At this time, the Ilizarov apparatus was applied to the leg (Figure 3). We combined K-wires with Schanz pins to make the fixator more rigid and stable at the proximal and distal ends of the tibia, with only K-wires in the segment for transport to decrease deep lesions in soft tissues during transport. The assembly must be extended to the hind- and forefoot in patients in whom a loss of tissue in the distal tibia requires extensive resection, and the length of the distal tibial fragment is only a few centimeters in length, and if the bone defect is more than 5 cm, to prevent equins deformity.

**Figure 3:** Postoperative radiograph and clinical photograph demonstrating the application of the Ilizarov circular frame, utilizing flexible intramedullary K-wires as guides for the transported bone segment.

By the end of the distraction phase, the part of the device that is in the forefoot was removed. When the infection had subsided and the wound was healed, the second step consisted of performing a corticotomy at the proximal or distal tibial metaphysis according to the resection site, distal or proximal (Figure 4).

**Figure 4:** Radiographic and clinical appearance after the distraction osteogenesis phase, illustrating the transit of the bone segment and the developing regenerated bone.

In the noninfected cases, *debridement* and corticotomy were performed in one step with two 2.5 mm K-wires intramedullary to guide the segment of transport. After a delay of 5 to 10 days, distraction is begun by turning the nuts on the metal rods 0.5mm, two times per day, for a total distraction of 1mm per day. Knee and ankle kinesitherapy and muscle strengthening are started immediately, and standing and walking start a few days after the operation.

Patients were initially followed every 2 weeks till the lengthening was completed, and thereafter monthly till the consolidation of the regenerate. Once bone ends were in near apposition, a docking procedure was performed as necessary. Compression over the docking site was maintained by ¼ turn 2x/week.

In our study, the bone ends at the defect site were cut in a transverse manner at the primary debridement or at the time of adjustment or docking site procedure, to make a maximal contact area when the transported segment comes into contact with the other fragment, and to avoid bone graft as shown in Figure 5.

**Figure 5:** Final clinical and radiographic outcomes demonstrating complete consolidation at the docking site, mature regenerated bone calcification, and restoration of limb alignment following the removal of the Ilizarov frame.

Once the distraction is stopped, the frame remains attached to allow the new bone to harden. Dynamization was performed in all cases before the removal of the frame. This time is usually one month per cm of new bone formed. After this time the fixator removal was performed either in the Outpatient Department (OPD) without anesthesia for cooperative patients with stable clinical signs, or in the Operating Room (OR) under General Anesthesia (GA). GA was strictly reserved for pediatric patients, highly anxious individuals, or cases requiring significant pin tract debridement.

## 4. Results

The clinical outcomes of the 79 patients were evaluated following a mean of 2.8 prior failed surgical procedures (range: 0–16). The mean length of the tibial bone defect was 9.3 cm (range: 5–16 cm). Infection was successfully eradicated in all 37 cases presenting with infected nonunion. Furthermore, definitive bone union was achieved in all 79 patients. The mean external fixation index was 1.3 months/cm. Statistical analysis revealed no significant difference in the external fixation index between infected and non-infected cases (P = 0.630), as detailed in Table 1.

**Table 1.** Association Between Infected and Non-Infected Cases Regarding External Fixation Index

| Status | Mean Time for Frame Removal (Days) | Standard Deviation (SD) | P-value |
|---|---|---|---|
| **Infected** | 347 | 85.5 | 0.630 |
| **Non-Infected** | 359 | 124.9 | |

Note to typesetter: The P-value applies to the comparison between the two groups

Outcomes were classified using the Association for the Study and Application of the Method of Ilizarov (ASAMI) scoring system [16]. Bone results were categorized as excellent in 68 patients (86.1%), good in 6 (7.6%), fair in 1 (1.3%), and poor in 4 (5.1%) (Table 2). Functional results were excellent in 46 patients (58.2%), good in 28 (35.4%), and fair in 5 (6.3%). Notably, there were no patients with poor functional results or total failure (Table 3).

All the patients had achieved consolidation of the regenerate at the distracted site before union at the compression site, even those whom acute compression at the fracture site, followed by lengthening.

**Table 2.** Evaluation of the bony results using the ASAMI scoring system (n = 79) [16].

| Bone results | No. of patients | % | Criteria |
|---|---|---|---|
| Excellent | 68 | 86.08 | Union, no infection, deformity < 7°, limb length discrepancy (LLD) < 2.5 cm |
| Good | 6 | 7.59 | Union plus any two of the following: absence of infection, deformity < 7°, LLD < 2.5 cm. |
| Fair | 1 | 1.27 | Union plus any one of the following: absence of infection, deformity < 7°, LLD < 2.5 cm. |
| Poor | 4 | 5.06 | Nonunion/refracture/union plus infection plus deformity > 7° plus LLD > 2.5 cm |

**Table 3.** Evaluation of the functional results using the ASAMI scoring system (n = 79)

| Functional results | No. of patients | % | Criteria |
|---|---|---|---|
| Excellent | 46 | 58.23 | Active, no limp, minimum stiffness (loss of < 15° knee extension/< 15° ankle dorsiflexion), no reflex sympathetic dystrophy (RSD), insignificant pain. |
| Good | 28 | 35.44 | Active, with one or two of the following: limp, stiffness, RSD, significant pain |
| Fair | 5 | 6.33 | Active, with three or all of the following: limp, stiffness, RSD, significant pain |
| Poor | 0 | 0 | Inactive (unemployment or inability to return to daily activities because of injury) |
| Failure | 0 | 0 | Amputation |

There was no relation between the infected and non-infected cases for the external fixator index (P > 0.05).

Complications were divided into problems, obstacles, and true complications according to a system proposed by Paley [17] (Table 4). Problems are difficulties encountered while the fixator is still in place (distraction and consolidation period) and are resolved without operative treatment. Obstacles are complications encountered while the fixator is in place and are resolved by operative treatment. True complications are those that arise after the removal of the fixator and resolved operative treatment.

**Table 4**. Complications Encountered During and After Treatment.

| Problems | Obstacles | True Complications |
|---|---|---|
| 1. Pain during the distraction period 41/79<br>2. Infection at the regenerate site 2/79<br>3. Pin tract infection – 79/79<br>4. Infected at the regenerate site - 2/79<br>5. Knee Contracture – 1/79<br>6. Angulation of fragments – 5/79 | 1. Insufficient regeneration – 2/79<br>2. Premature Ossification of regenerate bone – 4/79<br>3. Skin invagination over gap site – 34/79<br>4. Adjustment was done in 27/79.<br>5. Bone graft -1/79 | 1. Equines deformity – 2/79<br>2. Refracture – 3/79 |

Premature ossification of the corticotomy site occurred in four patients. These were all treated by repeat corticotomy. Skin invagination over the gap site was seen in thirty-four (43%) patients, all of whom had big bone defects. All were resolved by a docking procedure, which necessitated removal of fibrous tissue from the gap site and adjustment of the bone ends if necessary.

A bone graft was performed in one case. Adjustment was done in 27 (34.2%) patients. Refracture at the docking site, which occurred after the removal of the fixator, was noted in 3 (3.8%) cases, two of which were managed by cast for three months, and the third was managed by reapplication of the Ilizarov apparatus after refreshment of the fracture site, and it was removed after 5 months. Achilles tendon lengthening for equinus deformity was performed in one patient. Pin tract infection was the most common complication. All the patients had a feeling of pain during the distraction period and required oral analgesics.

A total of 41 patients who required lengthening of more than 5 cm complained of pain during distraction, especially at the knee joint and at the site of the K-wires near the docking site. This was treated with analgesia (acetaminophen-codeine combination) as necessary or to divide distraction into times or four times. No neurovascular problems were caused by either intraoperative pin insertion. No patient developed compartment syndrome. Invagination was noted in 34 (43%) cases, bone graft was performed in one case. Adjustment was done in 27 (34.2%) patients.

The Ilizarov External Fixator (IEF) provides sufficient stability and beneficial cyclical axial loading and minimizes shear forces, all of which are important for fracture healing. The axial deviation during bone transport occurred in 27 patients and frame adjustments for modification were performed in these patients.

## 5. Discussion

Most authors reported that although the autogenous bone grafts, allografts, bone graft substitutes, and vascularized fibular bone grafts are effective in the treatment of selected long bone nonunions with defects, these techniques have their limitations. Tibial nonunion with a bone defect is usually not an isolated problem. It may be associated with infection, shortening, soft tissue dystrophy, disuse osteoporosis, equinus deformity, angular deformity, and neurovascular damage. Joshi and Kostakis [18] and Li *et al.,* [19] reported that a conventional autogenous bone grafting has limitations for the treatment of large defects, the time-period for graft incorporation is prolonged, and the quantity of available autogenous graft is limited, as well as producing significant donor site morbidity, late stress fracture, and nonunion at the graft site, blood loss, and paraesthesia [18-21], superficial infection [20, 22-24], hematoma [18, 25], poor cosmesis [20, 26] and most commonly, acute or chronic donor site pain [18-20, 22-24, 26-30]. Moreover, Toh and Jupiter [31] and Corona *et al.* [32] believed that in an infected defect, scarring of the soft-tissue envelope compromises revascularization of the graft.

Also, Ziran *et al.* [33] supported that the use of AlloMatrix/demineralized bone matrix as an alternative for autogenous bone graft in the treatment of nonunions resulted in an unacceptably high rate of complications, especially if there is a large volumetric defect or a history of any prior contamination of the tissue bed.

Most authors believe that a free vascularized fibula grafting is technically challenging and confers its own set of inherent risks and potential complications. Adequate hypertrophy of the incorporated fibular graft may take several years, and prolonged bracing, problems with union, as well as fatigue/stress fracture are common, which range between 10% and 25% [34-39]. Furthermore, donor site morbidity has been well documented, and up to 10% of patients may subsequently develop ankle pain, instability, and/or progressive valgus deformity if fibula harvest is not performed with proper technique [34, 40, 41].

In our study, the mean external fixation index was 1.3 months/cm, and the complications per patient were 2.6. The first data were better than the average data that were recorded in the literature, but the complication rate per patient was higher. Excellent rate in bone results was 86.1%, and good 7.6%, and excellent rate in functional results was 58.2%, meanwhile, good 35.4% in our study; these data were also better than the average data that was recorded in the literature. This current study reports an average defect of 9.4 centimeters (5 – 17 cm) as listed in Table 5.

Lavini *et al.* [42] performed application of the technique of acute compression and lengthening was performed only in two patients, and both patients had developed pus formation at the site of compression, which we explain as due to soft tissue stacking. So, it was limited to the extent of bone defects. Although acute compression beyond 4-5 cm could cause overmuch soft tissue stacking and arteriolar occlusion, which could affect new bone formation and the healing of fracture ends. Magadum *et al.* [43] represented in their literature, a huge bone defects (mean 10 cm, maximum 17 cm) were treated successfully by acute compression and distraction.

Bone transport could achieve good results in the treatment of both small and massive bone defects [44-48]. So, bone transport has no limit for the extent of bone defects, and it gives the surgeon the confidence to remove as much bone as necessary to ensure complete removal of pathologic bone. We achieved not only filling the bone defect and nonunion, but also the eradication of infection in all 37 cases. Green [49]

proved that, as Ilizarov is often quoted, "osteomyelitis burns in the flames of the regenerate". Sveshnikov *et al.* [50] found that the vascularity is noted to increase in the vicinity of the corticotomy site. This neovascularization promotes healing and also helps eradicate infection. Green *et al.* [51] conducted that wound breakdown, invagination, pin-track infection, skin invagination, and axial deviation were common complications in the course of bone transport. Soft tissue interposition also prevents compression and new bone formation at the docking site.

**Table 5.** Comparison of Clinical Outcomes in Established Bone-Transport Studies.

| Reference | Number of patients | Mean of bone loss (cm) | External fixator index (months/cm) | Complication rate per patient |
|---|---|---|---|---|
| Sen *et al.* [52] | 17 | 5.6 | 1.4 | 1.2 |
| Atesalp *et al.* [53] | 43 | 9.7 | 1.4 | 1.1 |
| Cattaneo *et al.* [54] | 28 | 4.0 | 2.2 | 0.6 |
| Green *et al.* [51] | 17 | 5.1 | 1.9 | 3.5 |
| Zorn & Iii [55] | 21 | 6.5 | 2.6 | 1,4 |
| Biermann *et al.* [10] | 25 | 4.1 | 2.1 | 2.1 |
| Saleh & Rees [56] | 8 | 6.5 | 2.5 | 2.2 |
| Dendrinos *et al.* [57] | 28 | 6.0 | 1.7 | 2.5 |
| Polyzois *et al.* [58] | 42 | 6.0 | 1.6 | 1.4 |
| Song *et al.* [11] | 27 | 8.3 | 1.0 | 0.5 |
| Paley & Maar [15] | 19 | 10.0 | 1.6 | 2.9 |
| Yin *et al.* [59] | 66 | 6.27 | 1.38 | 1.08 |
| **Present Study** | **79** | **9.3** | **1.3** | **2.6** |

Our study adopts regular adjustment and skin elevation at the docking site, and we have noted shortening of the time for healing at the docking site. The external fixator index was 1.3 cm/month. During adjustment, elevation of the skin due to invagination and docking site procedure, we had noted not only an increase in the vascularity at the bone ends, but also the medullary canal at both ends became close, and the shape becomes difference comparison with the initial shape after resection at the first operation. We think that after the last procedures, we refresh both ends and perform multiple drilling, opening the medullary canal if possible, and continue compression, and we achieve healing in a good position and faster than without these procedures. We avoid the necessity for a bone graft. The only case where a bone graft was done was due to healing at the docking site with a diameter less than half the diameter of the bone after prematurity of the regenerate.

If the contact area at the docking site is small, bone grafting will be necessary [11, 56, 60-63]. Several studies have reported that 80 to 100% Song *et al.* [11], Schultz *et al.* [64], 10% Thirumal and Shong [60], and 50% Maini *et al.* [61] of cases have required bone grafting at the docking site. The leading edge of the transported segment is relatively avascular. Green *et al.* [51] and Paley *et al.* [15] reported that this can delay union, unless the sclerotic end is trimmed, and 50% of patients reportedly undergo debridement of the leading edge of the transported segment.

In our study, most of the bone defects were due to high-energy trauma 93.7% (gunshot and RTA 67.1% and 26.6% respectively), open fractures were 84.8%, the distal third was involved 51.9%, the middle 38.05 whereas the proximal third was involved only in 10.1% only and this fact is due to poor soft tissue at the distal part of the leg.

Ilizarov technique studies reported that most complications were related to the docking site, such as nonunion, delayed union, malalignment, low cross-sectional area, and soft-tissue invagination [15, 16, 51, 55, 65-67]. Despite the many complications, the Ilizarov method has shown its versatility and reliability in the treatment of cases with problematic bone defects, infection, shortening, nonunions, complex trauma, and deformities.

The initial treatment by AO external fixation was in 59 (74.7 %) cases, to stabilize the injured segment of the limb with minimal additional trauma, avoiding introducing foreign bodies in the damage zone. Moreover, one-planer of the external fixation practically doesn't restrict surgical access to an injured limb and doesn't disturb further closure of skin defects. Simplicity of the application, less traumatic methods of external fixation, and relatively short duration of the operation are significant for patients with multiple traumas and in cases of mass hospitalization.

Only primary fixation by the Ilizarov technique was applied in 8 (10.1%) cases. 12 (15.2%) cases as a sequel of infection after open reduction and internal fixation by plates or due to osteomyelitis. The need for distraction osteogenesis due to bone defects was most frequent in young adults (21–40 years).

Ilizarov bone transport is an effective salvage tool in obtaining union in patients with an infected nonunion and as a primary tool in patients with large segmental bone loss due to trauma. The lengthy treatment time and considerable number of complications must be fully understood by both the surgeon and the patient prior to undertaking this complex treatment process.

## 6. Conclusion

The Ilizarov technique remains a highly effective and versatile modality for managing complex tibial defects. It uniquely enables the simultaneous resolution of massive segmental bone loss, active infection, nonunion, and compromised soft tissue envelopes. Furthermore, the adjunctive use of flexible intramedullary K-wires significantly mitigates the need for subsequent surgical adjustments at the docking site. Crucially, the empirical outcomes of this study demonstrate that the initial presence of active infection does not adversely affect the external fixation index or the overall success of bone transport when utilizing this protocol.

## Ethical Approval

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the Medical Ethics Committee of the Faculty of Medicine, Thamar University, Yemen (Approval No.: TUMEC-223), level 1.

## Patient Consent

Written informed consent was obtained from all patients prior to their enrollment and surgical intervention. Because the study cohort included adolescent participants (ages ranging from 12 to 60 years), written informed consent for patients under the age of 18 was strictly obtained from their parents or legal guardians. Patient confidentiality and data anonymity were maintained throughout the entirety of the study.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Funding

This work was not funded by a particular grant agency, whether in the public, commercial, or not-for-profit sector.

## Conflict of Interest

The authors declare no conflicts of interest.

## References

[1] Kocialkowski, A., Marsh, D.R., Shackley, D.C. (1998) Closure of the skin defect overlying infected non-union by skin traction, *British Journal of Plastic Surgery* **51**: 307-310.

[2] Lenoble, E., Lewertowski, J.M., Goutallier, D. (1995) Reconstruction of Compound Tibial and Soft Tissue Loss Using a Traction Histogenesis Technique, *Journal of Trauma and Acute Care Surgery* **39**: 356-360.

[3] Kesemenli, C., Subasi, M., Kirkgoz, T., Kapukaya, A., Arslan, H. (2001) Treatment of Traumatic Bone Defects by Bone Transport, *Acta Orthopaedica Belgica* **67**: 380-386.

[4] Mahaluxmivala, J., Nadarajah, R., Allen, P.W., Hill, R.A. (2005) Ilizarov external fixator: acute shortening and lengthening versus bone transport in the management of tibial non-unions, *Injury* **36**: 662-668.

[5] Keating, J.F., Simpson, A.H.R.W., Robinson, C.M. (2005) The management of fractures with bone loss, *The Journal of Bone & Joint Surgery British Volume* **87-B**: 142-150.

[6] El-Gammal, T.A., Shiha, A.E., El-Deen, M.A., El-Sayed, A., Kotb, M.M., Addosooki, A.I., Ragheb, Y.F., Saleh, W.R. (2008) Management of traumatic tibial defects using free vascularized fibula or Ilizarov bone transport: A comparative study, *Microsurgery* **28**: 339-346.

[7] Tonoli, C., Bechara, A.H.S., Rossanez, R., Belangero, W.D., Livani, B. (2013) Use of the Vascularized Iliac-Crest Flap in Musculoskeletal Lesions, *BioMed Research International* **2013**: 237146.

[8] Ebraheim, N.A., Haman, S.P., Sabry, F.F., Emara, K. (2005) Tibiofibular synostosis procedure in the management of complex tibia fractures, *American Journal of Orthopedics (Belle Mead, N.J.)* **34**: 493-497.

[9] Masquelet, A.C. (2003) Muscle reconstruction in reconstructive surgery: soft tissue repair and long bone reconstruction, *Langenbeck's Archives of Surgery* **388**: 344-346.

[10] Biermann, J., Prokuski, L., Marsh, J. (1994) Chronic Infected Tibia1 Nonunions With Bone Loss Conventional Techniques Versus Bone Transport, *Clinical Orthopaedics & Related Research* **301**: 139-146.

[11] Song, H.R., Cho, S.H., Koo, K.H., Jeong, S.T., Park, Y.J., Ko, J.H. (1998) Tibial bone defects treated by internal bone transport using the Ilizarov method, *International Orthopaedics* **22**: 293-297.

[12] Rigal, S., Tripon, P. (2005) Traitement des pertes de substance osseuse traumatiques des diaphyses par transport osseux segmentaire, *Annales Orthopédiques de l'Ouest* **37**: 163-5.

[13] Ilizarov, S., Katz, H., Freudigman, P., Tracey Watson, J., Weitzman, A., Robert Rozbruch, S. (2006) Simultaneous Treatment of Tibial Bone and Soft-tissue Defects With the Ilizarov Method, *Journal of Orthopaedic Trauma* **20**: 194-202.

[14] Trigui, M., Ayadi, K., Ellouze, Z., Gdoura, F., Zribi, M., Keskes, H. (2008) Treatment of bone loss in limbs by bone transport, *Revue de chirurgie orthopedique et reparatrice de l'appareil moteur* **94**: 628-634.

[15] Paley, D., Maar, D.C. (2000) Ilizarov bone transport treatment for tibial defects, *Journal of Orthopaedic Trauma* **14**: 76-85.

[16] Paley, D., Catagni, M.A., Argnani, F., Villa, A., Bijnedetti, G.B., Cattaneo, R. (1989) Ilizarov Treatment of Tibial Nonunions With Bone Loss, *Clinical Orthopaedics and Related Research* **241**: 146-165.

[17] Paley, D. (1990) Problems, obstacles, and complications of limb lengthening by the Ilizarov technique, *Clinical Orthopaedics and Related Research* **250**: 81-104.

[18] Joshi, A., Kostakis, G.C. (2004) An investigation of post-operative morbidity following iliac crest graft harvesting, *British Dental Journal* **196**: 167-171.

[19] Li, Y., Shen, S., Xiao, Q., Wang, G., Yang, H., Zhao, H., Shu, B., Zhuo, N. (2020) Efficacy comparison of double-level and single-level bone transport with Orthofix fixator for treatment of tibia fracture with massive bone defects, *International Orthopaedics* **44**: 957-963.

[20] Schnee, C.L., Freese, A., Weil, R.J., Marcotte, P.J. (1997) Analysis of Harvest Morbidity and Radiographic Outcome Using Autograft for Anterior Cervical Fusion, *Spine* **22**: 2222-2227.

[21] Nkenke, E., Weisbach, V., Winckler, E., Kessler, P., Schultze-Mosgau, S., Wiltfang, J., Neukam, F.W. (2004) Morbidity of harvesting of bone grafts from the iliac crest for preprosthetic augmentation procedures: A prospective study, *International Journal of Oral and Maxillofacial Surgery* **33**: 157-163.

[22] Skaggs, D.L., Samuelson, M.A., Hale, J.M., Kay, R.M., Tolo, V.T. (2000) Complications of Posterior Iliac Crest Bone Grafting in Spine Surgery in Children, *Spine* **25**: 2400-2402.

[23] Swan, M.C., Goodacre, T.E.E. (2006) Morbidity at the iliac crest donor site following bone grafting of the cleft alveolus, *British Journal of Oral and Maxillofacial Surgery* **44**: 129-133.

[24] Robertson, P.A., Wray, A.C. (2001) Natural History of Posterior Iliac Crest Bone Graft Donation for Spinal Surgery: A Prospective Analysis of Morbidity, *Spine* **26**: 1473-1476.

[25] Westrich, G.H., Geller, D.S., O'Malley, M.J., Deland, J.T., Helfet, D.L. (2001) Anterior Iliac Crest Bone Graft Harvesting Using the Corticocancellous Reamer System, *Journal of Orthopaedic Trauma* **15**: 500-506.

[26] Hill, N.M., Geoffrey Horne, J., Devane, P.A. (1999) Donor Site Morbidity in the Iliac Crest Bone Graft, *Australian and New Zealand Journal of Surgery* **69**: 726-728.

[27] Silber, J.S., Anderson, D.G., Daffner, S.D., Brislin, B.T., Leland, J.M., Hilibrand, A.S., Vaccaro, A.R., Albert, T.J. (2003) Donor Site Morbidity After Anterior Iliac Crest Bone Harvest for Single-Level Anterior Cervical Discectomy and Fusion, *Spine* **28**: 134-139.

[28] Sasso, R.C., LeHuec, J.C., Shaffrey, C., The Spine Interbody Research Group (2005) Iliac Crest Bone Graft Donor Site Pain After Anterior Lumbar Interbody Fusion: A Prospective Patient Satisfaction Outcome Assessment, *Clinical Spine Surgery* **18**: S77-S81.

[29] Kager, A.N., Marks, M., Bastrom, T., Newton, P.O. (2006) Morbidity of Iliac Crest Bone Graft Harvesting in Adolescent Deformity Surgery, *Journal of Pediatric Orthopaedics* **26**: 132-134.

[30] Heary, R.F., Schlenk, R.P., Sacchieri, T.A., Barone, D., Brotea, C. (2002) Persistent iliac crest donor site pain: independent outcome assessment, *Neurosurgery* **50**: 510-517.

[31] Toh, C.L., Jupiter, J.B. (1995) The Infected Nonunion of the Tibia, *Clinical Orthopaedics and Related Research* **315**: 176-191.

[32] Corona, P.S., Pujol, O., Vicente, M., Ricou, E., de Albert, M., Maestre Cano, D., Salcedo Cánovas, C., Martínez Ros, J. (2022) Outcomes of two circular external fixation systems in the definitive treatment of acute tibial fracture related infections, *Injury* **53**: 3438-3445.

[33] Ziran, B.H., Smith, W.R., Morgan, S.J. (2007) Use of Calcium-Based Demineralized Bone Matrix/Allograft for Nonunions and Posttraumatic Reconstruction of the Appendicular Skeleton: Preliminary Results and Complications, *Journal of Trauma and Acute Care Surgery* **63**: 1324-1328.

[34] Hariri, A., Mascard, E., Atlan, F., Germain, M.A., Heming, N., Dubousset, J.F., Wicart, P. (2010) Free vascularised fibular graft for reconstruction of defects of the lower limb after resection of tumour, *The Journal of Bone & Joint Surgery British Volume* **92-B**: 1574-1579.

[35] Wood, M.B. (2007) Free Vascularized Fibular Grafting—25 Years' Experience: Tips, Techniques, and Pearls, *Orthopedic Clinics of North America* **38**: 1-12.

[36] Chen, C.M., Disa, J.J., Lee, H.-Y., Mehrara, B.J., Hu, Q.-Y., Nathan, S., Boland, P., Healey, J., Cordeiro, P.G. (2007) Reconstruction of Extremity Long Bone Defects after Sarcoma Resection with Vascularized Fibula Flaps: A 10-Year Review, *Plastic and Reconstructive Surgery* **119**: 915-924.

[37] McClure, P.K., Abouei, M., Conway, J.D. (2021) Reconstructive Options for Tibial Bone Defects, *JAAOS - Journal of the American Academy of Orthopaedic Surgeons* **29**: 901-909.

[38] Chang, D.W., Weber, K.L. (2005) Use of a Vascularized Fibula Bone Flap and Intercalary Allograft for Diaphyseal Reconstruction after Resection of Primary Extremity Bone Sarcomas, *Plastic and Reconstructive Surgery* **116**: 1918-1925.

[39] Minami, A., Kasashima, T., Iwasaki, N., Kato, H., Kaneda, K. (2000) Vascularised fibular grafts, *The Journal of Bone & Joint Surgery British Volume* **82-B**: 1022-1025.

[40] Vail, T.P., Urbaniak, J.R. (1996) Donor-Site Morbidity with Use of Vascularized Autogenous Fibular Grafts, *The Journal of Bone & Joint Surgery British Volume* **78**: 204-11.

[41] Kanaya, K., Wada, T., Kura, H., Yamashita, T., Usui, M., Ishii, S. (2002) Valgus deformity of the ankle following harvesting of a vascularized fibular graft in children, *Journal of Reconstructive Microsurgery* **18**: 091-096.

[42] Lavini, F., Dall'Oca, C., Bartolozzi, P. (2010) Bone transport and compression-distraction in the treatment of bone loss of the lower limbs, *Injury* **41**: 1191-1195.

[43] Magadum, M., Yadav, C.B., Phaneesha, M., Ramesh, L. (2006) Acute compression and lengthening by the Ilizarov technique for infected nonunion of the tibia with large bone defects, *Journal of Orthopaedic Surgery* **14**: 273-279.

[44] Shetu, N.H., Balo, N.R., Rahaman, M.S., Mahmud, B.R., Islam, M.O. (2025) Clinical Outcomes of Bone Transport Over Intramedullary Nail by Ilizarov Method, *Scholars Journal of Applied Medical Sciences* **13**: 377-385.

[45] Abdel-Aal, A.M. (2006) Ilizarov bone transport for massive tibial bone defects, *Orthopedics* **29**: 70.

[46] Liu, T., Yu, X., Zhang, X., Li, Z., Zeng, W. (2012) One-stage management of post-traumatic tibial infected nonunion using bone transport after debridement, *Turkish Journal of Medical Sciences* **42**: 1111-1120.

[47] Xu, Y.-Q., Fan, X.-Y., He, X.-Q., Wen, H.-J. (2021) Reconstruction of massive tibial bone and soft tissue defects by trifocal bone transport combined with soft tissue distraction: experience from 31 cases, *BMC Musculoskeletal Disorders* **22**: 34.

[48] Bumbaširević, M., Tomić, S., Lešić, A., Milošević, I., Atkinson, H.D.E. (2010) War-related infected tibial nonunion with bone and soft-tissue loss treated with bone transport using the Ilizarov method, *Archives of Orthopaedic and Trauma Surgery* **130**: 739-749.

[49] Green, S.A. (1991) Osteomyelitis. The Ilizarov perspective, *The Orthopedic Clinics of North America* **22**: 515-521.

[50] Svesnikov, A., Barabash, A., Cheplenko, T., Smotrova, L., Larionov, A. (1984) Radionuclide studies of osteogenesis and circulation in substitution of large defects of the leg bones in experiment, *Ortopediia, Travmatologiia i Protezirovanie* 33-37.

[51] Green, S.A., Jackson, J.M., Wall, D.M., Marinow, H., Ishkanian, J. (1992) Management of Segmental Defects by the Ilizarov Intercalary Bone

Transport Method, *Clinical Orthopaedics and Related Research* **280**: 136-142.

[52] Sen, C., Eralp, L., Gunes, T., Erdem, M., Ozden, V.E., Kocaoglu, M. (2006) An alternative method for the treatment of nonunion of the tibia with bone loss, *The Journal of Bone & Joint Surgery British Volume* **88-B**: 783-789.

[53] Atesalp, A.S., Basbozkurt, M., Erler, K., Sehirlionğlu, A., Tunay, S., Solakoğlu, C., Gür, E. (1998) Treatment of tibial bone defects with the Ilizarov circular external fixator in high-velocity gunshot wounds, *International Orthopaedics* **22**: 343-347.

[54] Cattaneo, R., Catagni, M., Johnson, E.E. (1992) The Treatment of Infected Nonunions and Segmental Defects of the Tibia by the Methods of Ilizarov, *Clinical Orthopaedics and Related Research* **280**: 143-152.

[55] Zorn, K., Iii, G. (1994) Segmental Tibia1 Defects Comparing Conventional and Ilizarov Methodologies, *Clinical Orthopaedics & Related Research* **301**: 118-123.

[56] Saleh, M., Rees, A. (1995) Bifocal surgery for deformity and bone loss after lower-limb fractures. Comparison of bone-transport and compression-distraction methods, *The Journal of Bone & Joint Surgery British Volume* **77-B**: 429-434.

[57] Dendrinos, G.K., Kontos, S., Lyritsis, E. (1995) Use of the Ilizarov technique for treatment of non-union of the tibia associated with infection, *The Journal of Bone & Joint Surgery British Volume* **77**: 835-846.

[58] Polyzois, D., Papachristou, G., Kotsiopoulos, K., Plessas, S. (1997) Treatment of tibial and femoral bone loss by distraction osteogenesis, *Acta Orthopaedica Scandinavica* **68**: 84-88.

[59] Yin, P., Zhang, Q., Mao, Z., Li, T., Zhang, L., Tang, P. (2014) The treatment of infected tibial nonunion by bone transport using the Ilizarov external fixator and a systematic review of infected tibial nonunion treated by Ilizarov methods, *Acta Orthopaedica Belgica* **80**: 426-435.

[60] Thirumal, M., Shong, H.K. (2001) Bone transport in the management of fractures of the tibia, *The Medical Journal of Malaysia* **56**: 44-52.

[61] Maini, L., Chadha, M., Vishwanath, J., Kapoor, S., Mehtani, A., Dhaon, B.K. (2000) The Ilizarov method in infected nonunion of fractures, *Injury* **31**: 509-517.

[62] Ring, D., Jupiter, J.B., Gan, B.S., Israeli, R., Yaremchuk, M.J. (1999) Infected Nonunion of the Tibia, *Clinical Orthopaedics and Related Research* **369**: 302-311.

[63] Li, R., Zhu, G., Chen, C., Chen, Y., Ren, G. (2020) Bone Transport for Treatment of Traumatic Composite Tibial Bone and Soft Tissue Defects: Any Specific Needs besides the Ilizarov Technique?, *BioMed Research International* **2020**: 2716547.

[64] Schultz, J.H., Schmidt, H.G., Jürgens, C., Kortmann, H.R. (1994) Change in the treatment method of infected pseudarthrosis with bone loss of the tibia, *Zentralblatt fur Chirurgie* **119**: 714-721.

[65] Aronson, J. (1997) Current Concepts Review - Limb-Lengthening, Skeletal Reconstruction, and Bone Transport with the Ilizarov Method, *The Journal of Bone & Joint Surgery British Volume* **79**: 1243-58.

[66] Miraj, F., Nugroho, A., Dalitan, I.M., Setyarani, M. (2021) The efficacy of ilizarov method for management of long tibial bone and soft tissue defect, *Annals of Medicine and Surgery* **68**: 102645.

[67] Hernández-Irizarry, R., Quinnan, S.M., Reid, J.S., Toney, C.B., Rozbruch, S.R., Lezak, B., Fragomen, A.T. (2021) Intentional Temporary Limb Deformation for Closure of Soft-Tissue Defects in Open Tibial Fractures, *Journal of Orthopaedic Trauma* **35**: e189-e194.

TUJNAS

Review Article

*Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 15– 20*

**DOI: https://doi.org/10.59167/sbta1126**

# A Systematic Review and Meta-analysis Survey of IDS/IPS Techniques for CAN and Vehicular Networks

**Younis Abdo Mohammed Nasser Al Shojaa[1]*, Khaled Al Soufy[2]**

[1]Information Technology Department, Faculty of Engineering and Information Technology, Al-Qalam University, Ibb, Yemen.

[2]Electrical Engineering Department, Faculty of Engineering, Ibb University, Ibb, Yemen.

***Corresponding Author:** Younis A. M. N. Al Shojaa, Information Technology Department, Faculty of Engineering and Information Technology, Al-Qalam University, Ibb, Yemen. Email younis.alshogaa@quni.edu.ye

Received: 18 November 2025. Revised: 16 April 2026. Accepted: 18 April 2026. Published: 29 June 2026.

## Abstract

As cyberattacks grow increasingly sophisticated and frequent, the need for advanced mechanisms to protect modern network systems has become more pressing, especially in sensitive environments such as automotive networks. This research presents a systematic review, complemented by a sensitivity-based quantitative trend analysis, of recent efforts to employ Machine Learning (ML) and Deep Learning (DL) techniques for intrusion detection and prevention within automotive networks, with a particular focus on the Controller Area Network (CAN) bus. The research employs a clear and reproducible methodology for identifying, screening, and evaluating relevant studies, guided by the PRISMA framework. It also presents a structured classification that groups current approaches according to model type, data sources, and evaluation strategies. Furthermore, the review provides a comparative analysis highlighting the key strengths, weaknesses, and experimental performance of various IDS/IPS techniques in automotive environments. In addition, the study includes a sensitivity-based quantitative summary of reported accuracy patterns in the literature, offering an exploratory and sensitivity-aware view of performance trends rather than a formal pooled estimate. The analysis encompassed 56 studies published between 2018 and 2025, covering diverse datasets, methodological designs, and multiple detection objectives. The findings reveal ongoing challenges, including real-time limitations, limited computational resources, and dataset variability, while also pointing to promising research avenues that could contribute to the development of more efficient AI-based intrusion detection and prevention systems for connected and autonomous vehicles.

**Index Terms—**CAN Bus, Vehicular Networks, In- Vehicle Communication, Intrusion Detection System, Sensitivity Analysis, Machine Learning, Deep Learning, Automotive Cybersecurity

## 1. Introduction

With the rapid development of modern technologies and the ever-increasing reliance of systems on networking across various sectors of life, industry, and transportation, cybersecurity has become more critical than ever. Maintaining data confidentiality, integrity, and availability is no longer a secondary option but a fundamental necessity, especially given the rise in cyberattacks and their significantly evolving methods. In this context, Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) represent the first line of defense, monitoring network traffic and detecting any suspicious behavior before it develops into an actual threat. However, traditional signature-based methods often struggle to address new and sophisticated attacks, particularly in sensitive environments that demand immediate response, limited resources, and high reliability. Artificial intelligence and ML techniques have gained increasing attention due to their ability to enhance the performance of detection systems and provide higher levels of generalization and flexibility in addressing unknown threats [1-5].

This research employs a hybrid structure that combines systematic review methodology with a sensitivity-based quantitative trend analysis. The established steps of systematic reviews were followed, including comprehensive database searches, screening processes, quality assessment, and study selection within the PRISMA framework. Simultaneously, the research incorporates a broader analytical perspective, consistent with the nature of survey studies, allowing for comparison of current trends and review of advanced methods and techniques in this field.

In addition to qualitative analysis, the study includes a sensitivity-based quantitative synthesis intended to summarize reported accuracy patterns across several IDS/IPS studies under explicit simplifying assumptions. This component is not intended to provide a formal pooled effect estimate; rather, it offers an exploratory view of performance trends across the reviewed literature. This dual approach provides both analytical depth and a structured quantitative perspective, contributing to a more comprehensive understanding of detection and prevention mechanisms within vehicle networks, particularly those based on the CAN

bus.

### 1.1 Scope

This research focuses on intrusion detection and prevention mechanisms within vehicle communication environments, with particular attention to the control bus known as CAN. Accordingly, the study is limited to data traffic within vehicles and their associated networks, excluding enterprise or public networks.

In this context, the research reviews the most recent developments in employing artificial intelligence techniques to detect malicious activity in vehicle systems. It also addresses ML and DL methods used to improve detection speed and accuracy and enhance system resilience, while acknowledging the limitations identified in previous studies. Furthermore, the research discusses the key challenges facing the development of intelligent detection and prevention solutions in resource- constrained vehicle environments and highlights research trends that warrant further attention in the future [1, 4-7].

### 1.2 Contributions

This review offers several contributions that distinguish it from previous studies in the field. First, it employs a clear and reproducible screening methodology applied to 56 studies published between 2018 and 2025, enabling researchers to verify and build upon the findings in the future. Second, it presents an improved classification for organizing ML and DL techniques according to network type in vehicles and their application areas, thus helping to organize a fragmented research field. Third, the review goes beyond traditional descriptive comparisons by linking methodological strengths and weaknesses to practical constraints, such as real-time requirements, memory limitations, resilience to repeated attacks, and feasibility in resource-constrained systems. Fourth, it provides a comprehensive analytical framework that connects datasets, evaluation metrics, and factors related to scientific publication, offering a practical model for comparing different IDS/IPS techniques. Finally, the review highlights promising research trends focused on developing lighter, faster, and more standardized solutions to support the next generation of intelligent detection and prevention systems in vehicle environments.

## 2. Methodology

This study employs a hybrid methodology that aligns with the principles of systematic reviews and survey analyses, complemented by a sensitivity-based quantitative trend analysis. It followed a structured systematic literature review (SLR), which included identifying studies, screening titles and abstracts, assessing eligibility, and ultimately selecting the final studies based on the PRISMA framework. Simultaneously, the study incorporated a subject-based survey classification and comparative analysis to identify technical trends and methodological categories for intrusion detection systems in CAN and vehicular networks. A sensitivity-based quantitative analysis was conducted to summarize reported accuracy trends across the included studies under transparent simplifying assumptions.

This section provides a detailed description of each methodological step in a transparent and reproducible manner, enabling future researchers to replicate or expand the study.

### 2.1 Information Sources and Search Strategy

This review adopted a systematic search methodology inspired by PRISMA principles, documenting the search and screening steps, as well as the acceptance and exclusion criteria, to ensure transparency and replicability. The search process relied on specialized and widely used databases in the fields of computer engineering and cybersecurity, including the IEEE Xplore platform, which includes relevant IEEE journals and conferences (such as IEEE Access, IV Conferences, VTC, and GLOBECOM), and the ACM Digital Library, which focuses on security and networking (such as the ACM AsiaCCS conference).

The study also utilized Elsevier's ScienceDirect database, which includes journals such as *Vehicular Communications*, *Internet of Things*, and *Future Generation Computer Systems*, along with open-access platforms like *MDPI (Sensors)* and *Hindawi (Security and Communication Networks)*, and the *arXiv* repository of primary research in artificial intelligence and cybersecurity.

The logical query syntax was customized for each database using keywords reflecting the scope of the study, such as:

"intrusion detection" / IDS / IPS AND (CAN OR "in-vehicle" OR automotive OR "controller area network") AND ("machine learning" OR "deep learning")

In addition, subtypes of attacks or environments (such as DoS, spoofing, and anomaly detection) were included. The search period was limited to 2018–2025, prioritizing peer-reviewed research in cybersecurity and intelligent vehicle systems. A citation tracking mechanism (reviewing the references of listed studies and the studies that cited them) was used to ensure that no work was missed in automated searches.

The screening process was initially based on titles and abstracts, followed by full-text assessment according to predefined criteria. These procedures help ensure the study can be replicated and its findings updated in line with PRISMA 2020 recommendations [6-8].

### 2.2 Eligibility Criteria

#### *2.2.1 Inclusion*

The study included peer-reviewed research or authoritative manuscripts published between 2018 and 2025 related to network traffic analysis in the field of cybersecurity, which relies on artificial intelligence (ML or DL) techniques for intrusion detection and/or intrusion prevention (IDS/IPS) within computer networks, vehicles, or industrial systems. For acceptance, studies had to provide a quantitative assessment such as accuracy, precision, recall, F1 value, AUC, or other similar metrics, in addition to a description of the data used, workloads, or comparison and evaluation environments. The study areas were particularly focused on in-vehicle networks (CAN/IoV) and wireless industrial sensor networks (WSNs), given their prevalence in recent research and their reliance on common standards and designs.

#### *2.2.2 Exclusion*

Editorial reports, short messages, research commentaries, and theoretical or conceptual studies lacking empirical evaluation were excluded. Research outside the scope of networking or cybersecurity, duplicates or replaced versions, and studies containing inconsistent information that prevents reliable analysis or replication of results were also excluded. Additionally, work with insufficient evaluation details that clearly precluded re- experiments or comparisons was not prioritized.

Rationale: Research on intrusion detection systems in vehicle control systems (CAN/IoV IDS) focuses on standardized metrics and comparable evaluation environments. Similarly, studies on industrial sensor networks (WSNs) rely on standardized performance indicators in ML and DL applications. Therefore, the inclusion criteria require the submission of quantitative results and clear, reliable evaluation environments to ensure reproducibility and reusability. To preserve domain specificity, studies based on CAN-native and vehicular datasets were treated as the core evidence base of this review, whereas general benchmark datasets were retained only for contextual comparison and were interpreted separately where necessary.

### 2.3 Screening Workflow

Figure 1 presents the finalized PRISMA-based screening workflow used in this review.

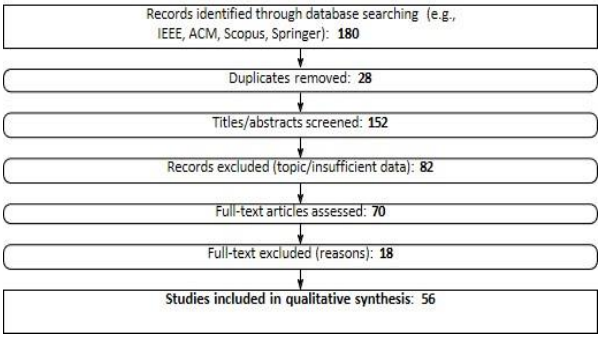



**Figure 1:** Final PRISMA-based screening workflow.

### 2.4 Data Extraction & Quality Appraisal

For each study, a set of elements was extracted, including: the datasets used, types of attacks, model family, adopted characteristics, training system, and key metrics such as accuracy, precision, recall, F1 value, and AUC, in addition to execution or response time, and computation and memory requirements. The risks of bias were also discussed from multiple perspectives, including the potential for data leakage, addressing class imbalances, the transparency of model parameter

adjustments, and the generalizability of the results across different environments. Some cited studies were authored by the present research group; however, they were retained solely because they satisfied the predefined inclusion criteria and contributed directly to the comparative scope of the review.

## 3. Taxonomy of AI Techniques

As seen, Figure 2 provides a multiaxis taxonomy.

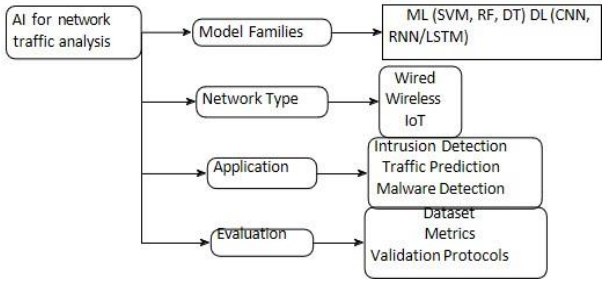



**Figure 2:** Mind-map taxonomy for AI in network traffic analysis.

## 4. Comparative Analysis

In this section, we present a concise comparative discussion of representative studies that have addressed ML, DL, and hybrid approaches for intrusion detection and prevention in CAN and vehicular network environments. For clarity and to maintain the flow of the main text, the detailed comparative tables are provided in the APPENDICES. Appendix A summarizes representative ML studies. Appendix B summarizes representative DL studies. Appendix C presents combined, hybrid, benchmarking, and overview studies.

Some representative studies have been selected to illustrate the current state of IDS/IPS research in CAN and vehicular network environments. These studies were chosen according to the predefined inclusion criteria and their relevance to the objectives of this review. Recent review studies have emphasized the rapid evolution of AI-based intrusion detection techniques in automotive systems, while comparative studies have highlighted the importance of evaluating different machine learning approaches under realistic conditions to identify their strengths and limitations [6, 9]. Accordingly, the selected studies provide a balanced and representative overview of current methodological trends, supporting a comprehensive comparison of ML-, DL-, and hybrid-based IDS/IPS techniques. These supporting tables allow easy reference while preserving the readability of the main discussion.

## 5. Quantitative Trend Analysis Framework

Although this research is primarily presented as a systematic review, a brief quantitative framework is included to clarify how reported accuracy values from the included IDS/IPS studies were summarized. Because many studies did not report sample sizes, confidence intervals, or variance estimates, this component should not be interpreted as a formal effect-size meta-analysis. Rather, it provides a sensitivity-based quantitative summary of reported performance trends under explicit variance assumptions.

For each study, denoted by $k$, the observed accuracy is represented by the coefficient $\hat{\theta}_k$.

When sample sizes are available, the variance can be calculated using the standard estimator for the Bernoulli variable as follows:

$$V_k = \frac{\hat{\vartheta}_k(1-\hat{\vartheta}_k)}{n_k} \qquad (1)$$

However, because several reviewed studies do not report $n_k$, a sensitivity-based variance estimation approach was adopted. This approach assumes a set of representative sample sizes:

$$n \in \{100, 500, 1000, 5000\},$$

This procedure allows for the evaluation of the resilience of the aggregated accuracy value when tested under different assumptions of statistical conditions. To examine the stability of trend-level summaries under different assumptions, fixed- and random-effect style calculations were explored as sensitivity tools rather than as formal pooled estimators.

In the fixed-effects model, the standardized effect value is calculated using the following formula:

$$\hat{\vartheta}_{FE} = \frac{\sum_k W_k \hat{\vartheta}_k}{\sum_k W_k}, \qquad W_k = \frac{1}{V_k} \qquad (2)$$

For the random-effects model, the DerSimonianLaird estimator is used to incorporate between-study variance ($\tau^2$):

$$\hat{\vartheta}_{FE} = \frac{\sum_k W_k^* \hat{\vartheta}_k}{\sum_k W_k^*}, \qquad W_k^* = \frac{1}{V_k + \tau^2} \qquad (3)$$

This framework is included to enhance transparency and to explain the computational logic underlying the sensitivity-based quantitative summary presented later in the paper. It is not intended as a full formal meta-analytic derivation.

## 6. Dataset and Evaluation Protocols

This section provides a concise overview of the datasets and evaluation protocols used in the studies included in this survey.

### 6.1 Datasets

The reviewed literature draws on both domain-specific vehicular datasets and general IDS benchmark datasets. CAN-specific and vehicular datasets, such as Car-Hacking, ROAD, real CAN traces, J1939-based traffic, and Internet of Vehicles (IoV) datasets, more directly reflect in-vehicle communication characteristics, message timing behavior, and attack surfaces relevant to automotive systems. These datasets are therefore the most informative for assessing IDS/IPS applicability in CAN and vehicular environments [1, 4, 10, 11].

By contrast, benchmark datasets such as KDD Cup, NSL-KDD, and CICIDS are more general-purpose intrusion detection datasets. Although they are not specific to in-vehicle communication, they remain useful for contextualizing broader model behavior, algorithmic trends, and comparative IDS performance under controlled benchmark settings. General benchmark datasets such as KDD Cup and NSL-KDD were retained only to contextualize broader IDS model behavior and were analytically distinguished from in-vehicle CAN-specific datasets.

### 6.2 Evaluation Protocols and Metrics

Across the reviewed studies, evaluation practices vary considerably in terms of train-test splitting strategies, attack composition, feature preprocessing, and validation protocols. Commonly reported metrics include accuracy, precision, recall, F1-score, detection rate, false alarm rate, AUC, and response time. While these metrics are useful for comparative interpretation, their meaning depends strongly on the dataset type, class balance, and experimental setup. This heterogeneity limits direct one-to-one comparison across studies and further highlights the need for more standardized evaluation frameworks in future research.

Although accuracy was used as the main quantitative indicator in the trend analysis due to its wider availability across the reviewed studies, it should not be interpreted in isolation. Its meaning must be considered alongside precision, recall, F1-score, false alarm rate, and dataset balance in order to avoid overestimating model performance under heterogeneous evaluation conditions.

## 7. Sensitivity-Based Quantitative Trend Analysis

To complement the qualitative review and provide a structured quantitative perspective on IDS/IPS performance in vehicle network studies, a sensitivity-based quantitative trend analysis was conducted. Reported accuracy values from representative studies were examined using transparent simplifying assumptions in order to summarize general performance tendencies across methodological categories. Given the inconsistent availability of sample sizes, confidence intervals, and variance measures in the source studies, this component should be interpreted as an exploratory quantitative summary rather than as a formal pooled meta-analysis. Table I summarizes representative studies included in this analysis, while Figure 3 provides a comparative visualization of their reported accuracy values.

This analysis is intended to reveal broad performance tendencies under sensitivity assumptions and should therefore be interpreted as trend-oriented rather than as a formal pooled meta-analytic estimate.

The quantitative summary suggests a general trend in which deep learning approaches often achieve higher reported performance, followed by hybrid approaches, while traditional machine learning methods remain attractive in lightweight and resource-constrained settings. The sensitivity analysis showed limited variation across the explored assumptions,

suggesting that the observed trend pattern remained broadly stable within the tested scenarios.

**Table I.** Sensitivity-Based Quantitative Summary of Representative IDS/IPS Studies

| Study | Technique | Metric | Score |
|---|---|---|---|
| Al-Quayed *et al.* [12] | DT, RF, SVM | Accuracy | 0.9948 |
| Alalwany *et al.* [6] | Stacking Ensemble Learning | Accuracy | 0.984 |
| Nair *et al.* [13] | ML+DL (SVM, RF, XGBoost + CNN, RNN) | Accuracy | 0.965 |
| Rani *et al.* [14] | Comparative ML/DL architectures | Accuracy | 0.970 |
| Sharmin *et al.* [15] | Survey + Benchmarking | Accuracy | 0.950 |
| Dayyeh *et al.* [16]. | Multi-stage FL IDS | Accuracy | 0.960 |
| Young *et al.* [ 17] | ML+DL (SVM, RF, kNN, DNN, CNN) | Accuracy | 0.950 |
| Tanksale [18] | IDS survey (ML, DL) | Accuracy | 0.940 |
| Apruzzese *et al.* [19] | IDS Evaluation Framework | Accuracy | 0.955 |
| Yang, *et al.* [20] | Hybrid Statistical + NN | Accuracy | 0.990 |

**Note:** The quantitative summary is exploratory and sensitivity-based; variance-related interpretation relies on approximations because several source studies did not report sample sizes or confidence intervals.

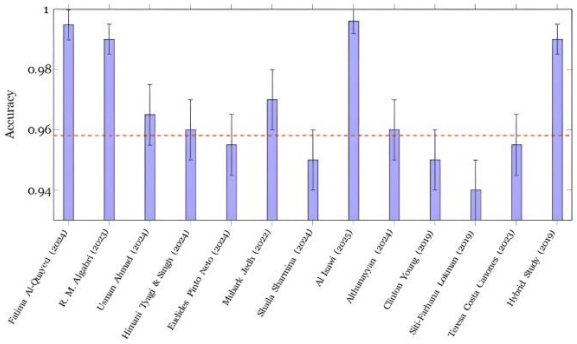



**Figure 3:** Comparative visualization of reported accuracy values across the included IDS/IPS studies.

### 7.1 Sensitivity Analysis

Due to the absence of reported sample sizes or confidence intervals in several studies, a sensitivity-based analysis was conducted to examine how reported accuracy values behave under different simplifying assumptions. To approximate variance, Bernoulli-based estimators were used under a set of hypothetical sample sizes:

$$\in\{100,\ 500,\ 1000,\ 5000\}.$$

Rather than producing formal pooled estimates, this analysis was designed to evaluate the stability of observed performance patterns under varying assumptions. The results showed limited variation across the explored scenarios, indicating that the general trend of reported performance remained broadly stable within the tested conditions.

This behavior suggests that the relative ordering of methodological approaches (e.g., DL, hybrid, and ML) is consistent across different sensitivity assumptions. However, these findings should be interpreted as indicative trends rather than statistically precise aggregate estimates, given the absence of consistent variance and sample-size information in the original studies.

### 7.2 Interpretation

The results suggest that deep learning techniques generally achieve higher reported performance in intrusion detection tasks, primarily due to their ability to capture complex temporal and structural patterns in vehicular network data. Hybrid approaches combining ML and DL also demonstrate strong performance, although they typically require higher computational resources.

Traditional machine learning methods remain effective in lightweight and resource-constrained environments, offering a practical trade-off between accuracy and efficiency. Overall, the findings highlight the importance of balancing detection performance with computational cost and deployment feasibility in CAN-based and vehicular systems.

These observations should be interpreted as general performance trends derived under sensitivity assumptions, rather than as definitive pooled performance estimates [1, 2, 21-23].

## 8. Critical Discussion

### 8.1 Reasons for the Success or Failure of Different Methods

The studies reviewed demonstrate that the performance of AI-based intrusion detection systems depends heavily on the type of model used, the characteristics of the data, and the operational constraints of the deployment environment. Traditional machine learning methods, particularly Random Forest (RF) and Support Vector Machine (SVM), remain widely adopted in intrusion detection for vehicular networks due to their robust classification capabilities and relatively low computational requirements. The reviewed studies indicate that RF models achieve reliable performance in detecting known attack patterns through ensemble learning, whereas their effectiveness may decrease when confronted with previously unseen or evolving attacks. In contrast, SVM-based models are capable of learning complex decision boundaries and can identify certain novel attack patterns; however, several studies have reported higher false-positive rates under heterogeneous traffic conditions, which may limit their applicability in real-world vehicular environments [1, 8, 11, 21].

To overcome these limitations, many recent studies have proposed hybrid architectures that integrate multiple machine learning algorithms or combine ML with deep learning techniques. Such approaches generally improve detection accuracy, robustness, and generalization capability, although these gains are often achieved at the expense of increased model complexity and computational overhead, which may reduce their suitability for real-time deployment in resource-constrained vehicular systems [8, 11, 21, 23].

To address these limitations, several studies have proposed hybrid approaches that combine multiple ML algorithms or integrate ML with deep learning (DL) techniques. These models often improve detection accuracy and robustness; however, they introduce additional complexity in model design and require higher computational resources, which can negatively impact their suitability for real-time applications.

In the domain of deep learning, convolutional neural networks (CNNs) demonstrate strong performance in capturing local patterns in CAN traffic and extracting low-level signal features, making them well-suited for complex intrusion detection scenarios. However, their effectiveness depends heavily on the availability of large, well-labeled datasets, and they require careful hyperparameter tuning to avoid overfitting. Recurrent neural networks (RNNs) and long short-term memory (LSTM) models are particularly effective in modeling temporal dependencies in sequential CAN messages. Despite their advantages, their relatively slow training and inference processes limit their applicability in time-critical environments.

Ensemble DL methods further enhance detection performance and improve resilience against diverse attack types, but they significantly increase computational and memory requirements. This makes their deployment challenging in resource-constrained environments such as in-vehicle networks and IoT-based vehicular systems.

### 8.2 External Validity and Reproducibility

The reviewed studies reveal clear challenges related to external validity and reproducibility. A considerable number of studies rely on synthetic, imbalanced, or domain-specific datasets that may not fully represent real-world vehicular traffic conditions. This increases the risk of overfitting and limits the generalizability of the reported results. Additionally, inconsistencies in data preprocessing, segmentation, and feature engineering may lead to unintended data leakage, which can artificially inflate performance metrics. Furthermore, some studies employ validation strategies that do not reflect real deployment scenarios, such as random cross-validation without considering temporal dependencies or data source separation. These practices reduce the reliability of performance claims and hinder fair comparison across different approaches [8, 24, 25].. To improve external validity, future research should adopt more rigorous evaluation protocols, such as time-aware data splitting, source-based

validation, and explicit reporting of preprocessing pipelines. Reporting inference time on target hardware and conducting ablation studies to quantify the impact of each component would further enhance reproducibility and transparency.

### 8.3 Practical Deployment Considerations

While many IDS/IPS approaches achieve high performance in controlled experimental settings, their deployment in real-world vehicular environments introduces additional challenges. Computational efficiency remains a critical concern, as many high-performing DL models require substantial processing power and memory resources, making them unsuitable for embedded systems such as electronic control units (ECUs).

To enable practical deployment, techniques such as model compression, pruning, and quantization are essential to reduce model size and inference cost without significantly degrading performance. Inference latency is another key factor, as intrusion detection systems must operate under strict real-time constraints. Models with high latency or those requiring batch processing may fail to meet these requirements.

Practical deployment of IDS/IPS solutions in vehicular environments remains challenging due to compatibility requirements with existing automotive communication protocols and system architectures [26, 27]. In addition, the absence of standardized evaluation frameworks and testing methodologies makes objective comparison between different approaches difficult and hinders their wider industrial adoption [28]. Furthermore, practical deployment requires addressing computational efficiency, privacy preservation, and low-latency processing, particularly in resource-constrained vehicular systems and edge-based environments[22, 29-31]. These findings collectively indicate that achieving practical and scalable deployment requires balancing detection performance with implementation feasibility.

Finally, scalability and cost considerations play an important role in deployment decisions. Cloud-based solutions offer high scalability but may introduce latency, privacy, and security concerns. Edge-based solutions reduce latency but require lightweight and efficient models capable of operating under limited resources. Achieving an optimal balance between detection accuracy, computational cost, latency, and deployment feasibility remains a central challenge in this field.

## 9. Challenges and Open Issues

Despite significant advancements in intrusion detection and protection technologies, several challenges still hinder the effective security of CAN networks [1, 5, 26, 27].

### 9.1 Resource Limitations

Many electronic control units (ECUs) have limited computing power and memory capabilities, making it difficult to implement resource-intensive security mechanisms.

### 9.2 Real-Time Requirements

Systems in vehicles rely on real-time communication, and introducing additional security protocols can increase latency, complicating implementation.

### 9.3 Lack of Standardized Standards

There are no widely agreed-upon security standards for CAN networks, which limit the ability to develop consistent and reliable protection solutions across different systems.

## 10. Limitations

Although this review provides a comprehensive overview of the latest developments in IDS/IPS techniques for CAN networks, several limitations should be acknowledged. First, many of the included studies did not report sample sizes or variance measures in a sufficiently consistent manner. As a result, the quantitative component of this review could not be conducted as a formal effect-size meta-analysis. Instead, a sensitivity-based quantitative trend analysis was used to explore reported performance patterns under explicit simplifying assumptions. Second, substantial heterogeneity exists across the reviewed studies in terms of datasets, feature representations, attack scenarios, evaluation metrics, and experimental protocols. This variability limits strict cross-study comparability and requires cautious interpretation of any quantitative summary. Third, some included studies rely on benchmark datasets that are not specific to in-vehicle communication environments. Although these studies were retained for contextual comparison, they should not be interpreted as equivalent to evidence derived from CAN-native or vehicular datasets. Accordingly, the quantitative findings presented in this review should be interpreted as broad trend indicators rather than definitive pooled estimates.

## 11. Future Research Trends

Based on the review and analysis, several promising research paths for the future can be suggested:

- **Developing Lightweight Security Solutions:** Future research should focus on designing effective security mechanisms that are compatible with the limited constraints of electronic control units within vehicles.
- **Testing and Validation in Real-World Environments:** Evaluation in real-world environments is a crucial step to ensure that detection and protection systems can operate effectively under varying and practical operating conditions.
- **Integrating ML Technologies:** The use of ML contributes to improving the ability to detect anomalous attacks while simultaneously reducing false alarm rates.
- **Prioritizing Research Areas:** It is important to direct research efforts according to priority and expected impact, focusing on the most prominent challenges facing vehicle cybersecurity.

## 12. Conclusion

This review has provided a comprehensive overview of current solutions for intrusion detection and prevention systems in CAN networks, employing a methodology that combines systematic review with a sensitivity-based quantitative trend analysis. This approach integrates structured evidence analysis with a broad analytical comparison of IDS/IPS methods implemented in these networks. The quantitative findings reported in this review should be interpreted as broad trend indicators that support comparative insight, rather than as definitive pooled performance estimates. Despite significant progress in this field, challenges remain in securing critical in-vehicle communication systems. Therefore, future research should focus on developing lightweight and highly efficient security solutions capable of addressing the growing cyber threats targeting modern vehicles.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Conflict of Interest

The authors declare that there are no conflicts of interest.

## References

[1] Kalkan, S.C., Sahingoz, O.K. (2020) In-vehicle intrusion detection system on controller area network with machine learning models. *In Proceedings of the 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT 2020)*, *IEEE*, Kharagpur, India, 1-3 July, 2020, pp. 1-6.

[2] Alfardus, A., Rawat, D.B. (2021) Intrusion detection system for can bus in-vehicle network based on machine learning algorithms. *In Proceedings of the 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, *IEEE*, 1-4 December, 2021, pp. 0944-0949.

[3] Al-Janabi, M., Ismail, M.A., Ali, A.H. (2021) Intrusion Detection Systems, Issues, Challenges, and Needs, *The International Journal of Computational Intelligence Systems* **14**: 560-571.

[4] Panigrahi, R., Borah, S., Bhoi, A.K., Ijaz, M.F., Pramanik, M., Jhaveri, R.H., Chowdhary, C.L. (2021) Performance assessment of supervised classifiers for designing intrusion detection systems: a comprehensive review and recommendations for future research, *Mathematics* **9**: 690.

[5] Micale, D., Costantino, G., Matteucci, I., Fenzl, F., Rieke, R., Patanè, G. (2022) Cahoot: a context-aware vehicular intrusion detection system. *In Proceedings of the 2022 IEEE international conference on trust,*

*security and privacy in computing and communications (TrustCom)*, *IEEE*, Wuhan, China, pp. 1211-1218.

[6] Alalwany, E., Mahgoub, I. (2022) Classification of normal and malicious traffic based on an ensemble of machine learning for a vehicle can-network, *Sensors* **22**: 9195.

[7] Rajapaksha, S., Kalutarage, H., Al-Kadri, M.O., Petrovski, A., Madzudzo, G., Cheah, M. (2023) Ai-based intrusion detection systems for in-vehicle networks: A survey, *ACM Computing Surveys* **55**: 1-40.

[8] Nagarajan, J., Mansourian, P., Shahid, M.A., Jaekel, A., Saini, I., Zhang, N., Kneppers, M. (2023) Machine Learning based intrusion detection systems for connected autonomous vehicles: A survey, *Peer-to-Peer Networking and Applications* **16**: 2153-2185.

[9] Kumar, A., Das, T.K. (2023) CAVIDS: Real time intrusion detection system for connected autonomous vehicles using logical analysis of data, *Vehicular Communications* **43**: 100652.

[10] Alalwany, E., Mahgoub, I. (2024) An effective ensemble learning-based real-time intrusion detection scheme for an in-vehicle network, *Electronics* **13**: 919.

[11] Shahriar, M.H., Xiao, Y., Moriano, P., Lou, W., Hou, Y.T. (2023) CANShield: Deep-learning-based intrusion detection framework for controller area networks at the signal level, *IEEE Internet of Things Journal* **10**: 22111-22127.

[12] Al-Quayed, F., Ahmad, Z., Humayun, M. (2024) A situation based predictive approach for cybersecurity intrusion detection and prevention using machine learning and deep learning algorithms in wireless sensor networks of industry 4.0, *IEEE Access* **12**: 34800-34819.

[13] Nair, R. (2023) Unraveling the Decision-making Process Interpretable Deep Learning IDS for Transportation Network Security, *Journal of Cybersecurity & Information Management* **12**.

[14] Rani, D., Kaushal, N.C. (2020) Supervised machine learning based network intrusion detection system for Internet of Things. *In Proceedings of the 2020 11th International conference on computing, communication and networking technologies (ICCCNT)*, *IEEE*, pp. 1-7.

[15] Sharmin, S., Mansor, H., Abdul Kadir, A.F., Aziz, N.A. (2024) Benchmarking frameworks and comparative studies of Controller Area Network (CAN) intrusion detection systems: A review, *Journal of Computer Security* **32**: 477-507.

[16] Dayyeh, R., AlSawareah, W., Kasasbeh, B., Qaddoura, R., Kamal, S. (2023) Comparative Analysis of Decision Trees, Random Forest, and k-Nearest Neighbors in Predictive Analytics for Orange Telecom's Customer Complaint Data. *In Proceedings of the 2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)*, *IEEE*, Zarqa, Jordan, pp. 1-6.

[17] Young, C., Olufowobi, H., Bloom, G., Zambreno, J. (2019) Automotive intrusion detection based on constant can message frequencies across vehicle driving modes. *In Proceedings of the 2019 ACM Workshop on Automotive Cybersecurity (AutoSec 2019)*, *Association for Computing Machinery (ACM)*, Richardson, Texas, USA, March 27, 2019, pp. 9-14.

[18] Tanksale, V. (2024) Intrusion detection system for controller area network, *Cybersecurity* **7**: 4.

[19] Apruzzese, G., Laskov, P., Schneider, J. (2023) Sok: Pragmatic assessment of machine learning for network intrusion detection. *In Proceedings of the 2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*, *IEEE*, Delft, Netherlands, 3-7 July, 2023, pp. 592-614.

[20] Kocher, G., Kumar, G. (2022) A hybrid deep learning approach for effective intrusion detection systems using spatial-temporal features, *Adv. Eng. Sci.* **54**: 1503-1519.

[21] Ahmed, N., Hassan, F., Aurangzeb, K., Magsi, A.H., Alhussein, M. (2024) Advanced machine learning approach for DoS attack resilience in internet of vehicles security, *Heliyon* **10**: e28844.

[22] Alemerien, K., Al-Suhemat, S., Almahadin, M. (2024) Towards optimized machine-learning-driven intrusion detection for Internet of Things applications, *International Journal of Information Technology* **16**: 4981-4994.

[23] Kousar, A., Ahmed, S., Altamimi, A., Khan, Z.A. (2024) A Novel Light-Weight Machine Learning Classifier for Intrusion Detection in Controller Area Network in Smart Cars, *Smart Cities* **7**: 3289-3314.

[24] El-Gayar, M.M., Alrslani, F.A., El-Sappagh, S. (2024) Smart collaborative intrusion detection system for securing vehicular networks using ensemble machine learning model, *Information* **15**: 583.

[25] Musa, U.I., Musa, A.I., Galadima, Y.I., Kantunsung, R.O., Gupta, S., Dua, S. (2024) Machine Learning-Based Cybersecurity Optimization in Internet of Things-Enabled Autonomous Vehicles. *In Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR)*, *IEEE*, Muscat, Oman, 14-15 May, 2024, pp. 1-6.

[26] Rajapaksha, S., Madzudzo, G., Kalutarage, H., Petrovski, A., Al-Kadri, M.O. (2024) CAN-MIRGU: a comprehensive CAN bus attack dataset from moving vehicles for intrusion detection system evaluation. *In Proceedings of the 2nd Symposium on Vehicle Security and Privacy (VehicleSec 2024)*, San Diego, California, USA, February 26 to March 1, 2024, pp. 43.

[27] Wasicek, A., Pesé, M.D., Weimerskirch, A., Burakova, Y., Singh, K. (2017) Context-aware intrusion detection in automotive control systems. *In Proceedings of the 5th ESCAR USA Conference (Embedded Security in Cars)*, Ypsilanti, Michigan, USA, 21-22 June, 2017, pp. 21-22.

[28] Ossen, S. (2025) Enabling Low-Latency High-Throughput Real-Time Stream Processing Using Smart Network Compute Elements. *Department of Computer Science*, Ph.D. Thesis, Indiana University, Bloomington, Indiana, USA.

[29] Liu, H., Lang, B., Liu, M., Yan, H. (2019) CNN and RNN based payload classification methods for attack detection, *Knowledge-Based Systems* **163**: 332-341.

[30] Faker, O., Dogdu, E. (2019) Intrusion detection using big data and deep learning techniques. *In Proceedings of the 2019 ACM Southeast Conference (ACMSE 2019)*, Kennesaw, Georgia, USA, 18-20 April, 2019, pp. 86-93.

[31] Kumar, A., Sharma, I. (2023) CNN-based approach for IoT intrusion attack detection. *In Proceedings of the 2023 International conference on sustainable computing and data communication systems (ICSCDS)*, *IEEE*, Erode, India, 23-25 March, 2023, pp. 492-496.

TuJNaS

Review Article

*Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 21– 28*

**DOI:** https://doi.org/10.59167/g1qw6p23

# Deep Learning for Respiratory Sound Analysis: A Systematic Review and Meta-Analysis (2019–2024)

**Shaima'a Mohammed Nasser Al-Jabali[1]*, Farhan Nashwan[2], Waleed M. Altalabi[3]**

[1]Information Technology Department, Faculty of Engineering and Information Technology, AL-Qalam University, Ibb, Yemen

[2]Electrical Engineering Department, Faculty of Engineering, IBB University, IBB, Yemen

[3]Biomedical Engineering Department, Sana'a Community College, Sana'a, Yemen

***Corresponding Author:** Shaima'a M. N. Al-Jabali, Information Technology Department, Faculty of Engineering and Information Technology, AL-Qalam University, Ibb, Yemen. Email: shaima.aljabali@gmail.com

Received: 18 November 2025. Revised: 12 April 2026. Accepted: 12 April 2026. Published: 29 June 2026.

## Abstract

**Background:** Respiratory diseases remain a significant global health burden, particularly in resource-limited settings. Objective: To review and quantitatively analyze deep learning-based models for interpreting respiratory sounds. **Methodology:** A systematic search was conducted in the PubMed, IEEE Xplore, ScienceDirect, and Scopus databases for the period 2019–2024, using PRISMA 2020 criteria. Of the 678 records accessed, 98 studies met the qualitative criteria, while 42 met the quantitative analysis criteria. **Results:** An exploratory meta-analysis conducted on a subset of studies reporting native accuracy (k = 9) yielded a pooled accuracy of approximately 90.0% (95% CI: 76.0%–97.0%), with substantial heterogeneity ($I^2$ = 99.2%). **Conclusion:** Deep learning techniques demonstrate promising diagnostic capabilities, but they still face challenges related to reliance on limited databases, poor representation of rare diseases, and difficulty in interpreting their outputs. **Implications:** Future research should focus on diversifying datasets, enhancing the integration of multimodal data, and developing interpretable or federated learning-based models to support their adoption in clinical settings.

**Index Terms:** Respiratory sounds; lung disease classification; Deep learning; Systematic review; Meta-analysis; Multitask learning; Explainable AI.

## 1. Introduction

Respiratory diseases are among the most significant global health challenges, including conditions such as asthma, chronic obstructive pulmonary disease (COPD), pneumonia, and bronchiectasis. These diseases continue to cause high rates of morbidity and mortality worldwide, particularly in low- and middle-income countries with limited healthcare resources. Early and accurate diagnosis remains essential for improving clinical outcomes and reducing economic burdens [1].

According to the 2019 Global Burden of Disease report, there were approximately 454.6 million active clinical cases of chronic respiratory diseases, with 4 million deaths in the same year, making these diseases the third among the leading cause of death globally [1]. This underscores the critical need for accurate, low-cost, and widely applicable diagnostic tools.

Unlike previous narrative reviews that merely presented the literature without quantitative analysis, this study applies a PRISMA 2020–guided systematic review framework and includes a structured quantitative synthesis of recent deep learning studies on respiratory sound analysis (2019–2024). It offers a comprehensive synthesis that includes the statistical integration of key indicators such as accuracy, variance, and subgroup analysis. This methodology represents a step toward bridging the gap between scattered evidence and reproducible, unified insights [2].

Traditional diagnostic methods—such as X-rays, computed tomography (CT) scans, and pulmonary function tests—remain effective, but they are often expensive or require specialized equipment. While auscultation with a stethoscope is a widely available and easy-to-use option, its accuracy is influenced by the practitioner's experience and can vary from one physician to another. To overcome these limitations, computational analysis of respiratory sounds has emerged as an objective and non-invasive alternative, transforming the auscultation process into a measurable and repeatable step, thus expanding its applicability in resource-constrained environments [2].

Recent years have witnessed significant advancements in the use of deep learning (DL) techniques for respiratory sound analysis. Convolutional network (CNN) models, recurrent network (RNN/LSTM) models, and newer models based on transformers or hybrid structures have proven their ability to extract accurate and meaningful sound patterns [3-5]. Their applications range from disease detection and classification of various pulmonary conditions to identifying sounds such as wheezing and rattling, and even to estimating disease severity, with studies ranging from analyzing wheezing sounds in children to modeling the severity of COPD [6, 7]. Recent review studies have further confirmed the effectiveness of deep learning–based auscultation systems and intelligent stethoscopes in improving diagnostic performance and automation capabilities [8, 9].

Despite this progress, several challenges remain, most notably:

- **Over-reliance on the ICBHI 2017 database** — which—while contributing to research—limits the generalizability of models to real-world settings [10, 11].

- **Data imbalance and scarcity** — with rare diseases being underrepresented, leading to model bias and inconsistent results. In addition, the availability of diverse and clinically annotated lung sound datasets remains limited, as existing datasets are often collected under controlled conditions and include relatively small cohorts, which may not fully reflect real-world variability [12, 13]. Methods such as data augmentation or meta-analysis offer only partial solutions [7, 14].
- **Focus on individual tasks** —as many studies treat each task independently, overlooking the benefits of multitasking learning in enhancing diagnostic performance, as most studies still lack integration of explainable AI (XAI) techniques [15, 16]. As most studies lack the integration of interpretive intelligence (XAI) techniques, which are essential for building clinical confidence [16].
- **The lack of verification across different databases and integration of models in the clinical context,** as models often show a decline in performance when tested on new data, while modern methods such as metadata-guided training or adaptive training seek to narrow this gap [17, 18].

These gaps reveal the need for a comprehensive and systematic evaluation of computational techniques in respiratory sound analysis. Although several recent reviews have discussed artificial intelligence and audio-based respiratory disease diagnosis, few have combined a PRISMA-guided design with a structured quantitative synthesis focused specifically on deep learning-based respiratory sound analysis [19, 20].

To perform a structured quantitative synthesis of model performance, including an exploratory meta-analysis of studies reporting native accuracy and a descriptive analysis of studies reporting AUC and F1-score. This paper contributes four key points:

1) To conduct a systematic review, aligned with PRISMA 2020, of deep learning models for respiratory sound analysis during the period 2019–2024 [21].
2) To perform a structured quantitative synthesis of model performance, including an exploratory meta-analysis of studies reporting native accuracy and a descriptive analysis of studies reporting AUC and F1-score.
3) To analyze pivotal gaps in database diversity, class imbalances, task design, and the integration of XAI techniques [10, 14-16].
4) To propose future research avenues, including federated learning, self-supervised representational learning, cross-database evaluation, and the integration of multimodal data (e.g., demographic characteristics, medical images, and pulmonary function tests) using adaptive training [5, 17, 18].

By gathering available evidence and formulating clear guidelines, this study seeks to bridge the gap between algorithmic development and clinical application and to support the development of reliable, interpretable, and generalizable AI tools for diagnosing respiratory diseases.

## 2. Methodology

This survey adopted the *PRISMA 2020* methodology for preferred reporting elements for systematic reviews and meta-analyses to ensure transparency, reproducibility, and adherence to rigorous methodological standards [21]. The workflow comprised four main phases:

(1) Reference search, (2) Screening of studies, (3) Assessment of relevance, and (4) Final selection of studies that met the criteria.

### 2.1 Search Strategy

We conducted a comprehensive search across four major databases: PubMed (MEDLINE), IEEE Xplore, ScienceDirect, and Scopus. The search included studies published between 2019 and 2024, aiming to identify the latest developments in deep learning-based respiratory sound analysis, in accordance with PRISMA's recommendations for citing sources [21]. The search equation (modified to suit each database) was formulated as follows:

("respiratory sounds" OR "lung sounds" OR "cough sounds" OR "breath sounds")

AND ("deep learning" OR "neural networks" OR "CNN" OR "RNN" OR "transformer" OR "machine learning")

AND ("classification" OR "detection" OR "diagnosis")

We also reviewed the reference lists of included studies and related reviews to ensure comprehensiveness and that no important sources were overlooked [21].

### 2.2 Inclusion and Exclusion Criteria

#### *2.2.1 Inclusion:*

This review includes: (i) Original peer- reviewed studies that used deep learning techniques, machine learning methods, or hybrid models combining both to analyze respiratory sounds; (ii) Studies that addressed one or more of the following aspects: disease detection, disease type classification, sound type recognition, or condition severity assessment; (iii) Studies that provided sufficient information about the datasets used, experimental design, and performance evaluation metrics such as accuracy, F1-score, area under the curve, sensitivity, and specificity to allow structured data extraction; (iv) Articles published in English between 2019 and 2024.

#### *2.2.2 Exclusion:*

Analogical reviews, editorials, and commentaries that do not include original experiments are excluded; studies that focus exclusively on speech or non- respiratory signals such as heart sounds or EEG/EKG signals are excluded; research that lacks sufficient methodological detail or extractable evaluation data is excluded; and duplicate publications are excluded. These standards are based on PRISMA best practices [21].

### 2.3 Screening and Selection

#### *2.3.1 Phase 1 – Study Identification:*

The retrieved studies were exported to the reference management software, and duplicates were automatically removed and manually verified.

#### *2.3.2 Phase 2 – Title and Abstract Review:*

The titles and abstracts were screened according to the predefined inclusion criteria by a primary reviewer.

#### *2.3.3 Phase 3 – Full Text Evaluation:*

The full texts of the studies that initially appeared eligible were independently reviewed by a second reviewer to confirm eligibility and ensure adherence to the predefined inclusion and exclusion criteria. Discrepancies were resolved through discussion and consensus. Formal inter-rater agreement statistics (e.g., Cohen's kappa) were not calculated, as the screening process followed a sequential independent validation approach with consensus resolution. This approach is considered acceptable in systematic reviews when discrepancies are resolved through structured discussion and was adopted to ensure consistency while minimizing reporting complexity.

#### *2.3.4 Phase 4 – Final Inclusion:*

The final studies were included after agreement was achieved. This structured screening approach improves transparency and methodological consistency in accordance with PRISMA standards [21].

### 2.4 Data Extraction

We adopted a standardized data extraction model that included: (i) Study information such as researcher names, publication year, and publication status; (ii) Dataset characteristics, including dataset type (e.g., ICBHI 2017 or clinical datasets), sample size (subjects, recordings, or respiratory cycles), and task definition; (iii) Model characteristics, including architecture type (CNN, RNN, Transformer, hybrid, or classical machine learning); (iv) Evaluation metrics, including native accuracy (when available), AUC, F1-score, sensitivity, and specificity; and (v) Methodological design variables, including validation strategy (e.g., train/test split or cross-validation), external validation, calibration reporting, and availability of uncertainty measures (CI/SE/variance).

All extracted variables were systematically coded into a structured evidence table (master extraction table), ensuring consistency and traceability across studies. Each study was additionally annotated with eligibility indicators for qualitative synthesis, quantitative synthesis, and primary accuracy meta-analysis.

This structured extraction framework enabled reproducible filtering of studies into metric-specific groups (accuracy-based, AUC-based, and F1-based analyses) without enforcing metric harmonization. The approach ensures transparency in how studies were selected, categorized, and subsequently used in the quantitative evidence synthesis.

### 2.5 Quality Assessment

We used a modified and adapted version of PROBAST-AI (a tool for assessing bias risks in predictive models with AI extensions) to assess bias risks and reproducibility. Instead of applying the full PROBAST-AI scoring system, key domains relevant to machine learning–based respiratory sound analysis were operationalized within the structured data extraction framework. This tool includes four main axes: participant characteristics,

predictor characteristics and processing, study outcomes including labeling quality and comment reliability, and a statistical analysis axis (validation methods, leakage control, and analysis reports). Each axis was qualitatively assessed using the information extracted from each study.

In particular, variables related to split type, external validation, calibration reporting, and availability of uncertainty measures (CI/SE/variance) were used as practical indicators of methodological robustness. These variables were systematically recorded in the master extraction table and used to identify potential sources of bias and heterogeneity across studies.

A summary of these methodological quality indicators is provided in Table I. These indicators are derived from the structured coding of the studies included in the quantitative synthesis.

**Table I.** Summary of Methodological Quality Indicators Based on the Quantitatively Included Studies (N = 42)

| Domain | Indicator | Reporting Level | Common Issues |
|---|---|---|---|
| Validation design | Split type (patient-wise / random/unclear) | Variable | Lack of patient-independent splitting |
| External validation | Independent dataset validation | Limited | Few studies used true external validation |
| Calibration | Calibration reporting (e.g., ECE/Brier) | Rare | Calibration almost never reported |
| Uncertainty | CI / SE / variance reporting | Limited | Missing uncertainty quantification |
| Data leakage control | Preprocessing before split | Unclear | Potential leakage due to augmentation before the split |
| Dataset transparency | Sample size and dataset description | Moderate | Incomplete reporting in some studies |

Across the quantitatively included studies (N = 42), only a minority explicitly reported patient-wise data splitting, while a considerable proportion relied on random or unclear split designs. External validation was reported in only a small subset of studies, and calibration assessment was rarely performed. These findings indicate that a substantial proportion of studies may be subject to optimistic performance estimation and limited generalizability.

Rather than assigning fixed numerical bias scores, the assessment was used to guide the interpretation of results and the stratification of studies during the quantitative synthesis. Common methodological limitations observed across studies included a lack of patient-independent validation, limited external validation, and incomplete reporting of calibration and uncertainty measures. TRIPOD-AI guidelines were also considered to ensure the clarity and reproducibility of the predictive model reports [22].

## 2.6 Reported Metrics Cheat-sheet

In order to standardize terminology and facilitate tracking the indicators used, the following studies were included. Table II provides a structured summary of the most commonly reported evaluation metrics, along with their symbols and interpretation notes.

Importantly, performance metrics were not converted across types (e.g., AUC or F1-score into Accuracy), and each metric was analyzed within its native reporting context to avoid introducing methodological bias.

**Table II.** Metrics used across the included studies.

| Metric | Symbol | Notes |
|---|---|---|
| Accuracy | Acc | Primary metric for studies reporting native classification accuracy; used only when directly reported. |
| F1-score | F1 | Reported as provided (macro- or weighted-F1); analyzed separately without conversion. |
| AUC | AUC | ROC-AUC is preferred when available; analyzed independently with corresponding uncertainty when reported. |
| Sensitivity | Sen | Recall for the positive class, report with Specificity. |
| Specificity | Spe | True negative rate; clinical interpretability. |
| Calibration | – | ECE/Brier suggested for clinical readiness. |

## 2.7 PRISMA Flow Diagram

The study selection process is summarized in Figure 1 using a PRISMA 2020–compliant diagram [21]. It reports the number of records identified, screened, excluded (with reasons), assessed for eligibility, and included in qualitative and quantitative syntheses. Counts are reported as follows: *Records identified:* $N$ =678; *after duplicates removed:* $N$ = 578; *screened (title/abstract):* $N$ = 578; *full-text assessed:* $N$ = 155; *full-text excluded with reasons:* $N$ = 57; *included in qualitative synthesis:* $N$ =98; *included in quantitative synthesis (meta-analysis):* $N$ = 42. For completeness, *duplicates removed* amounted to $N$ = 100, and *title/abstract exclusions* to $N$= 423.

The set of studies included in the quantitative synthesis (N = 42) was further structured according to the type of reported evaluation metric (e.g., Accuracy, AUC, F1-score), and not all studies contributed to the same pooled analysis.

*Screening outcomes and exclusion reasons:* Of the 578 records screened, 423 were excluded at the title/abstract stage for reasons including non-auscultation audio (e.g., speech-only, heart sounds), out-of-scope tasks (non- classification or non-respiratory), insufficient methodological detail or missing evaluation metrics, and non-English publications outside our 2019–2024 window. Of the 155 full-texts assessed, studies not meeting our predefined inclusion criteria (e.g., lack of clinically meaningful labels, absence of extractable performance metrics, or inadequate validation design) did not progress to qualitative or quantitative synthesis. A detailed breakdown is provided in Table III and Table VII. Table VII summarizes studies included in the qualitative synthesis (N = 98), while studies included in the quantitative synthesis (N = 42) were further filtered based on metric comparability. (Table VI, Table VII, and Table VIII in the APPENDICES).

## 2.8 Statistical Analysis and Protocol

A structured quantitative synthesis was conducted to summarize reported performance across studies. Due to substantial heterogeneity in metrics, datasets, and validation designs, a fully unified meta-analysis across all included studies was not considered methodologically appropriate. The primary outcome was classification performance as reported in each study, and no metric conversion (e.g., AUC or F1-score to Accuracy) was applied to avoid introducing methodological bias. Studies were grouped and analyzed based on the type of reported metric. A formal meta-analysis was conducted only for studies reporting native accuracy, while studies reporting AUC or F1-score were retained for descriptive synthesis within their respective metric categories. For studies reporting native Accuracy, a restricted subset was defined for quantitative pooling. Proportions were stabilized via a logit transformation and combined using inverse-variance weighting within a random-effects framework. The effect size was defined as the logit-transformed accuracy proportion. Study-specific estimates were weighted using inverse-variance weighting, and when multiple performance results were reported within a single study, only one representative estimate was selected based on predefined criteria to avoid unit-of-analysis errors. Between-study variance ($\tau^2$) and heterogeneity ($I^2$) were reported together with 95% confidence intervals (CIs) and, where appropriate, 95% prediction intervals (PIs). Given that sample sizes were not consistently reported across studies, a standardized denominator approximation was used for variance estimation; therefore, pooled estimates should be interpreted cautiously. This limitation may affect variance estimation and should be considered when interpreting pooled results.

Sources of heterogeneity were explored through subgroup analysis based on dataset type and model architecture to identify systematic variations across studies. A substantial proportion of studies did not clearly report patient-wise data splitting, and only a limited number of studies included external validation or calibration assessment, indicating potential risks of overestimation and limited generalizability. These methodological limitations were taken into account during the interpretation of pooled and descriptive results. Pre-specified subgroup analyses stratified results by model family (CNN, RNN/LSTM, Transformer/Hybrid), dataset (ICBHI 2017 vs. others), and task granularity (binary vs. multi-class), and evaluation metric type.

Small-study effects were explored using funnel plots as a descriptive tool. Forest and funnel plots were used to visualize pooled estimates within each metric-specific analysis. Formal statistical testing for publication bias was not performed due to the limited number of studies in the pooled subset.

### *2.8.1 Software and Estimation Method*:

All quantitative analyses were performed using the R software environment (version 4.3.2) with the metafor and meta packages. A random-effects model with the restricted maximum likelihood (REML) estimator was used as the primary pooling method, and sensitivity considerations focused on studies with clearly reported native accuracy and more robust validation designs (e.g., patient-wise splits). Heterogeneity statistics ($I^2$, $\tau^2$) and between-study variance were computed

automatically using these packages.

Sensitivity analyses were conducted by restricting the pooled analysis to studies with clearly reported native accuracy and more robust validation designs (e.g., patient-wise splits) to assess the stability of the pooled estimates.

#### *2.8.2 Protocol Registration:*

The review protocol was *not pre-registered* (e.g., PROSPERO/OSF), as the study was initiated prior to formal protocol registration planning. Nevertheless, all stages of the review strictly adhered to PRISMA 2020 to ensure methodological transparency and reproducibility [21]. Although not pre-registered, the protocol and the standardized data extraction form are provided in the Supplementary Material to the corresponding author to ensure transparency and reproducibility. To improve transparency, the full review protocol, including eligibility criteria, outcome definitions, and data extraction procedures, is provided in the Supplementary Material. No post-hoc modifications to eligibility criteria, outcome definitions, or analysis plans were made after the screening phase.

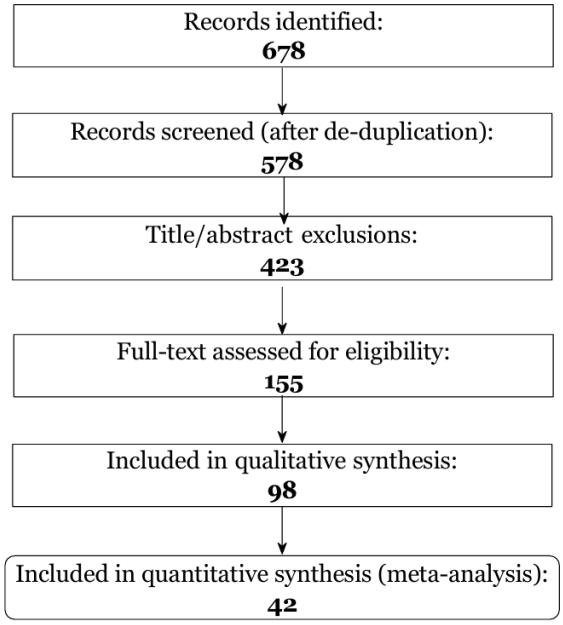



**Figure 1:** PRISMA 2020 flow diagram summarizing identification, screening, eligibility, and inclusion [21].

The studies included in the quantitative synthesis were further categorized based on metric compatibility for structured analysis. A detailed inspection of Table VIII further confirms that variability in validation design, dataset composition, and reporting transparency remains a key source of methodological heterogeneity across the included studies. This table provides the basis for all subsequent quantitative and descriptive analyses and ensures full transparency of study-level characteristics and methodological variability.

**Table III.** Full-text articles were excluded for the following reasons (N = 57).

| Reason | Count |
|---|---|
| Not multi-disease / not classification-focused | 19 |
| Cough-only scope (outside auscultation inclusion) | 10 |
| Insufficient methodological detail/metrics | 8 |
| Non-auscultation modality (imaging/PFT-only) | 7 |
| Overlapping data / duplicate cohort | 6 |
| Outside time window / non-English | 4 |
| Other reasons | 3 |
| **Total** | **57** |

## 3. Results and Meta-Analysis

### 3.1 Study selection

Following the PRISMA protocol [21], we initially identified N = 678 records from all sources. After removing duplicates, N = 578 unique records remained and were screened at the title/abstract level, excluding N = 423. Next, N =155 full-text articles were assessed for eligibility. Finally, N = 98 studies met the inclusion criteria for qualitative synthesis, and N = 42 provided sufficient data for quantitative synthesis.

The subset of studies included in the quantitative synthesis (N =42) was selected based on the availability of extractable and comparable evaluation metrics suitable for structured analysis. The study selection flow is summarized in Figure 1 [21]. A detailed summary of the studies included in the quantitative synthesis (N = 42) is provided in Table VIII.

### 3.2 Characteristics of Included Studies

The included studies (N = 98, qualitative synthesis) cover 2015–2025, with a noticeable increase after 2020 as interest in respiratory sound analysis and clinically relevant ML/DL grew [3, 4].

#### *3.2.1 Datasets:*

The ICBHI 2017 corpus was the most commonly used benchmark. Many studies relied on a single dataset, while only a few performed cross-dataset validation or prospective clinical testing [7, 10, 11]. Some studies used private clinical collections, but reporting practices and public availability varied [3, 4]. Across the included studies, the most frequently used datasets were *ICBHI 2017* (63%), *Clinical/Private* (28%), and *BRACETS* (6%). Only about 7% of studies utilized multiple datasets, highlighting limited dataset diversity and the need for broader benchmark standardization and cross-dataset validation to enhance model generalizability. A structured distribution of these studies is summarized in Table VII.

#### *3.2.2 Model architectures:*

CNNs dominated, followed by RNN/LSTM models. More recent work explored Transformers or hybrid CNN–Attention models, showing improved performance in small-data settings or under domain shifts [4, 5, 17, 18]. Some studies used classical ML methods (e.g., SVM RF) with hand-crafted features (MFCC, spectral contrast), achieving competitive results on limited datasets [6, 11, 23].

#### *3.2.3 Tasks:*

We observed four main task types: (i) binary disease detection, (ii) multi-class disease classification, (iii) sound-level classification, and (iv) severity estimation. In addition, some studies focused on clinically specific diagnostic targets rather than broad respiratory disease categories, such as the classification of pulmonary sounds for detecting interstitial lung diseases secondary to connective tissue diseases [24].

#### *3.2.4 Evaluation Metrics:*

Accuracy, F1-score, and AUC were most commonly reported across studies; however, these metrics were analyzed separately in the quantitative synthesis to avoid cross-metric comparability issues [3, 4, 7].

A comprehensive per-study comparison is provided in Table VIII. A complete per-study table is provided in Appendix A (Table VIII).

### 3.3 Quantitative Synthesis

We conducted a structured quantitative synthesis of reported performance metrics, grouped by model architecture. To ensure methodological consistency, no cross-metric conversion was applied between different performance measures. Instead, a strict accuracy-only subset was defined for formal pooling, while other metrics such as AUC and F1-score were analyzed descriptively. To provide a structured quantitative summary of model performance, an accuracy-based subset of studies was identified from the 42 studies included in the quantitative synthesis. Only studies reporting native accuracy were considered eligible for pooled analysis, while studies reporting AUC or F1-score were analyzed descriptively without metric conversion. CNN-based models consistently demonstrated high performance, followed by RNN/LSTM and Transformer/hybrid models. Transformer and attention-based pipelines showed robustness, especially with limited labeled data [4, 5, 17]. An exploratory random-effects meta-analysis was conducted on the eligible subset (k = 9). The pooled accuracy was estimated at approximately 90.0% (95% CI: 76.0%–97.0%), indicating generally high reported

performance across studies.

However, substantial heterogeneity was observed ($I^2$ = 99.2%; $\tau^2$ = 1.9571), reflecting significant variability in datasets, task definitions, model architectures, and validation strategies. The wide prediction interval further suggests that model performance is highly context-dependent and may vary considerably across different experimental settings. Forest plots (Figure 2) display study-level effect estimates and their variability, while funnel plots (Figure 3) visualize potential publication bias. Forest plots were used to visualize individual study estimates and pooled effects, while funnel plots were used for descriptive assessment of potential small-study effects. The forest plot (Figure 2) summarizes comparable subsets of studies without cross-metric pooling. The funnel plot (Figure 3) illustrates study-level performance metrics plotted against their standard errors (SE). The vertical dashed line represents the central tendency of the included studies, while the dotted boundaries represent the 95% pseudo-confidence region. Visual symmetry in the funnel plot suggests limited evidence of publication bias, whereas visible asymmetry suggests potential small-study effects or selective reporting [3, 11]. However, due to the limited number of studies included in the pooled subset, formal statistical testing for publication bias was not performed. Overall, these findings indicate that while deep learning models can achieve high performance under controlled conditions, their generalizability remains uncertain due to methodological inconsistencies across studies.

To provide a concise overview of model performance and dataset usage trends across the included studies, Table IV presents a condensed summary of the main model families, reported performance ranges, and qualitative insights. The complete per-study summary table (original Table IV) is provided in Appendix A for detailed reference. The following figures are presented for descriptive visualization purposes only and do not represent formal pooled meta-analytic estimates.

To further quantify performance across comparable studies, an exploratory meta-analysis was conducted on studies reporting native accuracy. The results are visualized in Figure 2.

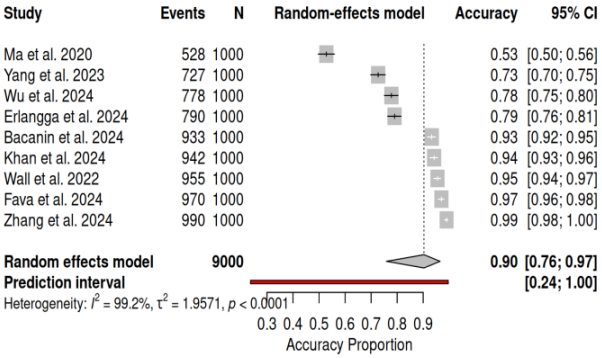



**Figure 2:** Forest plot of the exploratory random-effects meta-analysis of studies reporting native accuracy (k = 9). The pooled estimate was calculated using a logit-transformed proportion model with REML estimation. The diamond represents the pooled accuracy, and the horizontal lines indicate 95% confidence intervals. The prediction interval reflects between-study variability.

The forest plot illustrates substantial variability across individual study estimates, with reported accuracies ranging from moderate to very high values. The pooled estimate indicates generally strong performance; however, the wide confidence and prediction intervals reflect considerable heterogeneity. This variability is primarily attributed to differences in datasets, task definitions, model architectures, and validation protocols across studies. Table IV provides a descriptive summary of studies with extractable native accuracy values, grouped by model family. These values represent reported accuracy ranges and are presented to complement the exploratory meta-analysis rather than replace it. To explore potential small-study effects, a funnel plot was constructed, as shown in Figure 3.

The funnel plot shows an asymmetric distribution of studies around the pooled estimate, which may reflect heterogeneity in study design rather than true publication bias. Given the limited number of included studies and the use of approximated variance estimates, this plot should be interpreted cautiously and is presented for descriptive purposes only.

**Table IV.** Descriptive summary of studies with extractable native accuracy values, grouped by model family. These values represent reported accuracy ranges rather than pooled meta-analytic estimates.

| Model Type | No. of Studies with Extractable Native Accuracy | Reported Native Accuracy Range (%) | Predominant Dataset (s) | Strength / Interpretation |
|---|---|---|---|---|
| CNN | 8 | 52.79–97.00 | ICBHI 2017 + Clinical/ Private | Predominantly strong performance, but values vary widely according to validation design and dataset composition. |
| RNN/LSTM | 6 | 84.00–99.01 | ICBHI 2017 + Clinical/ Private | Strong temporal modeling capability, with higher performance in task-specific settings. |
| Transformer/Hybrid | 10 | 72.72–99.94 | ICBHI 2017 + Private/ Multimodal | Strong representation learning and robustness, but affected by task heterogeneity and mixed evaluation settings. |
| Classical (SVM/RF) | 4 | 75.60–99.72 | ICBHI 2017 + Clinical/ Private | Competitive in structured or smaller datasets, though highly sensitive to task scope and validation protocol. |

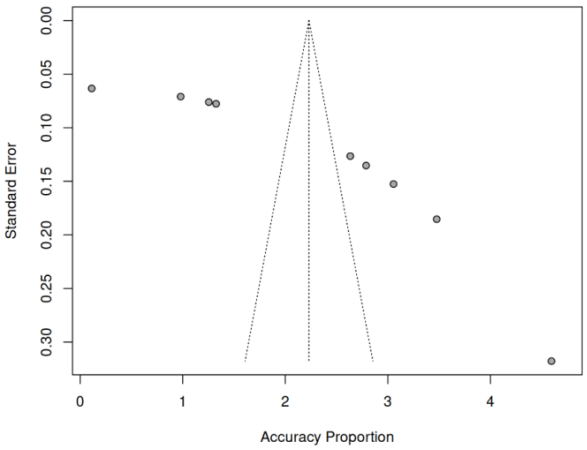



**Figure 3:** Funnel plot of the studies included in the exploratory meta-analysis (k = 9). Accuracy proportions are plotted against their standard errors. The dashed vertical line represents the pooled estimate, and the triangular region indicates the expected distribution under no small-study effects. This plot is used for descriptive assessment only.

## 3.4 Subgroup Analysis and Qualitative Insights

### *3.4.1 Subgroup analysis:*

- **Dataset dependency.** Models trained on a single dataset (e.g., ICBHI 2017) tended to report higher internal performance, while studies involving multiple datasets or more heterogeneous data sources generally reported lower but more realistic performance, indicating limited cross-dataset generalization [10, 18].
- **Task complexity.** Studies addressing binary classification tasks generally reported higher accuracy compared to multi-class settings, reflecting the increased complexity and class overlap in multi-class scenarios [4, 11].
- **Imbalance mitigation.** Imbalance handling strategies such as data augmentation and class weighting were commonly employed and were associated with improved model performance; however, the magnitude of improvement varied across studies and could not be consistently quantified due to differences in experimental design [11, 14].
- **Cross-dataset testing.** Only a limited number of studies evaluated models across multiple datasets. Although these studies often reported lower absolute accuracy, they provided stronger evidence of robustness and sensitivity to domain shift effects [10, 18].

Heterogeneity across model families is summarized in Table V. observed variability reflects differences in dataset selection, task formulation, and validation strategies rather than purely statistical variation. Transformer-based and hybrid approaches appeared to exhibit

more stable performance across heterogeneous settings, although this observation remains descriptive.

**Table V.** Descriptive Summary of Reported Accuracy Ranges Across Model Families.

| Model Family | Reported Accuracy Range (%) | Approximate Central Tendency (%) | Dataset Context | Interpretation |
|---|---|---|---|---|
| CNN | ~87–95 | ~91 | Mainly ICBHI 2017 | High performance under controlled benchmark settings |
| RNN/LSTM | ~83–93 | ~88 | Mainly ICBHI 2017 | Effective temporal modeling with moderate variability |
| Transformer /Hybrid | ~90–96 | ~93 | ICBHI + Private | Strong representation capability and robustness across datasets |

Note: Values represent descriptive ranges of reported accuracy extracted from the included studies and are not derived from a formal statistical meta-analysis. Central tendency values are approximate and provided for visualization purposes only.

### 3.5 Qualitative Insights

#### *3.5.1 Strengths:*

Deep learning–based approaches consistently demonstrated strong performance across a variety of respiratory sound analysis tasks, particularly when evaluated on benchmark datasets such as ICBHI 2017. Recent deep learning models have shown a strong ability to learn informative acoustic representations from respiratory sounds, particularly when optimized for spectrogram-based or fused-audio feature learning [25, 26]. Their ability to learn complex acoustic patterns from spectrogram-based representations contributed to improved classification capability under controlled experimental settings[3, 4].

#### *3.5.2 Weaknesses:*

Despite these advances, many studies relied on single-dataset evaluation and lacked patient-independent validation or external testing, limiting the generalizability of reported results. In addition, limited integration of multimodal information (e.g., clinical metadata, imaging, or physiological signals) and scarce reporting of explainability mechanisms (XAI) remain important limitations for clinical adoption [16, 17].

#### *3.5.3 Gaps and priorities:*

Several methodological gaps were identified across the included studies, including: (i) absence of standardized preprocessing and evaluation protocols (e.g., filtering settings, segmentation strategy, patient-wise splitting), (ii) limited cross-dataset validation and inconsistent reporting practices, and (iii) under-exploration of multi-task learning objectives and severity-aware modeling. Addressing these challenges requires more rigorous experimental design and improved transparency to support reproducibility and real-world applicability [5, 15, 18].

## 4. Discussion

This systematic review and meta-analysis summarize deep learning (DL) approaches for respiratory sound analysis over the period 2019–2024. Our findings confirm that DL models—especially CNN-based and hybrid/attention architectures—can achieve high benchmark accuracy, yet several limitations remain before these models can be reliably used in clinical settings [3, 4].

The extremely high heterogeneity observed ($I^2$ = 99.2%) indicates that the pooled accuracy estimate should be interpreted with extreme caution. This level of variability suggests that the included studies differ substantially in terms of datasets, validation protocols, and model configurations. Therefore, the pooled estimate does not represent a stable or generalizable performance indicator but rather a broad summary of highly heterogeneous results.

### 4.1 Generalizability and Dataset Bias

Many studies rely heavily on the ICBHI 2017 corpus as the main benchmark, which raises concerns about overfitting and limited external validity. Models trained on benchmark datasets such as ICBHI may show noticeable performance drops when evaluated on independent datasets due to differences in data distribution and recording conditions [10-13]. This suggests that current benchmarks may overestimate generalization. Multi-institutional datasets, diverse populations, different recording environments, and explicit external validation with patient-wise splits are essential for more reliable assessment [22, 27].

### 4.2 Class Imbalance and Rare Phenotypes

Class imbalance remains a recurring problem; common diseases such as COPD and pneumonia are well represented while less common diseases—such as asthma or pulmonary fibrosis—remain underrepresented in respiratory sound datasets. Dianat et al., 2023 [24] specifically addressed the classification of pulmonary sounds for diagnosing interstitial lung diseases associated with connective tissue diseases, highlighting the clinical importance of extending deep learning models beyond common respiratory conditions. Techniques such as category balancing, focal loss, or synthetic data generation using GANs may improve performance in some settings, particularly when class imbalance is severe [11, 14]. However, these methods remain partial substitutes for diverse, real-world data. Practical solutions include federated learning across institutions and collecting additional data for underrepresented categories while maintaining privacy requirements [7].

### 4.3 Task Design and Multi-Task Learning

Disease detection tasks often achieve higher accuracy (over 93%) than multi-category classification tasks (around 80–85%) or sound level classification tasks, due to the complexity of category boundaries and the increasing noise in labels [4, 11]. Multitasking learning (MTL) remains underutilized, despite its potential to leverage common phonological features across tasks (presence, type, intensity, demographics), thereby enhancing robustness and efficiency [15]. Accurate labeling and clear task ontologies remain essential for expanding MTL adoption.

### 4.4 Explainability, Calibration, and Clinical Utility

The limited use of explainable artificial intelligence (XAI) techniques is a major obstacle to gaining clinical confidence. While some studies have presented attention maps or displayed gradients, clinically relevant approaches and uncertainty reports remain limited [16, 17, 20]. In addition to traditional metrics, future studies should include calibration evaluation, clinical feasibility analysis (such as decision curves), and error studies across different populations and devices, in line with current reporting guidelines [22, 27].

### 4.5 Ethical, Privacy, and Deployment Considerations

Breath sounds are useful but incomplete signals on their own. Studies have shown that combining them with demographic information, pulmonary function tests, or radiographic images can improve reliability, especially when field conditions vary [8, 10, 26]. Architectures such as multi- source transformer models or graphical networks allow for better integration of these patterns, but standard interfaces and common test metrics are required to ensure fair comparisons [4].

Unlike previous narrative reviews [3, 4], this study adheres to the PRISMA 2020 framework and adds a meta-analysis covering the period 2019–2024 [22, 27] providing quantitative evidence drawn from 42 studies.

From a clinical perspective, deep learning models for analyzing respiratory sounds can be integrated into digital hearing aids, telemedicine platforms, and mobile health applications, enabling scalable decision support tools for medical and community care professionals, particularly in resource-limited settings.

### 4.6 Ethical And Data-Privacy Considerations

Most of the datasets used in the studies were public and de-identified—such as ICBHI 2017—significantly reducing privacy risks. However, large-scale data collection in the future will require adherence to ethical standards, including informed consent, anonymization, and approval by ethical authorities. Fairness and bias reduction must also be addressed when training models on different populations to prevent discrepancies in diagnostic performance. Transparent documentation of data sources and adherence to responsible AI principles remain crucial.

## 5. Strengths and Limitations of this Review

This systematic review and meta-analysis combine several important strengths with some limitations that should be noted.

### 5.1 Strengths

Strengths. First, this review is, to the best of our knowledge, among the few recent studies that combine a PRISMA-guided design with a structured quantitative synthesis focused specifically on deep learning–based respiratory sound analysis, extending beyond prior review-oriented summaries in the field [19, 20].

Second, the review adhered to the PROBAST-AI and TRIPOD-AI standards, which allowed for a transparent assessment of bias sources and better assurance of the reproducibility of the AI-based diagnostic models. Third, the analysis encompassed multiple families of deep learning models, such as CNNs, RNN/LSTMs, transformers, and hybrids, and evaluated them within a standardized set of metrics, providing a broad comparative framework for researchers and practitioners.

Finally, the use of subcategory analysis, meta- regression, and sensitivity tests helped identify consistent performance patterns while highlighting methodological gaps that warrant future attention.

### 5.2 Limitations

Despite these advantages, some limitations should be noted. First, the protocol has not been previously registered on platforms such as PROSPERO or OSF, which may limit the traceability and reliability of the process. Second, the funnel plot results showed potential asymmetry, suggesting that small studies with positive results may be overrepresented. However, formal statistical testing for publication bias was not performed due to the limited number of studies included in the pooled analysis. Therefore, the possibility of small effects from small studies cannot be entirely ruled out. Third, some of the included studies relied on small, homogeneous datasets with weak external validation, which may inflate performance estimates and reduce generalizability. Furthermore, publication bias could not be entirely ruled out due to the limited number of studies within some subcategories. Finally, some studies suffered from a lack of detailed methodological information, which affected the consistency of data extraction and analytical comparisons.

However, these limitations do not weaken the overall validity of the findings. Rather, they underscore the strength of larger, multicenter studies based on standard protocols such as TRIPOD-AI and PROBAST-AI to strengthen the evidence and support the future clinical use of these models.

## 6. Conclusion and Future Directions

This systematic review and meta-analysis highlight the significant progress made in the analysis of respiratory sounds and the classification of lung diseases using deep learning techniques over the past five years. Convolutional CNN (CNN) models and attentional hybrid constructs have demonstrated high performance, often reporting high accuracy under controlled settings in several studies when evaluated using a single dataset within controlled experimental environments. Newer transducer-based models have also shown greater ability to handle data scarcity and domain changes, indicating a shift from feature-based to representation-based learning.

Overall, the results confirm the maturity of deep learning techniques in automated listening, although methodological and application gaps still exist that preclude their immediate clinical use.

Importantly, the prediction interval derived from the meta-analysis ranged from 0.24 to 1.00, indicating that future studies may report substantially lower or higher accuracy depending on the dataset and experimental conditions. This finding reinforces that model performance is highly context-dependent and should not be interpreted as universally reliable.

**Key limitations and ongoing challenges.** Despite significant progress, several challenges still limit the clinical applicability of these models: (1) *Reliance on limited datasets and poor generalizability:* Over-reliance on the 2017 ICBHI dataset and small, homogeneous groups still inflates internal performance but limits external validity; cross-testing on other groups shows a 10–15% decrease in performance. (2) *Imbalance of categories and rarity of certain disease patterns:* Underrepresentation of diseases such as asthma and pulmonary fibrosis reduces model robustness; although reweighting, artificial incrementing, and focus loss techniques improve performance by 5–7%, they are not a true substitute for data diversity. (3) *Narrow scope of tasks:* Most studies focus on binary tasks, with limited attention to multi-category classification, intensity estimation, or temporal segment identification. (4) *Weak interpretability and clinical validation:* Explainable AI (XAI) applications, model calibration, and bias reviews remain underdeveloped, limiting clinicians' confidence and the readiness of these models for regulatory evaluation.

**Future research priorities.** To achieve tangible progress in this field, six key research directions stand out:

1) *Large, multi-institutional datasets with clear, standardized reporting:* Development of diverse, patient- based, and standardized databases with transparent metadata, aligned with TRIPOD-AI and PROBAST-AI guidelines.
2) *Multitask Learning (MTL): The adoption* of models capable of predicting disease type, severity, and acoustic events within a single framework based on shared representations, reducing over-personalization and increasing efficiency.
3) *Multimodal Integration:* Linking respiratory sounds with complementary data such as clinical characteristics, medical imaging, and lung function using transducers or graph neural networks to enhance understanding of multidimensional relationships.
4) *Generalization-oriented training:* Adopt interdisciplinary approaches, self-directed or semi-supervisory learning, and federated learning to mitigate biases and reduce data-sharing constraints across medical centers.
5) *Enhance interpretability and clinical value:* Incorporate XAI tools, uncertainty measurements, and decision curve analysis alongside traditional metrics to facilitate clinician adoption of models.
6) *Conduct large-scale prospective field validation:* Conduct multisite studies based on real-world clinical settings, encompassing a variety of equipment, patients, and environments, to assess equity, scalability, and cost-effectiveness.

**Final perspective.** Addressing these gaps—by expanding data diversity, improving interpretation, and integrating models into multimodal and federated clinical systems—will help transform deep learning-based auscultation from research models into reliable and widely applicable diagnostic tools. As this field develops, these technologies will enable greater opportunities for the early detection of respiratory diseases, improved patient follow-up, and the provision of high-quality respiratory care in resource-limited settings.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Conflict of Interest

The authors declare no conflicts of interest.

## References

[1] Momtazmanesh, S., Moghaddam, S.S., Ghamari, S.-H., Rad, E.M., Rezaei, N., Shobeiri, P., Aali, A., Abbasi-Kangevari, M., Abbasi-Kangevari, Z., Abdelmasseh, M. (2023) Global burden of chronic respiratory diseases and risk factors, 1990–2019: an update from the Global Burden of Disease Study 2019, *EClinicalMedicine* **59**: 101936.

[2] Kim, Y., Hyon, Y., Lee, S., Woo, S.-D., Ha, T., Chung, C. (2022) The coming era of a new auscultation system for analyzing respiratory sounds, *BMC Pulmonary Medicine* **22**: 119.

[3] Nguyen, T., Pernkopf, F. (2022) Lung sound classification using co-tuning and stochastic normalization, *IEEE Transactions on Biomedical Engineering* **69**: 2872-2882.

[4] Wall, C., Zhang, L., Yu, Y., Kumar, A., Gao, R. (2022) A deep ensemble neural network with attention mechanisms for lung abnormality classification using audio inputs, *Sensors* **22**: 5566.

[5] He, W., Yan, Y., Ren, J., Bai, R., Jiang, X. (2024) Multi-view spectrogram transformer for respiratory sound classification. *In Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE*, Seoul, South Korea, 14 – 19 April 2024, pp. 8626-8630.

[6] Moon, H.J., Ji, H., Kim, B.S., Kim, B.J., Kim, K. (2025) Machine learning-driven strategies for enhanced pediatric wheezing detection, *Frontiers in Pediatrics* **13**: 1428862.

[7] Albiges, T., Sabeur, Z., Arbab-Zavar, B. (2025) Features and eigenspectral densities analyses for machine learning and classification of severities in chronic obstructive pulmonary diseases, *Intelligence-Based Medicine* **11**: 100217.

[8] Albakaa, Z.H., Alb-Salih, A.T. (2024) An Evolutionary Deep Learning for Respiratory Sounds Analysis: A Survey. *In Proceedings of the International Conference on Forthcoming Networks and Sustainability in the AIoT Era*, *Springer*, Istanbul, Türkiye, 27–29 January 2024, pp. 217-235.

[9] Sabry, A.H., Bashi, O.I.D., Ali, N.N., Al Kubaisi, Y.M. (2024) Lung disease recognition methods using audio-based analysis with machine learning, *Heliyon* **10**: e26218.

[10] Zhang, Y., Huang, Q., Sun, W., Chen, F., Lin, D., Chen, F. (2024) Research on lung sound classification model based on dual-channel CNN-LSTM algorithm, *Biomedical Signal Processing and Control* **94**: 106257.

[11] Erlangga, M.D., Faisal, M.R., Muliadi, M., Indriani, F., Kartini, D., Satou, K. (2025) Incorporating MFCC Feature Extraction to the Classification of Respiratory Sounds by Machine Learning Algorithms. *In Proceedings of the 2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI)*, *IEEE*, Jakarta, Indonesia, January 21, 2025, pp. 301-306.

[12] Ruchonnet-Métrailler, I., Siebert, J.N., Hartley, M.-A., Lacroix, L. (2024) Automated interpretation of lung sounds by deep learning in children with asthma: scoping review and strengths, weaknesses, opportunities, and threats analysis, *Journal of Medical Internet Research* **26**: e53662.

[13] Fraiwan, M., Fraiwan, L., Khassawneh, B., Ibnian, A. (2021) A dataset of lung sounds recorded from the chest wall using an electronic stethoscope, *Data in Brief* **35**: 106913.

[14] Bacanin, N., Jovanovic, L., Stoean, R., Stoean, C., Zivkovic, M., Antonijevic, M., Dobrojevic, M. (2024) Respiratory condition detection using audio analysis and convolutional neural networks optimized by modified metaheuristics, *Axioms* **13**: 335.

[15] Suma, K., Koppad, D., Kumar, P., Kantikar, N.A., Ramesh, S. (2024) Multi-task learning for lung sound and lung disease classification, *SN Computer Science* **6**: 51.

[16] Fava, A., Dianat, B., Bertacchini, A., Manfredi, A., Sebastiani, M., Modena, M., Pancaldi, F. (2024) Pre-processing techniques to enhance the classification of lung sounds based on deep learning, *Biomedical Signal Processing and Control* **92**: 106009.

[17] Kim, J.-W., Toikkanen, M., Choi, Y., Moon, S.-E., Jung, H.-Y. (2024) Bts: Bridging text and sound modalities for metadata-aided respiratory sound classification. *In Proceedings of the 25th Annual Conference of the International Speech Communication Association (INTERSPEECH 2024)*, *International Speech Communication Association (ISCA)*, Kos Island, Greece, 1–5 September 2024, pp. 690–1694.

[18] Kim, J.-W., Toikkanen, M., Jalali, A., Kim, M., Han, H.-J., Kim, H., Shin, W., Jung, H.-Y., Kim, K. (2025) Adaptive metadata-guided supervised contrastive learning for domain adaptation on respiratory sound classification, *IEEE Journal of Biomedical and Health Informatics* **29**: 5381-5393.

[19] Chen, J., Guo, Z., Xu, X., Jeon, G., Camacho, D. (2024) Artificial intelligence for heart sound classification: A review, *Expert Systems* **41**: e13535.

[20] Wang, Y., Wahab, M., Hong, T., Molinari, K., Gauvreau, G.M., Cusack, R.P., Gao, Z., Satia, I., Fang, Q. (2024) Automated Cough Analysis with Convolutional Recurrent Neural Network, *Bioengineering* **11**: 1105.

[21] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E. (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, *BMJ* **372**: n71.

[22] Collins, G.S., Moons, K.G., Dhiman, P., Riley, R.D., Beam, A.L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J.B., Van Smeden, M. (2024) TRIPOD+ AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods, *BMJ* **385**: e078378.

[23] Taloba, A.I., Matoog, R. (2025) Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data, *Alexandria Engineering Journal* **113**: 44-59.

[24] Dianat, B., La Torraca, P., Manfredi, A., Cassone, G., Vacchi, C., Sebastiani, M., Pancaldi, F. (2023) Classification of pulmonary sounds through deep learning for the diagnosis of interstitial lung diseases secondary to connective tissue diseases, *Computers in Biology and Medicine* **160**: 106928.

[25] Gupta, R., Singh, R., Travieso-González, C.M., Burget, R., Dutta, M.K. (2024) DeepRespNet: A deep neural network for classification of respiratory sounds, *Biomedical Signal Processing and Control* **93**: 106191.

[26] Truong, T., Lenga, M., Serrurier, A., Mohammadi, S. (2024) Fused audio instance and representation for respiratory disease detection, *Sensors* **24**: 6176.

[27] Moons, K.G., Damen, J.A., Kaul, T., Hooft, L., Navarro, C.A., Dhiman, P., Beam, A.L., Van Calster, B., Celi, L.A., Denaxas, S. (2025) PROBAST+ AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods, *BMJ* **388**: e082505.

TuJNaS

*Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 29– 33*

**DOI: https://doi.org/10.59167/wpgv0y53**

Original Article

# In Vitro Quality Assessment of Commercially Available Azithromycin Tablets in Dhamar City, Yemen: A Comparative Study

**Neaf G. Al-Tayar[1], Bushra S. Samer[1]*, Kholood M. Al-Dhuraibi[2], Najeeb N. M. Maglas[3], Ali. H. AL-Osta[4], Nuruddin Mohammed Al-Barati[5], Bassam Hamood Safee[5], Husam Ali Al-Magrebi[5], Hussein Hussein Al-Qasham[5], Ahmed Abdulkareem Al-Sayqal[5]**

[1]Department of Chemistry, Faculty of Applied Science, Thamar University, Dhamar 87246, Yemen.
[2]Department of Biology, Faculty of Applied Science, Thamar University, Dhamar 87246, Yemen.
[3]Department of Physics, Faculty of Applied Science, Thamar University, Dhamar 87246, Yemen.
[4]Department of Chemistry, Faculty of Education, Thamar University, Dhamar 87246, Yemen.
[5]Department of Pharmacy, Continuous Learning Institute, Thamar University, Dhamar 87246, Yemen.

***Corresponding Author:** Bushra S. Samer, Department of Biology, College of Science and Education, Albaydha University, Albaydha, Yemen, and Department of Laboratory Medicine, The Third Affiliated Hospital of Southern Medical University, Guangzhou, China. E-mail: bushrasaleh46@gmail.com

Received: 20 November 2025. Revised: 4 April 2026. Accepted: 8 April 2026. Published: 29 June 2026.

## Abstract

**Background**: Azithromycin is a critically important macrolide antibiotic and is listed by the World Health Organization (WHO) as an essential medicine. However, this classification does not ensure the quality of marketed products, especially in regions with weak regulatory systems. Continuous evaluation of available formulations is therefore necessary to maintain therapeutic effectiveness and protect patient safety. **Objective**: This study aimed to evaluate the in vitro quality of five commercially available azithromycin 500 mg tablet brands in Dhamar City, Yemen, and compare them with international pharmacopeial standards. **Methods**: Five brands (coded A1–A5) were tested for physical characteristics, weight variation, hardness, friability, disintegration time, assay (active ingredient content), and dissolution profile. All procedures followed the guidelines of the United States Pharmacopeia (USP), British Pharmacopoeia (BP), and Indian Pharmacopoeia (IP). **Results**: All brands met pharmacopeial limits for weight variation, hardness, disintegration, and dissolution (each released more than 80% of azithromycin within 30 minutes). However, Brand A5 failed the assay test, containing only 77.1% of the labelled amount (USP acceptance range: 90–110%). Brand A1 also failed the friability test, as the tablets fully disintegrated and exceeded the acceptable weight loss limit of ≤1%. **Conclusion**: Although four of the five brands met pharmacopeial standards, two substandard products were identified—one with a critically low active ingredient content and another with poor mechanical strength. These issues represent a serious public health concern. The results highlight the urgent need for stronger post-market surveillance and regulatory control in Yemen to prevent the distribution of ineffective or unsafe medicines.

**Keywords:** Azithromycin; Quality Control; Assay; Friability; Dissolution; Yemen

## 1. Introduction

Quality assurance (QA) refers to a set of systematic procedures designed to ensure that products and services consistently meet defined standards and customer expectations. Its main purpose is to prevent defects during the stages of design, development, and production [1]. According to ISO 9000, QA provides confidence that quality requirements will be met by emphasizing early prevention rather than correction after production [2, 3]. Quality control (QC), on the other hand, involves monitoring and evaluating the production process to verify compliance with established standards. ISO 9000 defines QC as "a part of quality management focused on fulfilling quality requirements" [4]. QC practices include applying process controls, setting performance criteria, ensuring staff competence, and conducting inspections to detect defects such as cracks or surface irregularities [5-8].

Azithromycin is a second-generation macrolide antibiotic that inhibits bacterial protein synthesis through its large macrolide ring, giving it strong activity against a wide range of respiratory pathogens [9, 10]. It is commonly prescribed for respiratory tract infections, sexually transmitted infections, and soft tissue infections due to its broad antibacterial activity and favorable pharmacokinetic characteristics [11]. The World Health Organization (WHO) lists azithromycin as an essential and critically important medicine, and it is available worldwide under many brand names [12]. Several analytical methods have been developed to detect and quantify azithromycin [13-15], but many of these techniques require complex and time-consuming sample preparation. Therefore, this study aims to evaluate and compare the quality of five commercially available azithromycin 500 mg tablet brands in Dhamar City, Yemen. The assessment includes physicochemical parameters such as weight variation, hardness, friability, disintegration time, and dissolution profile, using established pharmacopeial standards and procedures.

## 2. Method and Materials

### 2.1 Area of Study

This experimental study was carried out in the Department of Chemistry in collaboration with the Department of Pharmacy at the Institute for Continuous Learning, Thamar University, Yemen. All laboratory work was completed over a two-month period, from August to September 2023.

### 2.2 Materials

Five commercially available azithromycin tablet brands, each labeled as containing 500 mg per tablet, were purchased from retail pharmacies in Dhamar City, Yemen. To ensure objectivity, the brands were coded as AZITHROMYCIN STAR, AZROMAX, AZICURE, ZITHROX, and AZBACT. All brands were evaluated using standard quality control tests, including physical examination (appearance, color, break-line, and edge integrity), weight variation, content uniformity (assay), thickness, hardness, friability, disintegration time, and dissolution. All analyses were performed according to official pharmacopeial guidelines [16, 17] to determine pharmaceutical quality and in vitro performance. The following instruments were used during testing: an electronic analytical balance, a digital friability tester (Tnnco), a Monsanto hardness tester, a disintegration test apparatus, a dissolution test apparatus, and a UV–visible spectrophotometer (Systronics Smart).

### 2.3 Analytical Methods

In this study, several standard tests were performed to evaluate the quality of all selected azithromycin tablet brands.

#### *2.3.1 Physical Examination*

The physical characteristics of the tablets were assessed through visual inspection. For each brand, ten tablets were randomly selected and examined with the naked eye to evaluate their appearance, color, break-line, edge integrity, and the presence of any cracks or deformities. This procedure was applied to all brands to maintain consistency. Physical examination provides an initial indication of product quality and helps detect manufacturing defects that may affect patient acceptability, dosing accuracy, or product stability [18, 19].

#### *2.3.2 Weight Variation Test*

The weight variation test verifies that tablets contain a uniform amount of active ingredient and is a key parameter in quality control. According to the British Pharmacopoeia (BP, 2023) [16], tablets with an average weight of more than 80 mg should not deviate by more than ±5%, while tablets weighing 80 mg or less should not deviate by more than ±10%. In this study, twenty tablets from each brand were individually weighed using an analytical balance. The average tablet weight was calculated, and the deviation of each tablet from this average was recorded. The percentage weight variation was calculated using the following equation:

$$Weight\ variation\ (\%) = \frac{(Iw - Aw)}{Aw} \times 100$$

where:

$Iw$ = Individual weight of tablet
$Aw$ = Average weight of tablet

This test helps identify inconsistencies in the manufacturing process, such as variations in die filling, compression force, or powder flow, which can affect content uniformity and, consequently, the therapeutic efficacy of the tablets [20].

#### *2.3.3 Assay of the active ingredient*

For the quantitative determination of azithromycin, twenty tablets from each brand were randomly selected, accurately weighed, and finely powdered. An amount of the powdered sample equivalent to 667 mg of azithromycin was transferred into a 200 mL volumetric flask. A suitable volume of diluent was added, and the mixture was sonicated for 15 minutes to ensure complete extraction of the active pharmaceutical ingredient. The solution was then allowed to cool to room temperature and diluted to volume with Diluent A to obtain a final concentration of 0.4 mg/ml. The resulting solution was filtered through a 0.45 µm membrane filter to remove insoluble excipients. The same procedure was applied to all tested brands. A standard solution of azithromycin was prepared under identical conditions for quantitative comparison. Chromatographic analysis was performed using a high-performance liquid chromatography (HPLC) system equipped with a UV detector. Separation was achieved on a reversed-phase C18 column (250 mm × 4.6 mm, 5 µm particle size). The mobile phase consisted of a mixture of acetonitrile and phosphate buffer (pH 6.0) at a 60:40 (v/v) ratio, delivered at a flow rate of 1.0 mL/min. The detection wavelength was set at 210 nm, with an injection volume of 20 µL and a total run time of 10 minutes. System suitability testing was performed prior to analysis to ensure the reliability of the chromatographic system. The parameters evaluated included theoretical plate count (N ≥ 2000), tailing factor (≤ 2), and relative standard deviation (RSD ≤ 2%) for replicate injections. Diluent A consisted of a mixture of phosphate buffer (pH 6.0) and acetonitrile in a ratio of 50:50 (v/v). Quantification of azithromycin was carried out by comparing the peak area of the sample solutions with that of the corresponding standard solution. The method was performed in accordance with the pharmacopeial monograph for azithromycin tablets (e.g., United States Pharmacopeia). This method ensures accurate quantification of the active ingredient and allows reliable comparison with pharmacopeial specifications [21].

#### *2.3.4 Hardness Test*

The mechanical strength (hardness) of the tablets was evaluated using a Monsanto-type hardness tester (China). Ten tablets from each brand were randomly selected and individually tested to determine the force required to break each tablet. The mean hardness for each brand was calculated using the following formula:

$$Hardness\ (kg/cm^2) = \frac{Total\ hardness\ of\ all\ tablet}{Number\ of\ tablets}$$

This test determines the tablet's resistance to chipping, abrasion, and breakage during handling, packaging, and transportation [22].

#### *2.3.5 Friability Test*

Friability testing was carried out using an Electro Lab EF-Friabilator following United States Pharmacopeia (USP) guidelines. Ten tablets from each brand were accurately weighed and placed in the apparatus, which was operated at 25 rpm for 4 minutes (equivalent to 100 revolutions). After the test, the tablets were reweighed, and the percentage friability (%F) was calculated using the following equation:

$$\%\ of\ Friability = \frac{weight\ before\ test - weight\ after\ test}{weight\ before\ test} \times 100$$

Friability assesses the tablet's ability to withstand mechanical stress during handling and transportation. For conventional tablets, a friability value not exceeding 1% is considered acceptable [23].

#### *2.3.6 Disintegration Time Test*

The in vitro disintegration time was evaluated using a USP disintegration tester with a disc and distilled water as the medium. For each brand, three tablets were placed individually into the tubes of the basket rack, which was then lowered into a 1-liter beaker containing water maintained at 37 ± 0.5 °C. The disintegration time was recorded as the time required for each tablet to completely break down into fine particles and pass through the mesh at the bottom of the tube. The average disintegration time of the three tablets was taken as the disintegration time for that brand [17].

#### *2.3.7 Dissolution Study*

In vitro dissolution profile of azithromycin tablets was determined using USP Dissolution Apparatus II (paddle method). Each of the six vessels was filled with 1000 mL of phosphate buffer (pH 6.0) maintained at 37 ± 0.5°C. After the medium reached equilibrium, one tablet from each brand was placed into the vessels, and the paddle speed was set to 75 rpm. Samples (13 mL) were collected at 10, 20, and 30 minutes using a syringe and replaced with an equal volume of fresh buffer to maintain sink conditions. The samples were filtered through a 0.45 µm membrane filter, and 11.25 mL of each filtrate was transferred to a 25 mL volumetric flask and diluted with diluent (USP 36/NF 31) [24] to obtain a concentration of approximately 0.25 mg/mL, similar to the standard. A 50 µL volume of both the sample and standard solutions was injected into a UV–visible spectrophotometer, and absorbance was measured at 298 nm. The percentage of drug release was calculated using the following formula:

$$\%\ Content\ of\ drug\ release = [Ru/Rs] \times [Cs/L] \times V \times 100$$

where:

Ru = response of the sample, Rs = response of the standard, Cs = concentration of the standard, L= label claim, and V = volume of dissolution medium, and V= volume of dissolution medium. The obtained results were compared with the United States Pharmacopeia (USP) acceptance criteria for azithromycin tablets [17].

## 3. Results and Discussion

The quality of pharmaceutical products is a key factor determining their safety, efficacy, and therapeutic reliability. In this study, five commercially available azithromycin tablet brands marketed in Dhamar City, Yemen, were evaluated for their compliance with international pharmacopeial standards, including the United States Pharmacopeia (USP), British Pharmacopoeia (BP), and Indian Pharmacopoeia (IP). The parameters assessed included physical characteristics, weight variation, assay (active ingredient content), hardness, friability, disintegration time, and dissolution profile.

### 3.1 Physical Examination

As shown in Table 1, all tested brands exhibited acceptable physical characteristics, with uniform oblong shapes, smooth surfaces, and intact edges. These features suggest adherence to good manufacturing practices during the compression and coating stages. However, Brand A1 showed surface fragility during the friability test, indicating a possible deficiency in binder concentration or insufficient compression force during manufacturing. A noticeable color difference was also observed in Brand A5, which appeared yellowish compared with the white tablets of the other brands. This variation may be related to differences in excipients or the intentional use of colorants for product identification or stability purposes. Although color differences do not directly influence drug efficacy, they may affect patient confidence and adherence, especially when inconsistencies appear between different batches [18]. Overall, the physical examination results indicate that most brands meet pharmacopeial appearance requirements and maintain acceptable structural integrity, reflecting generally satisfactory formulation quality.

**Table 1.** Physical Characteristics of Different Azithromycin Tablet Brands.

| Brands | Description |
|---|---|
| **A1** | White colour, and oblong shape, tablets with regular edges |
| **A2** | White colour, and oblong shape, tablets with regular edges |
| **A3** | White colour, and oblong shape, tablets with regular edges |
| **A4** | White colour, and oblong shape, tablets with regular edges |
| **A5** | Yellow colour, and oblong shape, tablets with regular edges |

### 3.2 Weight Variation

As shown in Table 2, all five azithromycin tablet brands met pharmacopeial specifications, with weight variation falling within the acceptable ±5% limit specified in the USP and BP standards [16, 17]. The mean tablet weights ranged from 0.680 g for Brand A4 to 0.881 g for Brand A3, reflecting slight differences in excipient composition and formulation strategies among manufacturers. The low standard deviation values (0.013–0.021 g) indicate excellent uniformity in die filling and compression during tablet production. Such consistency suggests a well-controlled manufacturing process with stable powder flow properties across all brands. Maintaining uniform tablet weight is essential to ensure accurate dosing of the active pharmaceutical ingredient (API), which supports therapeutic reliability and reduces the risk of underdosing or overdosing [20]. The results confirm that all tested brands demonstrated acceptable weight uniformity, indicating compliance with international pharmacopeial requirements and reflecting effective control of production parameters.

**Table 2.** Weight Variation of Different Azithromycin Tablet Brands.

| Brands | Weight (Mean ± SD) |
|---|---|
| **A1** | 0.685 ± 0.0134 |
| **A2** | 0.744 ± 0.0134 |
| **A3** | 0.881 ± 0.0144 |
| **A4** | 0.68 ± 0.0179 |
| **A5** | 0.871 ± 0.0212 |

### 3.3 Assay of the Active Ingredient

As shown in Table 3 and Figure 3, the assay results revealed clear variability among the five azithromycin tablet brands. Four brands—A1, A2, A3, and A4—were within the USP specification range of 90–110% of the labelled claim [17], indicating proper formulation, good uniformity of drug distribution, and adequate chemical stability during manufacturing and storage. However, Brand A5 contained only 77.1% of the labelled amount, which is far below pharmacopeial requirements and therefore failed the assay test. This low drug content is a serious quality and public health concern. Subtherapeutic antibiotic doses may lead to ineffective treatment, longer illness duration, and an increased risk of antimicrobial resistance (AMR)—a growing global challenge, especially in low-resource settings such as Yemen [25-27]. Possible causes of this deviation include poor mixing during formulation, inaccurate weighing of the active ingredient, or degradation of azithromycin due to improper storage conditions, such as high humidity or elevated temperatures [19, 28]. Similar issues have been reported in previous studies from Nigeria and Ghana, where some azithromycin brands were found to be subpotent and non-compliant with pharmacopeial assay limits [28, 29]. Overall, these findings highlight the need for strong regulatory oversight, strict quality assurance practices, and routine post-market surveillance to ensure that all antibiotic products meet international quality standards before they reach consumers.

**Table 3.** Assay Results of the Active Ingredient in Different Azithromycin Brands

| Brands | Assay (%) |
|---|---|
| **A1** | 96.60 |
| **A2** | 90.11 |
| **A3** | 97.50 |
| **A4** | 102.30 |
| **A5** | 77.10 |

### 3.4 Hardness and Friability

As shown in Table 4, the hardness results indicate that all azithromycin tablet brands had sufficient mechanical strength, with values above 9.6 kg—well above the USP minimum requirement of about 4 kg [17]. This confirms that the tablets can withstand normal mechanical stress during handling, packaging, and transportation. Among the tested products, A2 and A4 recorded the highest hardness values (14.6 kg and 14.99 kg, respectively), which may be linked to the use of higher compression forces or greater amounts of binding agents.

**Table 4.** Hardness of Different Azithromycin Tablet Brands

| Brands | Hardness (kg) |
|---|---|
| **A1** | 13.9 |
| **A2** | 14.6 |
| **A3** | 13.44 |
| **A4** | 14.99 |
| **A5** | 9.64 |

Despite its relatively high hardness (13.9 kg), Brand A1 failed the friability test because it completely disintegrated during the 4-minute tumbling process. This unusual result suggests poor internal cohesion or uneven binder distribution, leading to structural weakness even though the tablet surface appears firm. According to pharmacopeial standards, friability should not exceed 1% for conventional tablets [16, 30]. As presented in Table 5 and Figure 5, the remaining brands (A2–A5) showed friability values between 0.114% and 0.522%, which are within acceptable limits. These results indicate that their formulations are robust and reflect good manufacturing practices with effective control of granulation and compression steps. These observations are consistent with previous studies conducted in India and Bangladesh, where most azithromycin tablet brands complied with both hardness and friability requirements [31, 32]. Overall, the present results suggest that the majority of the tested products have suitable mechanical integrity for normal handling and distribution. However, Brand A1 requires further formulation improvements to meet friability standards.

**Table 5.** Friability of Different Azithromycin Tablet Brands

| Brand | Initial weight (g) | Final weight (g) | Result (%) |
|---|---|---|---|
| **A1** | Failed | Failed | Failed |
| **A2** | 7.65 | 7.61 | 0.522 |
| **A3** | 8.81 | 8.80 | 0.114 |
| **A4** | 6.81 | 6.80 | 0.15 |
| **A5** | 8.65 | 8.62 | 0.35 |

### 3.5 Disintegration Time

As shown in Figure 1, all five azithromycin tablet brands complied with pharmacopeial limits, with disintegration times ranging from 3 minutes (A5) to 15 minutes (A2), well within the 30-minute maximum specified for film-coated tablets by the USP and BP [16, 17]. The rapid disintegration observed for Brands A3 (5 min) and A5 (3 min) suggests the presence of effective disintegrants, such as croscarmellose sodium or sodium starch glycolate, which enhance water uptake and promote tablet swelling, leading to faster breakdown. Rapid disintegration is generally associated with improved dissolution rates and higher bioavailability, which can result in a faster onset of therapeutic action [33, 34]. Conversely, Brand A2 had the longest disintegration time (15 min). While still pharmacopeial compliant, this may slightly delay drug release and onset of action. Differences in disintegration behavior among brands likely reflect variations in formulation design, including disintegrant type and concentration, binder ratios, coating thickness, and compression force applied during tableting [18, 33]. Overall, the results suggest that all brands meet official standards, with A3 and A5 potentially offering superior in vitro disintegration performance.

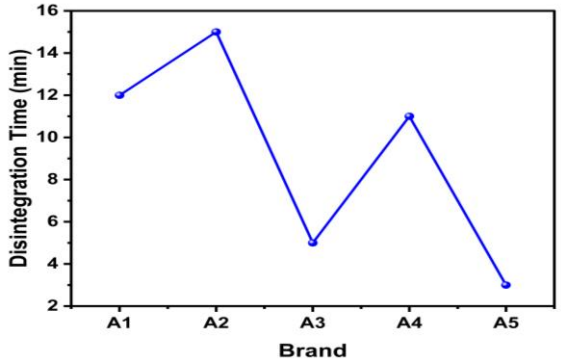



**Figure 1:** Disintegration Times of Different Azithromycin Tablet Brands.

### 3.6 Dissolution Profile

As presented in Figure 2, the dissolution profiles of all five azithromycin tablet brands complied with USP and BP specifications, which require that at least 80% of the labelled drug content be released within 30 minutes for immediate-release tablets [16, 17]. All brands released more than 80% of the drug within 30 minutes under standard dissolution conditions. Brand A5 showed the highest dissolution efficiency (97.5%), followed by Brand A3 (93.75%), indicating optimized formulation parameters that enhance wetting, disintegration, and drug solubilization [34]. In contrast, Brand A2 had the lowest release (82.0%), which, while acceptable, may reflect higher binder content, denser granule structure, or reduced tablet porosity, slowing fluid penetration and drug release. The dissolution results correlate with disintegration behavior: tablets that disintegrated faster (A3 and A5) also dissolved more quickly, highlighting the strong relationship between disintegration and drug availability. These findings are consistent with previous studies in Nigeria and Uganda, which reported that most marketed azithromycin tablets met pharmacopeial dissolution standards [15, 35, 36]. However, Brand A5, despite its excellent dissolution, failed the assay test (77.1% of labelled content). This indicates that even though the tablet dissolves efficiently, the total drug content is insufficient for therapeutic effectiveness. This underscores a critical quality concern: high dissolution cannot compensate for subpotent tablets, as this may still lead to subtherapeutic dosing and contribute to antimicrobial resistance.

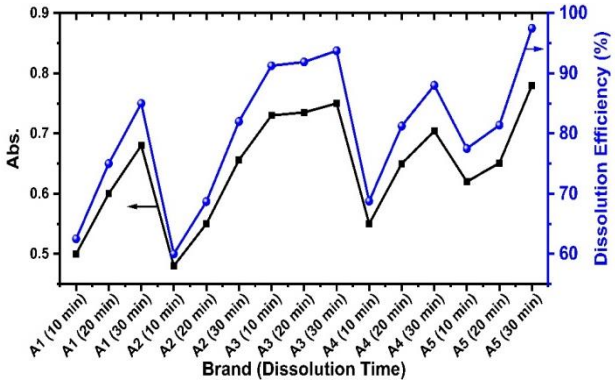



**Figure 2:** Dissolution Profile Data of Different Azithromycin Tablet Brands.

## 4. Conclusion

This study assessed the in vitro quality of five commercially available azithromycin 500 mg tablet brands in Dhamar City, Yemen, against USP, BP, and IP standards. All brands met pharmacopeial requirements for weight variation, hardness, disintegration, and dissolution, with more than 80% drug release within 30 minutes. However, two critical issues were identified. Brand A5 contained only 77.1% of the labelled azithromycin, well below the USP acceptable range of 90–110%, posing a significant risk of therapeutic failure and contributing to antimicrobial resistance. Brand A1 failed the friability test, completely disintegrating during mechanical stress, which indicates poor mechanical strength and unsuitability for handling or patient use. These results demonstrate that not all marketed brands meet essential quality standards, despite claims of compliance. The presence of substandard and physically unstable products highlights a critical gap in pharmaceutical regulatory oversight in Yemen. Routine post-market surveillance, mandatory laboratory testing of random samples, and public disclosure of non-compliant products are strongly recommended to protect public health and ensure the efficacy of essential medicines.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Conflict of Interest

The authors declare no conflicts of interest.

## References

[1] World Health Organization. (2024) *Quality assurance of pharmaceuticals: a compendium of guidelines and related materials*. Volume 1: Good manufacturing practices and inspection, 9240099425; World Health Organization, Geneva, Switzerland, ISBN: 978-92-4-009942-5, pp 534.

[2] Sanjeev Salvi, S., Kerkar, S. (2020) Quality Assurance and Quality Control for Project Effectiveness in Construction and Management, *International Journal of Engineering Research & Technology* **9**: 133–137.

[3] Arachchi, S.M., Chong, S.C., Madhushani, A. (2015) Quality assurance and quality control in ERP systems implementation, *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)* **11**: 70-83.

[4] Leong, T.K., Zakuan, N., Saman, M.Z.M. (2014) Review of quality management system research in construction industry, *International Journal of Productivity and Quality Management* **13**: 105-123.

[5] Teasdale, A., Elder, D., Nims, R.W. (2017) ICH Quality Guidelines: An Implementation Guide, John Wiley & Sons, Inc., *Hoboken, New Jersey, USA*, pp. 808.

[6] Montgomery, D.C. (2020) Introduction to Statistical Quality Control, 8th ed., *John Wiley & Sons, Inc.*, Hoboken, New Jersey, USA, pp. 768.

[7] Taylor, R., Shakoor, O., Behrens, R., Everard, M., Low, A., Wangboonskul, J., Reid, R., Kolawole, J. (2001) Pharmacopoeial quality of drugs supplied by Nigerian pharmacies, *The Lancet* **357**: 1933-1936.

[8] Neyestani, B. (2017) Seven Basic Tools of Quality Control: The Appropriate Techniques for Solving Quality Problems in the Organizations, **SSRN Article No. 2955721.** https://doi.org/10.5281/zenodo.400832.

[9] Hoepelman, I., Schneider, M. (1995) Azithromycin: the first of the tissue-selective azalides, *International Journal of Antimicrobial Agents* **5**: 145-167.

[10] Idkaidek, N.M., Najib, N., Salem, I., Jilani, J. (2014) Physiologically-based IVIVC of azithromycin, *American Journal of Pharmacological Sciences* **2**: 100-102.

[11] Retsema, J., Girard, A., Schelkly, W., Manousos, M., Anderson, M., Bright, G., Borovoy, R., Brennan, L., Mason, R. (1987) Spectrum and mode of action of azithromycin (CP-62,993), a new 15-membered-ring macrolide with improved potency against gram-negative organisms, *Antimicrobial Agents and chemotherapy* **31**: 1939-1947.

[12] Piscitelli, S., Danziger, L., Rodvold, K. (1992) Clarithromycin and azithromycin: new macrolide antibiotics, *Clinical Pharmacy* **11**: 137-152.

[13] Al-Rimawi, F., Kharoaf, M. (2010) Analysis of azithromycin and its related compounds by RP-HPLC with UV detection, *Journal of Chromatographic Science* **48**: 86-90.

[14] Jaber, S.H., Salih, Z.T., Salmo, H.M. (2012) Formulation of azithromycin suspension as an oral dosage form, *Iraqi Journal of Pharmaceutical Sciences* **21**: 61-69.

[15] Kalamuzi, P.M. (2018) A pharmaceutical equivalence study of the selected Azitiiromycin 500 mg brands on the Ugandan Market. *Bachelor of Pharmacy (BPharm) Research Project Report*, BPharm Thesis, Kampala International University, Ishaka, Bushenyi, Uganda.

[16] British Pharmacopoeia Commission. (2022) *British Pharmacopoeia*. The Stationery Office, London, UK, ISBN: 9780113230921, pp

[17] United States Pharmacopeial Convention. (2020) *The United States pharmacopeia: USP 43–National formulary: NF 38*. United States Pharmacopeial Convention, Inc., Rockville, Maryland, USA, ISBN-13 (5-Volume Print Set): 978-1936424948, pp 8760.

[18] Aulton, M.E., Taylor, K. (2013) Aulton's pharmaceutics: the design and manufacture of medicines, 4th ed., *Churchill Livingstone / Elsevier*, London, UK, pp. 908.

[19] World Health Organization. (2017) *A study on the public health and socioeconomic impact of substandard and falsified medical products*. World Health Organization, Geneva, Switzerland, ISBN: 978-92-4-151343-2, pp 72.

[20] Mekasha, Y.T., Chali, B.U., Feissa, A.B., Godena, G.H., Hassen, H.K., Wega, S.S. (2023) Quality evaluation of the Azithromycin tablets commonly marketed in Adama, and Modjo towns, Oromia Regional State, Ethiopia, *Plos One* **18**: e0282156.

[21] Huber, L. (2007) Validation and Qualification in Analytical Laboratories, 2nd ed., *CRC Press*, Boca Raton, Florida, USA, pp. 288.

[22] Amin, M.R., Biswas, S., Rahman, M.R., Bhuiyan, J.R., Rana, M.S. (2014) Study and impact evaluation of particle size distribution on physicochemical properties of three different tablet formulations through sieve technology, *International Journal for Pharmaceutical Research Scholars* **3**: 448–463.

[23] Almuzaini, T., Choonara, I., Sammons, H. (2013) Substandard and counterfeit medicines: a systematic review of the literature, *BMJ Open* **3**: e002923.

[24] United States Pharmacopeial Convention. (2013) *Vehicle for Oral Solution, Vehicle for Oral Solution Sugar-Free and Vehicle for Oral Suspension*. In United States Pharmacopeia 36–National Formulary 31 (USP 36–NF 31), United States Pharmacopeial Convention, Rockville, MD, USA, pp n.pag.

[25] World Health Organization. (2015) *Guidance on The Selection of Comparator Pharmaceutical Products for Equivalence Assessment of Interchangeable Multisource (Generic) Products*. WHO Expert Committee on Specifications for Pharmaceutical Preparations: 49th report, TRS No. 992; World Health Organization, Geneva, Switzerland, ISBN: 978-92-4-120992-2 pp 185–195.

[26] Newton, P.N., Green, M.D., Fernández, F.M. (2010) Impact of poor-quality medicines in the 'developing'world, *Trends in Pharmacological Sciences* **31**: 99-101.

[27] Okeke, I.N., Lamikanra, A., Edelman, R. (1999) Socioeconomic and behavioral factors leading to acquired bacterial resistance to antibiotics in developing countries, *Emerging Infectious Diseases* **5**: 18.

[28] Okeke, I., Lamikanra, A. (2001) Quality and bioavailability of ampicillin capsules dispensed in a Nigerian semi-urban community, *African Journal of Medicine and Medical Sciences* **30**: 47–51.

[29] Egbo, H.A. (2013) Quality Evaluation of a Selection of Antibiotics Distributed in Accra (Ghana) and Lagos (Nigeria). *Department of Chemistry*, Master's Thesis, University of Ghana, Legon, Ghana.

[30] Indian Pharmacopoeia Commission. (2020) *Indian Pharmacopoeia 2020*. Indian Pharmacopoeia Commission, Ministry of Health & Family Welfare, Government of India, Ghaziabad, Uttar Pradesh, India, ISBN: 978-93-89098-61-0, pp

[31] Singh, R., Saxena, M., Sahay, D., Singh, S. (2017) In-vitro study of quality control parameters of three different brands of azithromycin tablets, *International Journal of Basic & Clinical Pharmacology* **6**: 1572–1575.

[32] Bari, F.S.N.U., Islam, M.R., Moghal, M.M.R., Ira, I.J. (2017) Release Kinetics Study of Azithromycin from Bi-layered Tablets, *Bangladesh Pharmaceutical Journal* **20**: 54-63.

[33] Tarke, S., Shanmugasundaram, P. (2017) Formulation and evaluation of fast Dissolving tablets of Antihypertensive Drug, *Research Journal of Pharmacy and Technology* **10**: 155.

[34] Leuner, C., Dressman, J. (2000) Improving drug solubility for oral delivery using solid dispersions, *European Journal of Pharmaceutics and Biopharmaceutics* **50**: 47-60.

[35] Dressman, J.J., Kramer, J. (2005) Pharmaceutical Dissolution Testing, 1st ed., *CRC Press*, Boca Raton, Florida, USA, pp. 429.

[36] Ukwueze, S., Nwachukwu, I., Ezealisiji, K. (2018) Ultraviolet Spectrophotometric Evaluation of Different Brands of Azithromycin Dihydrate Tablets and Capsules Sold in Nigeria, *European Journal of Pharmaceutical and Medical Research* **5**: 112–118.

TuJNaS

*Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 34– 39*

**DOI:** https://doi.org/10.59167/tj3pbg80

Original Article

# Preliminary Evaluation of Acute Toxicity and Hypoglycemic Effects of *Salvia officinalis* L. Extract in Healthy Male Rabbits: An Animal Model Study

**Hisham A. Al-khawlani[1,*], Nabil Ali Al-Mekhlafi[2], Aisha Al-Habri[3], Amal Al-Nahary[3], Hadeel Al-mekhlafi[3], Hanan Al-Nahary[3], Malak Al-magmhi[3], Maram Al-Mashrama[3], Resalh Jubra[3], Safa'a Amer[3], Shatha Al-Fthahy[3], Wala'a Amran[3],**

[1]*Department of Pharmacy, Albaydha University Institute for Continuing Education, Albaydha University, Albaydha, Yemen.*
[2]*Department of Biochemical Technology, Faculty of Applied Science, Thamar University, Dhamar 87246, Yemen.*
[3]*Department of Pharmacy, Thamar University Institute for Continuing Education, Thamar University, Dhamar 87246, Yemen.*

***Corresponding Author:** Hisham A. Al-khawlani, Department of Pharmacy, Albaydha University Institute for Continuing Education, Albaydha University, Albaydha 87246, Yemen. Cell Phone: +967 774580010, E-mail: hisham.alkhawlanii@gmail.com

Received: 9 December 2025. Revised: 2 April 2026. Accepted: 3 April 2026. Published: 29 June 2026.

## Abstract

**Background:** *Salvia officinalis L.* is an aromatic perennial herb belonging to the Lamiaceae family, known for its diverse pharmacological properties, including potent antioxidant, antimicrobial, anti-inflammatory, neuroprotective, and anti-diabetic activities. Despite its extensive traditional use, the specific roles of *S. officinalis* in glucose management and its acute toxicity profile in rabbits require further elucidation. **Objective:** This study provided a preliminary assessment of the acute toxicity profile and the basal glycemic effects of a methanolic extract of *S. officinalis* in a normoglycemic rabbit model. **Methodology:** Twelve adult male White rabbits were randomly assigned to four groups (n=3). Group I served as the vehicle control, while Groups II, III, and IV received single oral doses of *S. officinalis* extract at 1000, 1500, and 2000 mg/kg, respectively. Animals were monitored over 14 days for toxicity, mortality, and body weight, alongside repeated measurements of fasting blood glucose (FBG). Data were analyzed using two-way repeated measures ANOVA. **Results:** *S. officinalis* extract demonstrated dose-dependent basal glucose modulation. The repeated measures analysis revealed highly significant effects for both time and dose ($P < 0.001$). The 1000 mg/kg dose was identified as the No Observed Adverse Effect Level (NOAEL), while the 1500 mg/kg dose was established as the Maximum Tolerated Dose (MTD). At 2000 mg/kg, an observed lethal threshold was recorded with a 66.7% mortality rate. **Conclusion:** The preliminary findings suggest that *S. officinalis* extract exhibits potential dose-dependent hypoglycemic activity in normoglycemic models. However, the manifestation of systemic toxicity at higher concentrations defines a narrow therapeutic window for the extract.

**Keywords:** *Salvia officinalis*; Hypoglycemic; Acute Toxicity; Rabbit Model; Safety Assessment

## 1. Introduction

Diabetes Mellitus (DM) is a group of metabolic diseases characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both [1, 2]. Globally, DM prevalence poses a significant public health challenge, with estimates of 537 million adults affected, expected to rise to 643 million by 2030 and 783 million by 2045. The increasing resistance to standard therapies and treatment costs has led to a renewed interest in complementary and alternative medicine, particularly traditional herbal remedies for DM management [3–5].

In this context, *Salvia officinalis* (common sage or garden sage), an aromatic perennial herb belonging to the Lamiaceae family, has garnered significant scientific attention [6]. Historically recognized for its diverse pharmacological spectrum, this herb exhibits potent antioxidant, antimicrobial, anti-inflammatory, neuroprotective, and notably, anti-diabetic activities [7]. The therapeutic potential of *S. officinalis* is primarily attributed to its complex phytochemical composition, which includes essential oils, flavonoids, and phenolic acids. These constituents are thought to exert their hypoglycemic effects through various mechanisms, such as improving insulin sensitivity, enhancing glucose uptake, and inhibiting α-glucosidase activity [8–15]

Despite the extensive traditional application and a growing body of in vitro and in vivo literature elucidating its general pharmacological effects, several critical aspects regarding the safe and effective use of *S. officinalis* in DM management remain inadequately addressed. Specifically, there is a notable lack of robust experimental evidence defining the acute safety profile (NOAEL, MTD) and the precise dose-response relationship of its methanolic extract. Therefore, this study aims to provide a preliminary characterization of the acute safety profile and explore the basal glycemic effects of *Salvia officinalis* methanolic extract in a normoglycemic rabbit model, establishing a pharmacological baseline for future larger-scale investigations.

## 2. Methodology

### 2. 1 Plant Material and Extract Preparation

Fresh leaves of *Salvia officinalis L.* were collected in February 2025 from a house garden in Dhamar city, Yemen. The leaves were formally identified and authenticated by Dr. Abdullah Al-Shawsh, Department of Botany, Thamar University, Yemen, Under voucher specimen (No. TU+AC-2024). Certificates of authentication were subsequently submitted and retained in the Department of Pharmacy, Thamar University, Yemen.

The collected *S. officinalis* leaves (Figure 1) were thoroughly washed with distilled water and shade-dried at ambient temperature [11]. The dried leaves were subsequently pulverized into a fine powder using a mortar and pestle, according to established methods [11, 16]. A methanolic extract was then prepared via maceration. To ensure a sufficient yield for high-dose administration, the extraction was conducted in successive cycles (10 g of powder per 100 mL solvent) for 48 hours with occasional agitation. The filtrate was concentrated using a rotary evaporator at 40 °C under reduced pressure. The crude methanolic extract was stored in airtight amber vials at 20 °C until further use, following standard protocols [4, 17, 18]. Finally, the total yield of the extract was determined [19].

### 2. 2 Experimental Animals

A total of twelve adult male White rabbits (1000–1500 g, 18 weeks old) were sourced from the local market. The animals were housed in metallic cages at the Thamar University Institute for Continuous Education. All procedures involving the animals adhered to the internationally recognized guidelines for the care and use of laboratory animals and were approved by the Animal Ethics Committee (AEC) of Thamar University [20, 21].

**Figure 1:** Leaves of the *Salvia officinalis*.

Rabbits were acclimated for one week under controlled conditions (23 ± 2 °C, 55 ± 5% humidity, 12:12 h light/dark cycle) [22] [23]. Following acclimatization, animals were randomly assigned to four groups (n = 3 per group) for the **Preliminary Acute Toxicity and Dose-Ranging Assessment** (Table 1). The use of four parallel groups with n=3 was adopted as a **preliminary screening method** to establish the **Maximum Tolerated Dose (MTD)**. Group I (Control): Received vehicle only (distilled water), Group II-IV: Received single oral doses of sage extract at 1000, 1500, 2000 mg/kg, respectively. The extract and vehicle were administered as a single oral dose (P.O.) via oral gavage

**Table 1**: Design and distribution of experimental animals grouping.

| Group | Category | Dose Administration |
|---|---|---|
| **Group I** | Vehicle control | Distilled water (vehicle) |
| **Group II** | Low-Dose Toxicity/Efficacy | Extract 1000 mg/kg BW (p.o.) |
| **Group III** | Mid-Dose Toxicity/Efficacy | Extract 1500 mg/kg BW (p.o.) |
| **Group IV** | High-Dose Toxicity/Efficacy | Extract 2000 mg/kg BW (p.o.) |

***Note:*** The *S. officinalis* extract was administered as a single oral dose (P.O.) via oral gavage. Animals were fasted overnight before dosing to stabilize baseline glucose.

### 2. 3 Preliminary Acute Toxicity Assessment and Body Weight

The acute oral toxicity was assessed using a dose-range finding study (limit test) to characterize the initial safety profile of the extract. In accordance with the 3Rs principles for animal reduction and the exploratory nature of this screening, the No Observed Adverse Effect Level (NOAEL) and Maximum Tolerated Dose (MTD) were identified based on the highest doses resulting in zero mortality. Clinical signs were monitored using the Functional Observational Battery at least twice daily. Observations for behaviour, water/food intake, and mortality were recorded over a period of 14 days. This preliminary screening focused on clinical and observational parameters, including mortality and physical indicators. Body weights were measured using an electrical balance at baseline Day 0 and at the end of the experiment (Day 14). Body weight data were statistically analyzed and expressed as mean ±SEM. Figure 2 shows a graphical representation of the study methodology.

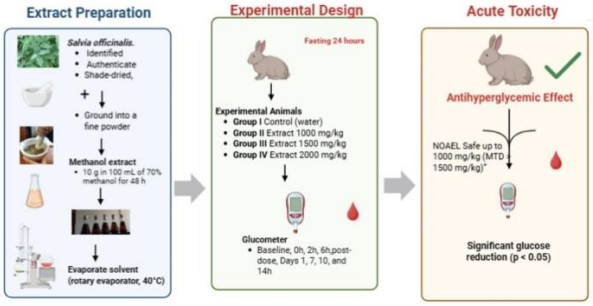



**Figure 2:** Graphical representation of *Salvia officinalis* ethanolic extract is safe and demonstrates significant antihyperglycemic activity in rabbits.

### 2. 4 Assessment of Effect on Basal Blood Glucose Levels

To evaluate the hypoglycemic efficacy of the *S. officinalis* extract, fasting blood glucose (FBG) levels were monitored in normoglycemic rabbits. Baseline measurements (0 h) were obtained immediately prior to extract administration. Post-treatment glucose levels were then recorded at acute intervals (0, 2, and 6 h) and sustained intervals (Days 1, 7, 10, and 14). Blood samples (~3mL) were collected from the marginal ear vein via lancet puncture, and glucose concentrations were determined using a validated portable glucometer (Accu-Chek, Roche Diagnostics, Germany) in accordance with the manufacturer's instructions.

### 2. 5 Ethical Approval

All experimental procedures were conducted in accordance with the ethical guidelines for the care and use of laboratory animals and were approved by the Animal Ethics Committee (AEC) of Thamar University (**Ref No: TU-2025-045**). Animals were monitored every 2 hours during the first 24 hours for clinical signs of toxicity. Humane endpoints were established to prevent unnecessary suffering; animals showing severe distress or moribund conditions were slated for immediate euthanasia using an overdose of anesthesia. Rabbits that succumbed to the high-dose toxicity (2000 mg/kg) were recorded, while surviving healthy animals were maintained under veterinary supervision for the 14-day observation period before being humanely retired from the study.

### 2. 6 Statistical Analysis.

Data were expressed as mean ± SEM and analyzed using SPSS software (version 16). For body weight data, differences among groups at Day 14 were analyzed by One-way ANOVA. For fasting blood glucose data, a **two-way repeated measures ANOVA** was employed to evaluate the main effects of Dose and Time, as well as their interaction. **Bonferroni post-hoc tests** were performed to identify significant differences between groups at specific time points. Mortality rates were compared by Fisher's exact test. A p-value of less than 0.05 ($p < 0.05$) was considered significant.

## 3. Results

### 3. 1 Plant Material and Extract Yield

The extraction process, involving a 48-hour maceration of 10 g of dried *S. officinalis* leaves in 70% methanol, resulted in a dark brown, gummy crude extract. The calculated percentage yield was 10.3% (w/w). This yield provided a sufficient quantity of the phytochemical constituents required for the subsequent acute toxicity and hypoglycemic assessments in the rabbit model.

### 3. 2 Preliminary Acute Toxicity and Clinical Observations

Oral administration of the methanolic extract of *S. officinalis* resulted in dose-dependent mortality and pronounced clinical manifestations (Table 2). All rabbits (3/3) in Group II survived the 14-day observation

period, establishing this dosage as the NOAEL. Conversely, Group III recorded a 33.3% mortality rate (1/3), while Group IV exhibited the highest fatalities at 66.7% (2/3). These findings identify 1500 mg/kg as the MTD and approximately 2000 mg/kg as the observed lethal threshold.

Clinical signs also followed a dose-dependent pattern. While the control and Group II exhibited normal behavior and stable vital signs, rabbits receiving 1500 and 2000 mg/kg displayed pronounced neurobehavioral alterations, including drowsiness, trembling, and aggressiveness. Severe manifestations, such as paralysis and asphyxiation, were exclusively recorded in Group IV, reflecting acute systemic toxicity.

**Table 2.** Acute Toxicity and Clinical Observations of *S. officinalis* Methanolic Extract in Male Rabbits Over 14 Days

| Mortality and clinical signs | Group I (Control) | *S. officinalis* methanolic extract | | |
|---|---|---|---|---|
| | | Group II (1000 mg) | Group III (1500 mg) | Group IV (High 2000 mg) |
| **Mortality** | 0/3 | 0/3 | 1/3 | 2/3 |
| **Behavioral Changes** | None | None | drowsiness, trembling, | Aggressiveness, paralysis, drowsiness, trembling, asphyxiation |
| **Body Temperature** | Normal | Normal | Normal | Elevated |
| **Heart Rate** | Normal | Increased | Increased | Increased |
| **Respiratory Rate** | Normal | Slightly Increased | Increased | Increased |

***Note:** N: normal, AN: abnormal, P: presence, A: aggressive.*

### 3. 3 Food and Water Consumption

As shown in Figures 3A and 3B, no statistically significant differences in mean food and water intake were observed across all treated groups compared to the control (P > 0.05). These results suggest that the extract did not impact appetite or hydration, indicating that observed weight changes were likely driven by metabolic or toxicological factors rather than reduced intake.

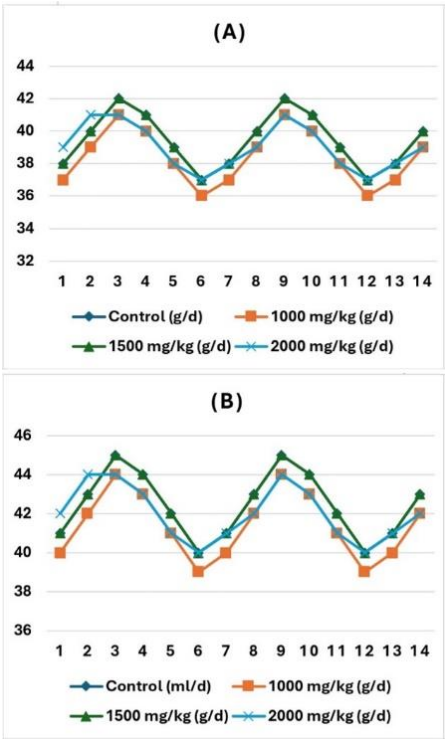



**Figure 3:** Food consumption expressed in g/d **(A)** and Water intake expressed in ml/d **(B)** of male rabbits orally administered with *S. officinalis* extract.

### 3. 4 The effect of Salvia officinalis extract on Rabbit's body weight

As demonstrated in Table 3, rabbits in the vehicle control group (Group I) exhibited a statistically significant increase in body weight from 1300 ± 20 g on day 0 to 1395 ± 25 g on day 14 (P < 0.001). In the Low-Dose Group II (1000 mg/kg), a marginal yet statistically insignificant increase in body weight was noted between day 0 and day 14 (P > 0.05). Conversely, the Mid-Dose Group (Group III) (1500mg/kg) exhibited a statistically significant reduction in body weight (P < 0.001), with a mean difference of −160±10, suggesting an early indication of systemic toxicity. In contrast to the control and low-dose groups, the High-Dose Group (Group IV) (2000 mg/kg BW) demonstrated the most severe and statistically significant reduction in body weight, dropping from 1302 ±21 g on day 0 to 1120±25 g on day 14 (P <0.001), strongly correlating with the observed mortality.

**Table 3:** The mean body weight and weight difference of male rabbits on day 0 and day 14.

| Groups | Body weight on day 0 (g) | Body weight on day 14 (g) | Mean difference ± SEM (g) | P value |
|---|---|---|---|---|
| **Group I** | 1300±20 | 1395±25 | +95±5 | <0.001 |
| **Group II** | 1305±22 | 1320±27 | +15±5 | >0.05 |
| **Group III** | 1310±20 | 1150±30 | - 160±10 | <0.001 |
| **Group IV** | 1302±21 | 1120±25 | - 182±6 | <0.001 |

***Note:** Data are expressed as mean ± SEM (n=3). The initial body weight range for all rabbits was 1000–1500 g. P-values compare Day 14 weight to Day 0 weight within the same group.*

### 3. 5 Hypoglycemic Efficacy and Basal Glucose Modulation

As summarized in Table 4, the methanolic extract exhibited a potent dose-dependent hypoglycemic effect across all treated groups compared to the control. In the Low-Dose Group II (1000 mg/kg), a significant reduction in FBG was observed at 2- and 6-hours post-administration (P < 0.05) compared to its baseline. The most pronounced reduction was recorded in the High-Dose Group IV (2000 mg/kg), where blood glucose levels dropped significantly from a baseline of 168.00 ± 1.36 mg/dL to a minimum of 60.00 ± 1.32 mg/dL at 2 hours (P < 0.001).

As illustrated in the longitudinal profile (Figure 4), the extract demonstrated a rapid onset of action within the first 6 hours, followed by a gradual stabilization phase. While all treated groups maintained lower FBG levels relative to their respective baselines through Day 14, the efficacy of the high-dose group was confounded by the previously noted systemic toxicity and mortality rate.

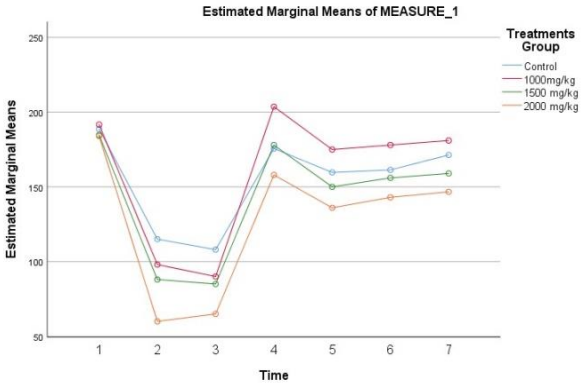



**Figure 4:** Longitudinal glycemic profile of normoglycemic rabbits following acute oral administration of *Salvia officinalis* L. extract. The graph depicts the dose-dependent reduction in fasting blood glucose (FBG) during acute (0–12 h) and sustained (1–14 days) observation periods. Error bars indicate SE.

## 4. Discussion

Medicinal plants have demonstrated significant utility within the realm of traditional medicine and hold relevance in economically disadvantaged nations [24–26]. The present study provides a preliminary characterization of the acute toxicity profile and the basal glycemic effects of a methanolic extract of *Salvia officinalis* in rabbits. By establishing the No Observed Adverse Effect Level (NOAEL) and the Maximum Tolerated Dose (MTD), the results yield preliminary insights into the safety threshold and metabolic response in a normoglycemic state associated with this botanical. Our findings revealed that the methanolic extract of *S. officinalis* diminishes blood glucose concentrations, especially at elevated dosages, while exhibiting a dose-dependent toxicity profile.

**Table 4:** Fasting blood glucose levels (mean ± SEM) in male rabbits following acute oral administration of *S. officinalis* extract.

| Group | Treatments | Blood glucose (mg/dl) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 0 | 2h | 6h | 1ST Day | 7TH Day | 10TH Day | 14TH Day |
| | | Mean ± SEM | Mean ± SEM | Mean ± SEM | Mean ± SEM | Mean ± SEM | Mean ± SEM | Mean ± SEM |
| **Group I** | Control Group | 188±2.88 A | 115±3.00 A | 108±2.65 A | 175±2.08 A | 159±0.58 A | 161±1.53 A | 171±2.31A |
| **Group II** | 1000 mg/kg | 191±2.87 A | 98±2.00 B | 90±2.00 B | 203±3.22 A | 175±3.0 B | 178±3.0 B | 181±3.0 B |
| **Group III** | 1500 mg/kg | 184±4.50 A | 88±2.00 C | 85±2.00 B | 178±1.0 A | 150±2.0 C | 156±3.0 A | 159±3.0 C |
| **Group IV** | 2000 mg/kg | 168±3.60 A | 60±2.00 D | 65±2.00 C | 158±1.0 A | 136±2.0 D | 143±3.0 C | 146±3.06 D |
| **P ± value** | | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |

***Note**: Data are expressed as Mean ± SEM (Standard Error of the Mean), with n=3 per group. Values within the same column followed by different superscript letters (a, b, c, d) are significantly different at p < 0.05, as determined by a two-way repeated measures ANOVA followed by a Bonferroni post-hoc test for multiple comparisons.*

The preliminary assessment of acute toxicity revealed a clear dose-dependent relationship between extract administration and systemic safety. While the 1000 mg/kg dose was well-tolerated, the emergence of neurobehavioral alterations and mortality at 1500 and 2000 mg/kg defines a narrow therapeutic window for the methanolic extract in this animal model. The observed clinical signs, including trembling and paralysis at high doses, suggest potential neurotoxic or systemic metabolic distress when safe thresholds are exceeded. Crucially, the significant reduction in body weight observed in Groups III and IV ($P < 0.001$) must be interpreted as a primary indicator of systemic toxicity. This finding implies an estimated MTD of approximately 1500 mg/kg, reflecting the extract's potential for dose-dependent toxicity [11] and reaffirming its safety profile at moderated dosages [7, 11]. This specific delineation of the MTD in male rabbits, alongside comprehensive clinical observations over 14 days, constitutes a critical contribution to the toxicological data of *S. officinalis*. While food and water intake remained stable, the rapid weight loss strongly correlated with the high mortality rate and clinical deterioration. This suggests that the extract, at high concentrations, may interfere with metabolic homeostasis or induce acute organ stress, highlighting the critical importance of dose optimization to avoid adverse outcomes.

Beyond the evaluation of toxicity, the present study illustrates that the methanolic extract of *Salvia officinalis* effectively lowers blood glucose levels in rabbits. Notably, the high-dose cohort (2000 mg/kg) exhibited a significant reduction in fasting blood glucose levels, a finding that is consistent with previous research highlighting the antidiabetic potential of *Salvia officinalis*. For example, previous studies [27] reported significant decreases in glucose levels in alloxan-induced diabetic rabbits subsequent to the administration of a 500 mg/kg dose of *Salvia officinalis* extract [28, 29]. The efficacy observed with the high-dose treatment in our investigation underscores the role of *Salvia officinalis* as a potent hypoglycemic agent in this model. However, the lack of a consistent dose-dependent response at elevated concentrations necessitates further exploration. It is conceivable that the intricate pharmacokinetics of the extract, encompassing absorption rates and metabolic pathways in healthy rabbits, may influence its efficacy across varying dosages.

Beyond the evaluation of phytochemical constituents, previous studies have indicated that the leaves of common sage contain elevated levels of flavonoids and saponins, among other phytochemicals [30]. The proposition that hypoglycemic activity may stem from the individual or synergistic effects of these compounds is well-documented [17, 31]. Furthermore, *S. miltiorrhiza* (Chinese sage) possesses a metabolite profile analogous to that of *S. officinalis*. Consequently, the observed declines in blood glucose levels in our study can be attributed to multiple potential mechanisms. Phytochemical analyses have identified flavonoids and saponins as key constituents known to exhibit antidiabetic properties; these may enhance insulin sensitivity and stimulate insulin secretion while concurrently diminishing hepatic glucose production [28, 32].

Additionally, the antioxidative characteristics of *Salvia officinalis* may mitigate oxidative stress, thereby safeguarding pancreatic beta-cells and sustaining insulin synthesis. Subsequent investigations delineating the specific pathways influenced by *S. officinalis* extracts could yield profound insights into its mechanistic underpinnings [33].

The findings of the current investigation are congruent with research conducted on diabetic models, such as [34], which indicated that the protective efficacy of *S. officinalis* extracts against weight reduction is contingent upon dosage and metabolic condition. The recorded weight decline at 1500 mg/kg and 2000 mg/kg in our healthy rabbits is plausibly attributable to the acute systemic stress and toxicological manifestations observed at these concentrations. This potentially culminates in the augmented degradation of structural proteins and muscular tissue, as previously elucidated within the framework of exacerbated catabolism [34, 35].

The significant reduction in fasting blood glucose observed in this study aligns with findings reported in other normoglycemic (non-diabetic) animal models. For instance, studies on healthy rats and rabbits have demonstrated that *Salvia officinalis* extract can exert a basal glucose-lowering effect without the need for pre-existing hyperglycemia. This distinguishes *S. officinalis* from some other medicinal plants that only show activity in diabetic models. The ability of the extract to modulate glucose levels in healthy subjects suggests a mechanism that enhances insulin sensitivity or inhibits glucose absorption, rather than merely replacing deficient insulin. Our results are consistent with those of [34], who observed similar hypoglycemic trends in healthy Wistar rats, further validating the potent metabolic regulatory role of this plant across different species.

Despite the significant findings of this investigation, certain limitations must be acknowledged. The sample size per group (n = 3) was relatively small; as this was a preliminary characterization of acute toxicity, the long-term impacts of *S. officinalis* extract on metabolic and organ functions were not assessed. Additionally, while a potent hypoglycemic effect was observed, the precise molecular mechanisms and the direct impact on serum insulin levels were not measured, as the study focused on initial safety and basal glucose modulation.

Based on these limitations, it is advisable that future research encompass broader studies utilizing diverse animal models of both sexes to investigate population variability in response to *S. officinalis*. Furthermore, comprehensive examinations into the sub-chronic and chronic impacts on hepatic and renal functions, alongside detailed histological assessments, are critical for an exhaustive safety evaluation. Finally, elucidating the specific signaling pathways and molecular mechanisms implicated in the observed glucose-lowering effects will furnish invaluable insights for potential clinical applications in diabetes management [36, 37].

## 5. Conclusions

This study provides preliminary experimental data that characterize the acute toxicity profile and basal glycemic effects of the methanolic extract of *Salvia officinalis* in rabbits. The findings underscore the presence of a **therapeutic window** that must be adhered to in order to maintain safety at efficacious dosages. Dose Safety and Toxicity Threshold: The No Observed Adverse Effect Level (NOAEL) for acute oral administration was successfully established at 1000 mg/kg body weight. Conversely, an observed lethal threshold was identified at approximately 2000 mg/kg, correlating with systemic toxicity, a 66.7% mortality rate, and significant body weight reduction. Basal Glucose Modulation: The extract demonstrated a significant basal blood glucose-lowering effect, even at the NOAEL dosage. Collectively, these results position *S. officinalis* as a promising candidate for further exploration in metabolic regulation, provided that dosages are optimized to avoid the identified toxicological thresholds.

## Conflict of Interest

The authors declare that they have no competing interests.

## Funding

This study was not funded by any governmental or non-governmental organizations.

## Acknowledgments

This work was supported by Thamar University, Republic of Yemen.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## CRediT authorship contribution statement

Hisham A. Al-khawlani **Conceptualization,** Hisham A. Al-khawlani; **methodology,** Hisham A. Al-khawlani.; validation, H.A.A **formal analysis**, Hisham A. Al-khawlani.; **investigation,** Aisha M. Al-Habri, Amal Y. Al-Nahary, Hadeel A. Al-mekhlafi, Hanan Y. Al-Nahary, Malak M. Al-magmhi, Maram M. Al-Mashrama, Resalh S. Jubran, Safa'a A. Amer, Shatha A. Al-Fthahy, Wala'a Y. Amran, Hisham A. Al-khawlani. Hisham A. Al-khawlani.; writing—**original draft preparation**, Hisham A. Al-khawlani.; writing—review and editing, Nabil Ali Al-Mekhlafi; visualization, Hisham A. Al-khawlani.; resources, Hisham A. Al-khawlani.

## List of Abbreviations

**FBG:** Fasting Blood Glucose, **NOAEL:** No Observed Adverse Effect Level, **MTD:** Maximum Tolerated Dose, **AEC:** Animal Ethics Committee, **SEM:** Standard Error of the Mean, **BW:** Body Weight; **DM:** Diabetes Mellitus.

## References


[1] Belhadj, S., Hentati, O., Hammami, M., Ben Hadj, A., Boudawara, T., Dammak, M., Zouari, S., El Feki, A. (2018) Metabolic impairments and tissue disorders in alloxan-induced diabetic rats are alleviated by *Salvia officinalis L.* essential oil, *Biomedicine & Pharmacotherapy* **108**: 985-995.

[2] Alarcon-Aguilar, F.J., Roman-Ramos, R., Flores-Saenz, J.L., Aguirre-Garcia, F. (2002) Investigation on the hypoglycaemic effects of extracts of four Mexican medicinal plants in normal and Alloxan-diabetic mice, *Phytotherapy Research* **16**: 383-386.

[3] Hossain, M.J., Al-Mamun, M., Islam, M.R. (2024) Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused, *Health Science Reports* **2024**: e2004.

[4] Eidi, M., Eidi, A., Zamanizadeh, H. (2005) Effect of *Salvia officinalis L.* leaves on serum glucose and insulin in healthy and streptozotocin-induced diabetic rats, *Journal of Ethnopharmacology* **100**: 310-313.

[5] Willcox, M.L., Elugbaju, C., Al-Anbaki, M., Lown, M., Graz, B. (2021) Effectiveness of Medicinal Plants for Glycaemic Control in Type 2 Diabetes: An Overview of Meta-Analyses of Clinical Trials, *Frontiers in Pharmacology* **12**.

[6] Anjali, Deepak, G., Pragi, Varun, K., Dimple, Monika (2024) A Review of Therapeutic Properties and Uses of *Salvia officinalis*, *Journal of Pharma Insights and Research* **2**: 146-154.

[7] Ghorbani, A., Esmaeilizadeh, M. (2017) Pharmacological properties of *Salvia officinalis* and its components, *Journal of Traditional and Complementary Medicine* **7**: 433-440.

[8] Ashkani-Esfahani, S., Noorafshan, A., Ebrahimi, A., Bahmani-Jahromi, M., Imanieh, M.-H., Ebrahimi, S., Hosseini, S., Tanideh, N. (2022) *Salvia Officinalis* Protects Pancreatic Beta-cells Against Streptozotocin-Induced Damage; A Stereological Study, *Jundishapur Journal of Natural Pharmaceutical Products* **17**: e109906.

[9] Ezema, C.A., Ezeorba, T.P.C., Aguchem, R.N., Okagu, I.U. (2022) Therapeutic benefits of *Salvia* species: A focus on cancer and viral infection, *Heliyon* **8**: e08763.

[10] Hasanein, P., Felehgari, Z., Emamjomeh, A. (2016) Preventive effects of *Salvia officinalis L.* against learning and memory deficit induced by diabetes in rats: Possible hypoglycaemic and antioxidant mechanisms, *Neuroscience Letters* **622**: 72-77.

[11] Rhaimi, S., Brikat, S., Lamtai, M., Ouhssine, M. (2023) Acute Oral Toxicity and Neurobehavioral Effects of *Salvia officinalis* Essential Oil in Female Wistar Rats, *Advances in Animal and Veterinary Sciences* **11**: 654–662.

[12] Sá, C.M., Ramos, A.A., Azevedo, M.F., Lima, C.F., Fernandes-Ferreira, M., Pereira-Wilson, C. (2009) Sage Tea Drinking Improves Lipid Profile and Antioxidant Defences in Humans, *International Journal of Molecular Sciences* **10**: 3937-3950.

[13] Christensen, K.B., Jørgensen, M., Kotowska, D., Petersen, R.K., Kristiansen, K., Christensen, L.P. (2010) Activation of the nuclear receptor PPARγ by metabolites isolated from sage (*Salvia officinalis L.*), *Journal of Ethnopharmacology* **132**: 127-133.

[14] Falya, Y., Sumiwi, S.A., Levita, J. (2013) Mini Review: Toxicity Study of Plant Extracts, *IOSR Journal of Pharmacy and Biological Sciences* **6**: 25-32.

[15] Amini, L., Mojab, F., Jahanfar, S., Sepidarkish, M., Raoofi, Z., Maleki-Hajiagha, A. (2020) Efficacy of *Salvia officinalis* extract on the prevention of insulin resistance in euglycemic patients with polycystic ovary syndrome: A double-blinded placebo-controlled clinical trial, *Complementary Therapies in Medicine* **48**: 102245.

[16] Ollanketo, M., Peltoketo, A., Hartonen, K., Hiltunen, R., Riekkola, M.-L. (2002) Extraction of sage (*Salvia officinalis L.*) by pressurized hot water and conventional methods: antioxidant activity of the extracts, *European Food Research and Technology* **215**: 158-163.

[17] Mokogwu, A.T.H., Adjekuko, C.O., Oshilonyah, U.H., Ikpefan, J.O., Eyenubo, O.B., Avwioro, O.G. (2022) Hypoglycaemic and Hypolipidemic Effects of Alcoholic Extract of Common Sage (*Salvia Officinalis*) In Streptozotocin –Induced Diabetic Rabbits, *African Journal of Biomedical Research* **25**: 243-247.

[18] Iravani, M., Mahinpour, R., Zahraei, Z., Toluei, Z., Asgari, F., Haghighipour, N. (2020) Comparison of cytotoxic and antioxidant activities and phenol content of four *Salvia L.* species from Iran, *Journal of Medicinal Plants* **19**: 59-68.

[19] Nutrizio, M., Gajdoš Kljusurić, J., Badanjak Sabolović, M., Bursać Kovačević, D., Šupljika, F., Putnik, P., Semenčić Čakić, M., Dubrović, I., Vrsaljko, D., Maltar-Strmečki, N., Režek Jambrak, A. (2020) Valorization of sage extracts (*Salvia officinalis L.*) obtained by high voltage electrical discharges: Process control and antioxidant properties, *Innovative Food Science & Emerging Technologies* **60**: 102284.

[20] Organisation for Economic Co-operation and Development OECD. (2025) *Guideline No. 497: Defined Approaches on Skin Sensitisation*. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, France, pp 20.

[21] Chow, P.K.H., Ng, R.T.H., Ogden, B.E. (2008) Using Animal Models in Biomedical Research, World Scientific Publication, *Singapore*, pp. 308.

[22] Obernier, J.A., Baldwin, R.L. (2006) Establishing an Appropriate Period of Acclimatization Following Transportation of Laboratory Animals, *Institute for Laboratory Animal Research (ILAR) Journal* **47**: 364-369.

[23] Javid, H., Moein, S., Moein, M. (2022) An investigationof the inhibitory effects of dichloromethane and methanol extracts of *Salvia macilenta, Salvia officinalis, Salvia santolinifola* and *Salvia mirzayanii* on diabetes marker enzymes, an approach for the treatment diabetes, *Clinical Phytoscience* **8**: 7.

[24] Al-Mekhlafi, N.A., Al-Badaii, F., Al-Ezzi, M.S., Al-Yamani, A., Almakse, E., Alfaqeeh, R., Al-Hatar, G., Al-Twity, M., Al-Masadi, M., Abdullah, M., Al-Qarhami, N. (2023) Phytochemical Analysis and Antibacterial Studies of Some Yemeni Medicinal Plants against Selected Common Human Pathogenic Bacteria, *Thamar University Journal of Natural & Applied Sciences* **8**: 14–18.

[25] Samer, B.S., Al-Mausmi, H., Gailan, A., Ziad, S., Saleh, M., Shalan, A., Hamdan, M., Gharab, A., Al-Ra'ai, A., Al-Ra'ai, N.A.-D., Al-Ezzi, M. (2024) Phytochemical Screening and Antibacterial Activity of Leaf Extracts from *Psiadia Punctulata*, *Thamar University Journal of Natural & Applied Sciences* **9**: 31–41.

[26] Abubakar, A.R., Haque, M. (2020) Preparation of Medicinal Plants: Basic Extraction and Fractionation Procedures for Experimental Purposes, *Journal of Pharmacy and Bioallied Sciences* **12**: 1-10.

[27] Lamia, A.K.R., Al-Mashhady, A. (2016) Hypoglycemic effect of *Salvia officinalis L.* extracts on induced diabetic rabbits, *International Journal of PharmTech Research* **9**: 252-260.

[28] Kanana, F.M., Maina, C.M., Kibet, J.M., Clement, J.M. (2020) Hypoglycaemic effects of *Salvia officinalis* extracts on alloxan-induced diabetic Swiss albino mice, *Journal of Medicinal Plants Research* **14**: 518-525.

[29] Skalli, S., Hassikou, R., Arahou, M. (2019) An ethnobotanical survey of medicinal plants used for diabetes treatment in Rabat, Morocco, *Heliyon* **5**: e01421.

[30] Bouteldja, R., Doucene, R., Aggad, H., Abdi, F.Z., Belkhodja, H., Belal, A., Abdali, M., Zidane, K. (2023) Phytochemical screening, acute toxicity

and antidiabetic activity of ethanolic extract of *Salvia officinalis L*. in Wistar Rat, *Agriculturae Conspectus Scientificus* **88**: 351-357.

[31] Hinad, I., S'Hih, Y., Elhessni, A., Mesfioui, A., Ouahidi, M.l. (2022) Medicinal plants used in the traditional treatment of diabetes in Ksar Elkebir Region (North-Western Morocco), *Pan African Medical Journal* **42**: 319.

[32] Gebru, A., Sibhat, G., Hiben, M. (2014) Preliminary phytochemical screening and blood glucose lowering activity of methanol extract of *Salvia tilifolia* Vahl aerial part, *International Journal of Pharmaceutical and Biological Archive* **5**: 122-125.

[33] Ben Khedher, M.R., Hammami, M., Arch, J.R.S., Hislop, D.C., Eze, D., Wargent, E.T., Kępczyńska, M.A., Zaibi, M.S. (2018) Preventive effects of *Salvia officinalis* leaf extract on insulin resistance and inflammation in a model of high fat diet-induced obesity in mice that responds to rosiglitazone, *PeerJ* **6**: e4166.

[34] Pereira, O.R., Catarino, M.D., Afonso, A.F., Silva, A.M.S., Cardoso, S.M. (2018) *Salvia elegans, Salvia greggii* and *Salvia officinalis* Decoctions: Antioxidant Activities and Inhibition of Carbohydrate and Lipid Metabolic Enzymes, *Molecules* **23**: 3169.

[35] Oliveira, G.O., Braga, C.P., Fernandes, A.A.H. (2013) Improvement of biochemical parameters in type 1 diabetic rats after the roots aqueous extract of yacon [*Smallanthus sonchifolius* (Poepp.& Endl.)] treatment, *Food and Chemical Toxicology* **59**: 256-260.

[36] Ebru K, F., A, A., K, C. (2017) Extraction and HPLC Analysis of Sage (*Salvia officinalis*) Plant, *Natural Products Chemistry & Research* **05**: 298.

[37] Dinel, A.-L., Lucas, C., Guillemet, D., Layé, S., Pallet, V., Joffre, C. (2020) Chronic Supplementation with a Mix of *Salvia officinalis* and *Salvia lavandulaefolia* Improves Morris Water Maze Learning in Normal Adult C57Bl/6J Mice, *Nutrients* **12**: 1777.

TuJNaS

*Thamar University Journal of Natural & Applied Sciences 11(1) (2026) 40– 45*

**DOI:** https://doi.org/10.59167/srw0t384

Original Article

# Immunomodulatory Effects of Carob (*Ceratonia siliqua*) Extract in Pregnant Rats: A Histological Evaluation of Embryonic Liver

**Zainab Hasan Majeed**

Department of Biology, College of Education for Pure Sciences, University of Kirkuk, Kirkuk, Iraq.

Email: zaiynab.hassan@uokirkuk.edu.iq

Received: 16 April 2026. Revised: 7 May 2026. Accepted: 10 May 2026. Published: 29 June 2026.

## Abstract

**Background**: Carob (*Ceratonia siliqua*) is a polyphenol-containing plant that has anti-oxidant and anti-inflammatory qualities, which can affect maternal immunity in pregnancy. Nevertheless, it has not been studied how it affects maternal immunity and the integrity of embryonic organs. Objective .The current paper aimed to test the immunomodulatory properties of carob extract on pregnant rats affected with induced hepatitis and to examine the histological alterations of embryonic liver tissue. **Methods**: Pregnant female Wistar rats (200-230 g) were chosen randomly and distributed into four groups (n=7/group): control (vehicle only), carob-only (400mg/kg/day carob extract), infected (induced hepatitis by concanavalin A), and carob + infected (carob pretreatment then ConA induced hepatitis). Carob aqueous extract was administered on gestational days 12-18. At gestational day 19 (GD19), hepatic enzyme levels (ALT, AST, ALP, GGT) and pro-inflammatory cytokine levels (IL-6, IL-12, IL-18, IL-29, TNF-α) were assessed in maternal blood using ELISA. Cesarean delivery was used to harvest embryos, and embryonic livers were subjected to histological analysis using hematoxylin and eosin staining. **Results**: Pretreatment with carob significantly attenuated maternal hepatic infection by reducing ALT (345 ± 30 to 120 ± 20 U/L, 65%) and AST (290 ± 25 to 105 ± 18 U/L, 64%). There was a significant decrease in pro-inflammatory cytokines, namely, IL-6 (reduced by 62.5% [320 ± 50 to 120 ± 30 pg/mL, $p < 0.01$]), TNF-α (reduced by 63% [2500 ± 400 to 920 ± 150 pg/mL, $p < 0.001$]), IL-12 (reduced by 50% [110 ± 20 to 55 ± 10, $p < 0.01$]) and IL-18 (reduction by 55% [400 ± 50 to 180 ± 30 pg/mL, $p < 0.01$]). Embryonic livers of carob-treated dams were observed to have preserved hepatic architecture with visible central veins and ordered hepatocyte cords, but those of embryos of damaged dams that did not receive carotenogenic treatment had a significant structural disorganization, cellular necrosis, and architectural distortion. **Conclusion**: Carob extract has strong immunomodulatory activity in pregnant rats, suppressing inflammatory changes and hepatic damage, which is associated with the maintenance of embryonic liver architecture. These effects have not yet been verified in other models, and the mechanisms behind them remain unclear.

**Keywords:** *Ceratonia siliqua*; Pregnancy; Embryonic Liver; Immunomodulation; Cytokines; Histology; Polyphenols

## 1. Introduction

Pregnancy is a distinctive immunological condition characterized by dynamic changes in maternal immune responses that are vital for ensuring embryonic development without compromising host defenses [1]. A significant change in the maternal immune system occurs, including alterations in the cytokine profile and immune cell populations, to develop tolerance to the semi-allogeneic embryo [2]. Disruption of this fine immunological balance by inflammatory states can negatively impact not only maternal health but also embryonic organ development [3, 4].

The cytokines that play a significant role in mediating inflammatory responses are pro-inflammatory cytokines, including interleukin-6 (IL-6), tumor necrosis factor-alpha (TNF-α), IL-12, IL-18, and IL-29 [5]. These cytokines can either bypass the placental barrier or, when elevated in maternal circulation, indirectly damage the embryo's tissue integrity by affecting placental function and nutrient transfer [6]. Such inflammatory insults are especially likely to affect the embryonic liver, which plays a major role in hematopoiesis during gestation [7]. Maternal inflammatory diseases have been linked to changes in hepatic morphology, hepatocyte disorganization, and long-term effects on offspring hepatic function [8].

Carob (*Ceratonia siliqua L.*) is a leguminous evergreen tree native to the Mediterranean region and has been used historically as a source of food and a traditional medicinal remedy [9]. Carob pods, particularly the polyphenols such as gallic acid, (+)-catechin, (-)-epicatechin, and their derivatives, are extraordinarily rich in bioactive compounds that together confer their high antioxidant activity [10, 11]. The biological effects of these polyphenolic compounds are achieved in several ways, including scavenging reactive oxygen species, blocking nuclear factor kappa-B (NF-kB) or mitogen-activated protein kinase (MAPK) signaling pathways, and regulating the production of inflammatory cytokines [12, 13].

Recent studies have shown that carob has therapeutic value in various experimental animals. According to Martić *et al.* [14], carob pulp extract reduced hepatic enzyme levels and inhibited lipid peroxidation in acetaminophen-induced hepatotoxicity. Rašković *et al.* [15] showed that carob supplementation also improved liver morphology and reduced inflammatory markers in diet-induced obese rats. Recent research by

Ahmed *et al.* [16] . Indicated that carob powder, together with thymoquinone, alleviated oxidative stress and reduced pro-inflammatory cytokines in an asthmatic pregnant rat model, suggesting its potential usefulness during pregnancy. Nevertheless, no research has examined the precise effects of carob extract on maternal immune parameters and embryonic liver histology in a setting of liver inflammation.

The immunomodulatory effects of carob have not yet been studied in pregnant models, and whether maternal treatments can influence embryonic organ integrity remains unknown. This is a knowledge gap, as there is growing interest in the use of natural compounds as potential therapeutic agents during pregnancy. Thus, we conjectured that carob extract treatment of pregnant rats would regularize the maternal inflammatory response and, by extension, prevent inflammation-related harm to the embryonic liver. This research set out to assess the immunomodulatory actions of carob extract on maternal parameters of inflammatory reactions in pregnant rats with induced immune-mediated hepatitis and to determine the resultant histological alterations of embryonic liver tissue.

## 2. Materials and Methods

### 2.1 Ethical Approval

The entire experimental procedure was performed in strict compliance with the National Institutes of Health Guidelines for the Care and Use of Laboratory Animals and was endorsed by the Institutional Animal Ethics Committee (Protocol No. IAEC/2024/05). The animals used were kept in humane conditions during the study, and suffering was reduced as much as possible.

### 2.2 Animals and Mating Protocol

Twenty-eight adult female Wistar rats (200 – 230 g, 10 – 12 weeks old) were obtained from the Animal House Facility at the College of Veterinary Medicine, University of Tikrit, Iraq. The animals were housed in polypropylene cages under controlled environmental conditions (temperature: 22 ± 2 °C; relative humidity: 50–60%; 12-hour light/dark cycle) with ad libitum access to standard rodent laboratory chow and filtered water. After a one-week acclimatization period, female rats were mated with fertile males at a ratio of 2:1. Vaginal smears were examined daily; the presence of spermatozoa was designated as gestational day 0 (GD0), and pregnancy was further confirmed by monitoring body weight gain [17].

### 2.3 Carob Extract Preparation

Mature carob (*Ceratonia siliqua L.*) pods were obtained from local markets. The pods were thoroughly washed, air-dried at room temperature in a well-ventilated area for 7 days, and ground into a fine powder using an electric grinder. To prepare the aqueous extract, 100 g of the powder was macerated in 1000 mL of distilled water at room temperature with periodic agitation every 72 hours. The suspension was filtered through Whatman No. 1 filter paper, followed by a 0.45 µm membrane filter for clarification. The filtrate was concentrated using a rotary evaporator at 40 °C under reduced pressure to obtain the crude extract (yield: 18.5% w/w). The resulting extract was lyophilized and stored at -20 °C until further use. For administration, the lyophilized powder was freshly resuspended in 0.5% carboxymethyl cellulose (CMC) as a vehicle. A dose of 400 mg/kg body weight was selected based on its established efficacy and safety in rodent models [14, 18], representing approximately 1/10 of the No-Observed-Adverse-Effect Level (NOAEL) for carob extracts in rodents [19].

### 2.4 This Experimental Design and Treatment Protocol

After pregnancy confirmation, rats were randomly assigned to four experimental groups (n = 7 per group) using a computer-generated randomization sequence:

**Group I** (Control): Received the vehicle (0.5% CMC, 1 mL/kg, p.o.) daily from GD12 to GD18, and a single injection of normal saline (1 mL/kg, i.v.) on GD18.

**Group II** (Carob Only): Received carob extract (400 mg/kg, p.o.) daily from GD12 to GD18, and normal saline (1 mL/kg, i.v.) on GD18.

**Group III** (ConA- infected): Received the vehicle (0.5% CMC, 1 mL/kg, p.o.) daily from GD12 to GD18, followed by a injection (15 mg/kg, i.v.) on GD18.

**Group IV** (Carob + ConA): Received carob extract (400 mg/kg, p.o.) daily from GD12 to GD18, followed by a ConA injection (15 mg/kg, i.v.) on GD18, administered 1 hour after the last carob dose.

The treatment period (GD12–GD18) was selected to encompass the critical stage of fetal hepatic organogenesis in rats [20].

**Induction of Immune-Mediated Hepatic Infection:** To induce acute hepatic inflammation, Concanavalin A (ConA) was used. ConA is a well-established T-cell mitogen that triggers rapid cytokine release and liver infection within 24 hours [21]. All intravenous injections were administered via the lateral tail vein under brief isoflurane anesthesia.

### 2.5 Sample Collection

On GD19 (24 hours after ConA or saline injection), pregnant rats received anesthesia with ketamine/xylazine solution (80/10 mg/kg, i.p.). The maternal blood was separated from the serum (about 5 mL) by cardiac puncture using plain tubes. Blood samples were clotted at room temperature for 30 minutes, then centrifuged at 3000 rpm for 4 °C. Serum aliquots were stored at -80 °C until biochemical and immunological tests. After the blood sample was collected, pregnant rats were euthanized by cervical dislocation, and embryos were harvested immediately through cesarean section. The viable embryo, resorption sites, and embryo weight were counted. Three random embryos were selected from each litter and subjected to histological examination. Embryonic livers were also dissected under a stereomicroscope and fixed immediately on 10% neutral-buffered formalin.

### 2.6 Maternal Hepatic Enzyme Analysis

Serum markers of hepatic infection were measured using an automated clinical chemistry analyzer (Beckman Coulter AU480, USA). The kinetic UV technique of NADH oxidation was used to measure the activities of Alanine aminotransferase (ALT) and aspartate aminotransferase (AST). Alkaline phosphatase (ALP) was ascertained using the p-nitrophenyl phosphatase technique, and gamma-glutamyl transferase (GGT) was ascertained using the γ-glutamyl-3-carboxy-4-nitroanilide technique. All assays were standardized with commercial standards and performed in duplicate; results are expressed in U/L.

### 2.7 Assessment of Cytokine Concentrations in Maternal Blood

Pro-inflammatory cytokines (IL-6, IL-12p70, IL-18, IL-29, and TNF-α) in the maternal serum were measured using commercially available rat-specific sandwich ELISA kits (Elabscience Biotechnology Co., Ltd., Wuhan, China) according to the manufacturer's instructions. In brief, 96-well precoated serum samples (diluted 1:2 in assay buffer) were incubated with capture antibody-coated microplates for 2 hours at 37°C, followed by incubation with predetermined recombinant cytokine standards. After three washing steps in phosphate-buffered saline containing 0.05 per cent Tween-20, biotin-conjugated detection antibodies were added and incubated for 1 hour at 37 °C. The solution of horseradish peroxidase (HRP)-conjugated streptavidin was introduced after washing, and the mixture was incubated at 37 °C for 30 minutes. Development was then performed using the colorimetric 3,3,5,5'-tetramethylbenzidine (TMB) substrate, and the reaction was left to develop in the dark for 15 minutes, after which it was stopped with 2N sulfuric acid. Optical density was measured at 450 nm using a microplate spectrophotometer (BioTek ELx800, USA). Cytokine levels (pg/mL) were determined using four-parameter logistic standard curves. Each sample was tested twice. The inter- and intra-assay coefficients of variation were less than 10%. Detection limits: IL-6 (5 pg/mL), IL-12p70 (10 pg/mL), IL-18 (2 pg/mL), IL-29 (15pg/mL), and TNF-α (5 pg/mL).

### 2.8 The Histological Study of Embryonic Liver.

Specimens of embryonic liver were fixed in 10% neutral-buffered formalin for 48 hours at room temperature. Fixed tissues were dehydrated in a graded series of ethanol (70, 80, 90, and 100 per cent), cleared twice in xylene, and embedded in paraffin wax. A rotary microtome (Leica RM2235, Germany) was used to prepare serial sections 5 mm thick, which were then mounted on glass slides. According to conventional histological procedures, sections were deparaffinized with xylene, rehydrated through a descending ethanol series, and stained with Harris hematoxylin and eosin (H and E) [22].

A qualified pathologist who was not aware of the experimental groups performed histological evaluation. The samples of liver were viewed using a light microscope (Olympus BX51, Japan) at magnifications of 1: 100, 1: 200, and 1: 400. The parameters that were assessed were as follows (1) overall hepatic architecture and structure; (2) the arrangement of the hepatocyte cords and morphology of the hepatocytes; (3) the visibility and integrity of the central veins; (4) the distribution of spaces in

a sinusoidal manner; (5) presence and degree of necrosis or degeneration; (6) the presence of the inflammatory cell infiltration; and (7) the presence of the hematopoietic activity (suitable to the gestational age). A digital camera system (Olympus DP72) was used to take representative photomicrographs.

The system of semi-qualitative scoring was used to evaluate the liver histology: 1, normal architecture with corded hepatocytes and visible central veins; 2, mild changes (slight disorganization, minimal affected cells); 3, moderate changes (obvious disorganization, scattered necrotic cells); 3 severe changes (assessed the significant disorganization of the architecture, extensive grade of necrosis or degeneration).

### 2.9 Statistical Analysis

GraphPad Prism (version 9.0; GraphPad Software, Inc., San Diego, CA, USA) was used to perform statistical analyses. The Shapiro-Wilk test was used to assess data normality. Data that follow a normal distribution are presented as mean ± standard deviation (SD). One-way analysis of variance (ANOVA) and the Tukey honestly significant difference (HSD) post hoc test were used to compare the groups. This was done using the Kruskal-Wallis test and Dunn's multiple comparison test, which were used to determine nonparametric scores for the histological scores. A p-value that was less than 0.05 was deemed to be statistically significant. To cause statistically significant differences between groups, power analysis (a = 0.05, b = 0.20, effect size = 1.5) was conducted with G*Power software, which also allowed determining the appropriate sample size (n = 7 per group).

The control and carob-only groups of all pregnant rats maintained normal behavior, food intake, and weight gain throughout the gestation period. Due to the administration of ConA on GD18, the damage to the dams in the infected group had clinical manifestations of acute hepatitis that included lethargy, piloerection, hunched posture, and decreased food intake within the period of 6-8 hours. Group IV exhibited a significant reduction in clinical manifestations of infected because of carob - pretreated pregnant rats compared with the non-treated infected group. In all experimental groups, there were no maternal deaths or premature births.

## 3. Results

Table 1 summarizes the results of pregnancy on GD19. There was no significant difference in the number of viable embryos per litter (10.1 – 11.4, p > 0.05). There were no significant differences in mean embryo weight (control, carob-only, carob + infected: 2.82 – 2.95 g), but the embryos of untreated injured dams were slightly but not significantly lower in weight (2.68 ± 0.31 g). Resorption rate was slightly higher in the infected group (8.5) than in controls (4.2), but the difference was not statistically significant (p = 0.08).

**Table 1.** Pregnancy outcomes on gestational GD19 (mean ± SD).

| Group | Viable Embryos/Litter | Mean Embryo Weight (g) | Resorption Rate (%) |
|---|---|---|---|
| **I. Control** | 11.4 ± 1.2 | 2.95 ± 0.28 | 4.2 ± 1.8 |
| **II. Carob Only** | 11.1 ± 1.3 | 2.89 ± 0.25 | 4.5 ± 1.6 |
| **III. ConA-Infected** | 10.1 ± 1.5 | 2.68 ± 0.31 | 8.5 ± 2.2 |
| **IV. Carob + ConA** | 11.0 ± 1.4 | 2.82 ± 0.29 | 5.1 ± 1.9 |

### 3.1 The Levels of Maternal Hepatic Enzymes

Table 2 shows the maternal serum hepatic enzyme levels. Severe hepatocellular infection in pregnant rats induced by ConA administration was evidenced by significant increases in serum transaminases. In the infected group (Group III), ALT levels rose to 345 ± 30 U/L, an increase of 7.7-fold compared with the control (45 ± 8 U/L, p < 0.001). Equally, AST increased to 290 ± 25 U/L, which was 5.8 times higher than the controls (50 ± 10 U/L, p<0.001).

These enzyme levels were considerably reduced by carob pretreatment (Group IV). The ALT decreased to 120 ± 20 U/L, which was 65 % lower than the untreated group (p < 0.01). AST reduced to 105 ± 18 U/L, which was a 64 per cent decrease (p < 0.01). Although the values remained higher than those in controls (p < 0.05), carob pretreatment provided significant protection of the liver.

There were similar patterns in cholestatic markers. The increase in ALP was 3 times (p<0.001) between 80 ± 15 U/L in controls and 240 ± 30 U/L in the infected group. Pretreatment with carob reduced ALP to 130 ± 20 U/L (45.8% inhibition of infected, p < 0.01). GGT was higher in controls 6 ± 2 U/L than in the infected group (18 ± 4 U/L, p < 0.001) and lower in carob pretreatment (5.1 per cent reduction, p < 0.01). It is important to note that the carob-only (Group II) did not show significant changes in any liver enzymes relative to controls (p > 0.05), which supported the claim that carob extract administration had no hepatotoxic effects.

**Table 2.** Serum hepatic enzyme levels of the maternal on the GD19.

| Group | ALT (U/L) | AST (U/L) | ALP (U/L) | GGT (U/L) |
|---|---|---|---|---|
| **I. Control** | 45 ± 8 | 50 ± 10 | 80 ± 15 | 6 ± 2 |
| **II. Carob Only** | 40 ± 5 | 48 ± 9 | 78 ± 20 | 5 ± 1 |
| **III. ConA-Infected** | 345 ± 30* | 290 ± 25* | 240 ± 30* | 18 ± 4* |
| **IV. Carob + ConA** | 120 ± 20*† | 105 ± 18*† | 130 ± 20*† | 8 ± 3 |

*Data are expressed as mean ± SD (n = 7 per group). ALT: alanine aminotransferase; AST: aspartate aminotransferase; ALP: alkaline phosphatase; GGT: gamma-glutamyl transferase. *p < 0.05 vs. Control group; †p < 0.01 vs. Infected group (one-way ANOVA with Tukey's post-hoc test).

### 3.2 Maternal Inflammatory Cytokine Profiles

Cytokine analysis of maternal serum revealed a high level of inflammatory response after ConA administration, which was strongly inhibited by carob pretreatment (Table 3). Pro-inflammatory cytokines were at baseline levels in the control group: IL-6 10 ± 2 pg/mL, TNF-α 22 ± 5 pg/mL, IL-12p70 5 ± 2 pg/mL, IL-18 20 ± 4 pg/mL, and IL-29 < 5 pg/mL.

Strong cytokine responses in ConA administration (Group III) were noted: IL-6 increased to 320 ± 50 pg /mL (p < 0.001 vs. control), TNF-α to 2500 ± 40 pg/mL (113-fold increase, p < 0.001), IL-12p70 to 110 ± 20 pg/mL (22-fold increase, p < 0.001), and IL-18 to 400 ± 50. These results support strong stimulation of the innate and adaptive inflammatory mechanisms.

All of the measured cytokines, IL-6 (62.5% reduction vs. Infected group, p < 0.01), TNF-α (63% reduction, p < 0.001), IL-12p70 (50% reduction, p < 0.01), IL-18 (55% reduction, p < 0.001), and IL-29 (57.9% reduction, p < 0.01) were significantly reduced by carob pretreatment (Group IV). The carob-sole sample showed no significant changes in any cytokines relative to controls (p<.05), indicating that carob extract alone did not cause any inflammatory effects.

**Table 3.** Maternal serum pro-inflammatory cytokine levels on GD19.

| Group | IL-6 (pg/mL) | TNF-α (pg/mL) | IL-12p70 (pg/mL) | IL-18 (pg/mL) | IL-29 (pg/mL) |
|---|---|---|---|---|---|
| **I. Control** | 10 ± 2 | 22 ± 5 | 5 ± 2 | 20 ± 4 | < 5 |
| **II. Carob Only** | 11 ± 3 | 24 ± 6 | 5 ± 2 | 21 ± 5 | < 5 |
| **III. ConA-Infected** | 320 ± 50* | 2500 ± 400* | 110 ± 20* | 400 ± 50* | 38 ± 8* |
| **IV. Carob + ConA** | 120 ± 30*† | 920 ± 150*† | 55 ± 10*† | 180 ± 30*† | 16 ± 4† |

*p < 0.05 vs. Control group; †p < 0.01 vs. Infected group (one-way ANOVA with Tukey's post-hoc test).

### 3.3 Histological Embryonic Liver Findings.

The histological analysis of embryonic liver tissues showed that there were morphological differences (according to the maternal treatment groups) (Figure 1). The summary of the histological scores is in Table 4. Control group (Figure 1A): Damaged embryonic livers. Standard hepatic architecture in control dams was appropriate to GD19. On histological examination, there were well-organized hepatocytes in cord-like structures emanating from visible central veins. Sinusoidal spaces were not destroyed and were regularly distributed. Hepatocytes had normal cell structures with basophilic cytoplasm and centrally placed nuclei. At this stage of gestation, hematopoietic cell clusters, which are typical of embryonic liver, were found throughout the parenchyma. No necrosis, degeneration, or inflammatory infiltration was seen. Mean histological score: 0.14 ± 0.38.

The infected group, as seen in Figure 1B), in striking contrast, embryonic livers of dams subjected to both untreated ConA-induced hepatitis exhibited serious histological defects. The classic hepatic

architecture was drastically impaired, and the organization of the hepatic cord pattern of hepatocytes was lost. Hepatocytes showed cellular disorganization, cytoplasmic vacuoles, and nuclear pyknosis, indicating cellular degeneration and necrosis. The presence of central veins was difficult to detect due to the surrounding tissue disorder. The sinusoidal spaces were irregularly distributed and typically dilated. These results show that inflammatory hepatitis of the maternal adversely affected the embryonic liver integrity. Mean histological score: 2.57 ± 0.51 ($p < 0$, 001 vs. control). In the Carob + infected group (Figure 1C), embryonic hepatic livers of dams pretreated with carob and induced hepatitis showed significantly preserved hepatic archetype. The tissue structure was very similar to embryonic liver morphology, with central veins and hepatocytes observable and organized in recognizable cords. Even though some cells with mild degenerative foci remained apparent, the architectural integrity was significantly better than in the untreated infected group. The distribution was relatively normal in the sinusoidal spaces. These results indicate that maternal carob supplementation showed protective effects on embryonic hepatic tissue despite the incidence of maternal inflammation. The treated Dams with carob extract alone (carob only group), as seen in Figure 1D, the embryonic liver histology of dams treated with carob extract alone showed no difference as compared to the control. The architectural appearance of the hepatocyte was good, with well-organized cords of hepatocytes; central veins were clearly visible, and sinusoidal spaces were not lost. There were no untoward histological alterations, indicating that the carob treatment of the maternal did not compromise the embryonic liver. Mean histological score ($p > 0.05$ vs. control) 0.14 ± 0.38.

**Table 4.** Semi-quantitative histological scoring of embryonic liver tissue in the different experimental groups.

| Group | Histological Score (0–3) | Architectural Integrity | Hepatocyte Organization | Necrosis / Degeneration |
|---|---|---|---|---|
| **I. Control** | 0.14 ± 0.38 | Normal | Normal | Absent |
| **II. Carob Only** | 0.14 ± 0.38 | Normal | Normal | Absent |
| **III. ConA-Infected** | 2.57 ± 0.51* | Severely disrupted | Disorganized | Extensive |
| **IV. Carob + ConA** | 1.00 ± 0.58*† | Moderately preserved | Partially organized | Mild/Focal |

*$p < 0.05$ vs. Control group; †$p < 0.01$ vs. Infected group (one-way ANOVA with Tukey's post-hoc test).

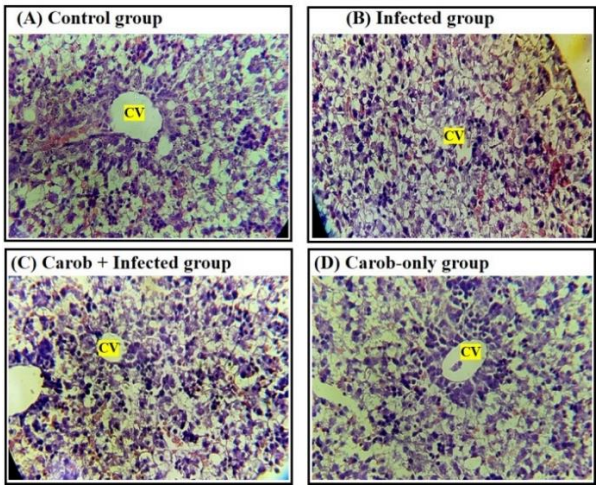



**Figure 1:** Embryonic liver sections stained with hematoxylin and eosin (400x), representative photomicrographs of the stage GD19. (A) Control group: normal embryonic hepatic architecture with well-organized hepatocytes in cord-like structures radiating at the center of the central vein (CV); complete sinusoidal spaces are observed. (B) infected group: A severe architectural disturbance that is characterized by the lack of order in hepatocytes, necrotic and degenerative processes; the central vein is not distinctly defined by the surrounding tissue disorder. (C) Carob + infected group: Significantly normal hepatic architecture with the central vein exposed, with hepatocytes being arranged into familiar cords, but some of the affected cells are also noticeable. (D) control group: Carob-only: Control liver histology. It is confirmed that maternal carob administration has no adverse effects, as embryonic liver histology is normal and similar to that of the control.

## 4. Discussion

The current study shows that the immunomodulatory power on pregnant rats with induced immune-mediated hepatitis is significant and associated with the maintenance of embryonic liver structure. The results provide new evidence of the effectiveness of dietary polyphenols in protecting fetal tissues from the negative effects of maternal inflammation.

Maternal hepatoprotective mechanisms: The significant reduction in maternal hepatic enzyme increases (65% of the increase in ALT and 64% of the increase in AST) in the presence of carob extract indicates that carob extract is a potent hepatoprotection during pregnancy. These results are consistent with previous non-pregnant models. Similar results were noted by Martić *et al.* [14], who found that transaminases significantly decreased after carob pretreatment in acetaminophen-induced hepatotoxicity, whereas Rašković *et al.* [15] found that hepatic parameters were improved in diet-induced obesity. These findings are generalized to the current study, which examines the more complex physiological background of pregnant women, where the liver's vulnerability can be modified by gestational hormonal and metabolic alterations [23].

The mechanisms underlying the hepatoprotective action of carob likely involve multiple pathways, as explained by the presence of its polyphenolic constituents. The most common polyphenols in carob are gallic acid and catechins, both of which have been well documented to have antioxidant properties capable of neutralizing reactive oxygen species generated during liver inflammation [24, 25]. Moreover, the compounds prevent NF-kB activation in hepatocytes and Kupffer cells, thereby inhibiting the transcription of pro-inflammatory genes [12, 26]. This dual action, as an antioxidant and an anti-inflammatory, could be the reason for the strong hepatoprotection observed in our study.

Immunomodulatory and cytokine-reduction: One of the key opportunities of this research is the significant reduction in maternal pro-inflammatory cytokines. The decreases in IL-6 (62.5%), TNF-α (63%), IL-12 (50%), IL-18 (55%), and IL-29 (58%) evidenced a comprehensive immunomodulatory effect. The roles of these cytokines in hepatic inflammation are interconnected: TNF-α triggers hepatocyte apoptosis and recruits inflammatory cells [27], IL-6 enhances the acute-phase response [28, 29], and Th1-mediated cytotoxicity is promoted by IL-12 and IL-18, which synergistically induce IFN-γ [30, 31]. Simultaneous decreases in these mediators indicate that carob blocks numerous branches of the inflammatory cascade.

Our findings are consistent with those of Atta *et al.* [32], who discovered that carob extract increased anti-inflammatory IL-10 and decreased TNF-α and IL-6 in a nephrotoxicity model. On the same note, Aboura *et al.* [13] also established that carob polyphenol-enriched infusions inhibited the expression of inflammatory cytokines in model colitis and obesity. This immunomodulatory profile, observed across a variety of inflammatory states, indicates a fundamental mechanism that inhibits NF-kB and MAPK signaling pathways, which regulate the expression of various pro-inflammatory mediators [12].

Embryonic liver architecture protection: The histological results of embryonic livers are the most important result of this paper. The embryos of dams of untreated hepatitis had hepatic disorganization, necrosis and loss of normal architecture, whereas those of carob-pretreated dams had hepatic integrity that was significantly preserved. This interaction between maternal immunomodulation and embryonic protection in the hepatic context implies that reduced maternal inflammatory load is translated into embryonic tissue protection. This embryonic protection may involve several mechanisms. To begin with, maternal pro-inflammatory cytokines would be reduced, and the attainment of the inflammatory load by the embryo via placental transfer or its indirect effect on placental function would be affected [32, 33]. The fact that maternal IL-6 crosses the placenta and reaches embryonic tissues has been demonstrated by Dahlgren *et al.* [34], supporting the applicability of maternal cytokine regulation. Second, carob polyphenols, in themselves, can travel across the placental barrier to act as direct antioxidants of embryonic tissues, as illustrated with structurally related flavonoids [35]. Third, better hepatic maternal function would sustain the metabolic homeostasis necessary to supply embryonic nutrients and support growth.

Development at GD19 in rat embryonic liver fundamentally involves the maturation of hepatocytes, the formation of hepatic lobular architecture, and the transformation into a metabolic organ rather than a hematopoietic organ [35, 36]. Only maternal inflammatory mediators would disrupt these processes, with effects on hepatic function in the postnatal period. The hepatic architecture of embryos treated with carob indicates that maternal supplementation could protect these processes.

Safety concerns: One of the main issues is that the carob administration in the maternal did not have a negative impact on embryonic liver histology. Carob-only embryonic liver showed no

difference from controls and exhibited normal architecture and no toxicity. This was supplemented by maternal hepatic enzymes and cytokines being the same in the carob-only group, which supports that carob extract is safe at the dose studied during the gestation period. This is in line with the generally accepted safety of carob as a food ingredient and with toxicological findings of no harm at doses up to 4000 mg/kg in rodents [19].

### 4.1 Future Directions

This study has raised several questions that require answers in future studies. Dose-response would also be useful in determining the best therapeutic dosages and safety margins. Postnatal outcome measures, such as tests of hepatic function and long-term health outcomes in offspring, would also be conducted to determine whether embryonic hepatoprotection confers long-term health benefits. The molecular mechanism of protection could be clarified through mechanistic studies of individual signaling pathways (NF-kB, MAPK, Nrf2) and placental delivery of carob components. The evaluation of anti-inflammatory cytokines and regulatory immune cell populations would better portray the immunological image. Lastly, phytochemical fractionation research might identify the bioactive compounds responsible for the noted actions and, eventually, lead to standardized therapeutic preparations.

## 5. Conclusion

Carob extract (*Ceratonia siliqua*) had strong immunomodulatory activity in rats with immune-mediated hepatitis during pregnancy, which was also indicated by a significant decrease in the markers of hepatic damage (ALT reduced by 65, AST reduced by 64%) and pro-inflammatory cytokines (IL-6, TNF-α, IL-12, IL-18, IL-29 reduced by 50 – 63%). Such maternal effects were associated with intact embryonic liver architecture, as indicated by histological analysis of intact hepatocyte organization and visible central veins in embryos of carob-treated dams, but not in embryos of untreated injured dams, which showed clear architectural disruption. Administration of carob to the maternal under isolated conditions did not cause any adverse effects on embryonic liver histology, indicating that the drug is safe at the dose studied during the gestation period in question. These results indicate that carob, with its anti-inflammatory and antioxidant effects mediated by polyphenols, may be a safe dietary supplement that can suppress maternal inflammation and safeguard embryonic liver tissue. Those are, however, only preliminary results obtained in an acute rodent model and need to be confirmed with dose-response studies, assessment of long-term offspring effects, and in other species before any clinical implications are considered. Additionally, studies are needed to clarify the mechanisms underlying these protective effects, including the molecular pathways and the bioactive compounds involved.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Funding

This study was not funded by a particular grant agency, whether in the public, commercial, or not-for-profit sector.

## Conflict of Interest

The authors declare no conflicts of interest.

## Acknowledgments

The paper humbly thanks the laboratory personnel in the Department of Histology and the institutional animal facility personnel who provided essential technical assistance in the work of the paper by diligently taking care of the experimental animals.

## References

[1] Mor, G., Cardenas, I. (2010) The immune system in pregnancy: a unique complexity, *American Journal of Reproductive Immunology* **63**: 425-433.

[2] PrabhuDas, M., Bonney, E., Caron, K., Dey, S., Erlebacher, A., Fazleabas, A., Fisher, S., Golos, T., Matzuk, M., McCune, J.M. (2015) Immune mechanisms at the maternal-fetal interface: perspectives and challenges, *Nature Immunology* **16**: 328-334.

[3] Hsiao, E.Y., Patterson, P.H. (2012) Placental regulation of maternal-fetal interactions and brain development, *Developmental Neurobiology* **72**: 1317-1326.

[4] Mahmoud, D.A., Saleh, S.S. (2021) Determination of the level of homocysteine and antioxidants in the blood serum of women with gestational diabetes, *Solid State Technology* **64**: 1129-1135.

[5] Brenner, C., Galluzzi, L., Kepp, O., Kroemer, G. (2013) Decoding cell death signals in liver inflammation, *Journal of Hepatology* **59**: 583-594.

[6] Pantham, P., Aye, I.L., Powell, T.L. (2015) Inflammation in maternal obesity and gestational diabetes mellitus, *Placenta* **36**: 709-715.

[7] Crawford, L.W., Foley, J.F., Elmore, S.A. (2010) Histology Atlas of the Developing Mouse Hepatobiliary System with Emphasis on Embryonic Days 9.5-18.5, *Toxicologic Pathology* **38**: 872-906.

[8] Ornoy, A., Reece, E.A., Pavlinkova, G., Kappen, C., Miller, R.K. (2015) Effect of maternal diabetes on the embryo, fetus, and children: congenital anomalies, genetic and epigenetic changes and developmental outcomes, *Birth Defects Research Part C, Embryo Today: Reviews* **105**: 53-72.

[9] Biner, B., Gubbuk, H., Karhan, M., Aksu, M., Pekmezci, M. (2007) Sugar profiles of the pods of cultivated and wild types of carob bean (*Ceratonia siliqua* L.) in Turkey, *Food Chemistry* **100**: 1453-1455.

[10] Papagiannopoulos, M., Wollseifen, H.R., Mellenthin, A., Haber, B., Galensa, R. (2004) Identification and quantification of polyphenols in Carob Fruits (*Ceratonia siliqua* L.) and derived products by HPLC-UV-ESI/MS$^n$, *Journal of Agricultural and Food Chemistry* **52**: 3784-3791.

[11] Goulas, V., Georgiou, E. (2019) Utilization of carob fruit as sources of phenolic compounds with antioxidant potential: Extraction optimization and application in food models, *Foods* **9**: 20.

[12] Rahman, I., Biswas, S.K., Kirkham, P.A. (2006) Regulation of inflammation and redox signaling by dietary polyphenols, *Biochemical Pharmacology* **72**: 1439-1452.

[13] Aboura, I., Nani, A., Belarbi, M., Murtaza, B., Fluckiger, A., Dumont, A., Benammar, C., Tounsi, M.S., Ghiringhelli, F., Rialland, M. (2017) Protective effects of polyphenol-rich infusions from carob (*Ceratonia siliqua*) leaves and cladodes of Opuntia ficus-indica against inflammation associated with diet-induced obesity and DSS-induced colitis in Swiss mice, *Biomedicine & Pharmacotherapy* **96**: 1022-1035.

[14] Martić, N., Zahorec, J., Stilinović, N., Andrejić-Višnjić, B., Pavlić, B., Kladar, N., Šoronja-Simović, D., Šereš, Z., Vujčić, M., Horvat, O. (2022) Hepatoprotective effect of carob pulp flour (*Ceratonia siliqua* L.) extract obtained by optimized microwave-assisted extraction, *Pharmaceutics* **14**: 657.

[15] Rašković, A., Martić, N., Tomas, A., Andrejić-Višnjić, B., Bosanac, M., Atanasković, M., Nemet, M., Popović, R., Krstić, M., Vukmirović, S. (2023) Carob extract (*Ceratonia siliqua* L.): Effects on dyslipidemia and obesity in a high-fat diet-fed rat model, *Pharmaceutics* **15**: 2611.

[16] Ahmed, A.F., Cevher, Ş.C., Peker, E.G.G., Balabanlı, B., Ebegil, M. (2024) Synergistic effects of thymoquinone and carob powder versus dexamethasone in the model of asthma in pregnant rats: new insights into their therapeutic effects, *Gazi Medical Journal* **35**: 393-400.

[17] Terry, K., Chatman, L., Foley, G., Kadyszewski, E., Fleeman, T., Hurtt, M., Chapin, R. (2005) Effects of feed restriction on fertility in female rats, *Birth Defects Research Part B: Developmental and Reproductive Toxicology* **74**: 431-441.

[18] Custódio, L., Fernandes, E., Escapa, A.L., Fajardo, A., Aligué, R., Alberício, F., Neng, N.R., Nogueira, J.M.F., Romano, A. (2011) Antioxidant and cytotoxic activities of carob tree fruit pulps are strongly influenced by gender and cultivar, *Journal of Agricultural and Food Chemistry* **59**: 7005-7012.

[19] EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA) (2010) *Scientific Opinion on the substantiation of health claims related to pectins and reduction of post-prandial glycaemic responses (ID 786), maintenance of normal blood cholesterol concentrations (ID 818) and increase in satiety leading to a reduction in energy intake (ID 4692) pursuant to Article 13 (1) of Regulation (EC) No 1924/2006*, European Food Safety Authority (EFSA Journal), 10; John Wiley & Sons, Hoboken, New Jersey, USA, https://efsa.onlinelibrary.wiley.com/doi/10.2903/j.efsa.2010.1747

[20] Si-Tayeb, K., Lemaigre, F.P., Duncan, S.A. (2010) Organogenesis and development of the liver, *Developmental Cell* **18**: 175-189.

[21] Tiegs, G., Hentschel, J., Wendel, A. (1992) A T cell-dependent experimental liver injury in mice inducible by concanavalin A, *The Journal of Clinical Investigation* **90**: 196-203.

[22] Gorovits, B., Azadeh, M., Buchlis, G., Fiscella, M., Harrison, T., Havert, M., Janetzki, S., Jawa, V., Long, B., Mahnke, Y.D. (2023) Evaluation of Cellular Immune Response to Adeno-Associated Virus-Based Gene Therapy, *An Official Journal of the American Association of Pharmaceutical Scientists (The AAPS Journal))* **25**: 47.

[23] Bancroft, J.D., Gamble, M. (2008) Theory and Practice of Histological Techniques, 6th ed., *Churchill Livingstone / Elsevier*, Philadelphia, Pennsylvania, USA, pp. 725

[24] Gao, L., Zhao, Y.C., Liang, Y., Lin, X.H., Tan, Y.J., Wu, D.D., Li, X.Z., Ye, B.Z., Kong, F.Q., Sheng, J.Z. (2016) The impaired myocardial ischemic tolerance in adult offspring of diabetic pregnancy is restored by maternal melatonin treatment, *Journal of Pineal Research* **61**: 340-352.

[25] Hasan, S., Junaid, F., Mahdi, B., Hussein, F. (2025) Therapeutic Applications of Medicinal Plants for the Treatment of Human Intestinal Diarrhea: Review Article, *South Asian Journal of Life Sciences* **13**: 20-24.

[26] Stavrou, I.J., Christou, A., Kapnissi-Christodoulou, C.P. (2018) Polyphenols in carobs: A review on their composition, antioxidant capacity and cytotoxic effects, and health impact, *Food Chemistry* **269**: 355-374.

[27] Koyama, Y., Brenner, D.A. (2017) Liver inflammation and fibrosis, *The Journal of Clinical Investigation* **127**: 55-64.

[28] Schwabe, R.F., Brenner, D.A. (2006) Mechanisms of liver injury. I. TNF-α-induced liver injury: role of IKK, JNK, and ROS pathways, *American Journal of Physiology-Gastrointestinal and Liver Physiology* **290**: G583-G589.

[29] Al-Quraghuli, A.Z.K., Mahmood, D.A. (2025) Determination of high-sensitivity C-reactive protein in chronic kidney disease patients in Al-Hawija General Hospital in Kirkuk City-Iraq, *International Journal of Biological and Pharmaceutical Sciences Archive* **10**: 173-187.

[30] Tilg, H., Moschen, A.R., Szabo, G. (2016) Interleukin-1 and inflammasomes in alcoholic liver disease/acute alcoholic hepatitis and nonalcoholic fatty liver disease/nonalcoholic steatohepatitis, *Hepatology* **64**: 955-965.

[31] Bachmann, M., Pfeilschifter, J., Mühl, H. (2018) A prominent role of interleukin-18 in acetaminophen-induced liver injury advocates its blockage for therapy of hepatic necroinflammation, *Frontiers in Immunology* **9**: 161.

[32] Atta, A.H., Atta, S.A., Khattab, M.S., El-Aziz, T.H.A., Mouneir, S.M., Ibrahim, M.A., Nasr, S.M., Emam, S.R. (2023) *Ceratonia siliqua* pods (Carob) methanol extract alleviates doxorubicin-induced nephrotoxicity via antioxidant, anti-inflammatory and anti-apoptotic pathways in rats, *Environmental Science and Pollution Research* **30**: 83421-83438.

[33] Robbins, J.R., Bakardjiev, A.I. (2012) Pathogens and the placental fortress, *Current Opinion in Microbiology* **15**: 36-43.

[34] Dahlgren, J., Samuelsson, A.-M., Jansson, T., Holmäng, A. (2006) Interleukin-6 in the maternal circulation reaches the rat fetus in mid-gestation, *Pediatric Research* **60**: 147-151.

[35] Vanhees, K., Godschalk, R.W., Sanders, A., van Doorn, S.B.v.W., van Schooten, F.J. (2011) Maternal quercetin intake during pregnancy results in an adapted iron homeostasis at adulthood, *Toxicology* **290**: 350-358.

[36] He, J., Cui, H., Shi, X., Jin, Q., Han, X., Han, T., Peng, J., Guo, S., Zhang, L., Zhao, Y. (2022) Functional hepatobiliary organoids recapitulate liver development and reveal essential drivers of hepatobiliary cell fate determination, *Life Medicine* **1**: 345-358.

TuJNaS

Volume 11, Issue 1, June 2026

# Appendices

# APPENDICES:

# A Systematic Review and Meta-analysis Survey of IDS/IPS Techniques for CAN and Vehicular Networks

**Younis Abdo Mohammed Nasser Al Shojaa[1*], Khaled Al Soufy[2]**

[1]Information Technology Department, Faculty of Engineering and Information Technology, Al-Qalam University, Ibb, Yemen.

[2]Electrical Engineering Department, Faculty of Engineering, Ibb University, Ibb, Yemen.

***Corresponding author:** Email younis.alshogaa@quni.edu.ye

# APPENDIX A

## REPRESENTATIVE MACHINE LEARNING (ML) STUDIES

**TABLE II.** REPRESENTATIVE MACHINE LEARNING (ML) STUDIES

| Authors (Year) | Technique | Dataset | Strengths | Weaknesses | Result |
|---|---|---|---|---|---|
| Kalkan and Sahingoz [1] | ML IDS | Custom | Multi-model support | Complex deployment | 95% |
| Alfardus and Rawat [2] | RF | Car-Hacking | Effective for known | Poor generalization | 95% |
| | | | attacks | | |
| Al-Janabi et al. [3] | SVM | KDD Cup | Performs well on | High false positives | 90% |
| | | | unknown attacks | | |
| Panigrahi et al. [4] | RF, SVM, etc. | NSL-KDD | RF achieves highest | Limited generalization | 99% |
| | | | accuracy | | |
| Micale et al. [5] | Context-aware | CAN | Context-based detection | Limited dataset size | >94% |
| | ML | | | | |
| Alalwany and Mahgoub [6] | Delay ML | CAN | Novel delay-based | Simulated dataset | 95% |
| | | | features | | |
| Rajapaksha et al. [7] | Context-aware | J1939 | Detects multiple attack | Simulated environment | >97% |
| | IDS | | types | | |
| Nagarajan et al. [8] | Robust ML | IoV | Secure and stable | Requires larger testing | >96% |
| Kumar and Das [9] | Supervised ML | CAV | Real-time performance | Limited unseen attack | 97% |
| | | | | coverage | |
| Ajibuwa et al. [10] | ML IDS | AVs | Detects AV-specific | Small dataset | 95% |
| | | | threats | | |
| Shahriar et al. [11] | CAN IDS | CAN | Practical design | Limited validation | 93% |
| Anthony et al. [12] | Non-tree ML | AVs | Explores rare algorithms | Scalability concerns | 94% |
| Alalwany and Mahgoub [13] | Ensemble ML | CAN | Boosts overall accuracy | Costly training phase | 96% |
| Huang et al. [14] | Anomaly ML | Vehicle | Lightweight solution | Limited data diversity | 95% |
| El-Gayar et al. [15] | Ensemble IDS | Vehicular | Collaborative | Complex deployment | 96% |
| | | | decision-making | | |
| Samir et al. [16] | ML IDS | CAN | Good accuracy | Restricted to CAN | 95% |
| Ahmed et al. [17] | ML IDS | IoV | Resistant to DoS attacks | Limited evaluation | 96% |
| Ahmad et al. [18] | ML IDS | CAV | Enhanced security | Generic evaluation | 95% |
| | | | features | | |
| Alemerien et al. [19] | Optimized ML | IoV | Optimized performance | Limited dataset | 95% |
| Kousar et al. [20] | Lightweight ML | CAN | Fast execution | Limited testing | 94% |
| Musa et al. [21] | ML IDS | IoAV | Handles class imbalance | Narrow scope | 94% |
| Adu-Kyere et al. [22] | Custom IDS | Vehicle | Real-time performance | Scalability issues | 93% |
| Abrar et al. [23] | GPS-IDS | GPS spoofing | Tailored GPS attack | Narrow focus | High |
| | | | detection | | |
| Al-Kadri [24] | CAN-MIRGU | CAN | High detection rate | Limited to CAN | 98.7% |
| Wasicek et al. [25] | Context-aware | CAN | High accuracy | Limited diversity | 97% |
| | AI | | | | |
| Ossen [26] | ASIC RF IDS | CAN | Real-time hardware | Hardware complexity | 97% |

# APPENDIX B

# REPRESENTATIVE DEEP LEARNING (DL) STUDIES

**TABLE III.** REPRESENTATIVE DEEP LEARNING (DL) STUDIES

| Authors (Year) | Technique | Dataset | Strengths | Weaknesses | Result |
|---|---|---|---|---|---|
| Liu et al. [27] | RNN | Custom | Sequential modeling | Slow training | 95% |
| Faker and Dogdu [28] | DL Ensemble | CICIDS2017 | Robust to varied attacks | High computational cost | 97% |
| | | | complex scenarios | | |
| Kumar and Sharma [29] | CNN | NSL-KDD, | Learns hierarchical | High data requirements | 99% |
| | | CICIDS2017 | features | | |
| Raza et al. [30] | Federated CNN | CICIDS2017, | Privacy-preserving | Synchronization | 99.1% |
| | | TON_IoT | detection | overhead | |
| Longari et al. [31] | RNN-AE | Car-Hacking | Lightweight temporal | Slightly lower accuracy | 98% |
| | | CAN | modeling | | |
| Altaie and Hoomod [32] | OR-CNN | NIDS V.10 2017 | High accuracy, real-time | Dataset-dependent | 99.9% |
| | | | capability | performance | |
| Shankar et al. [33] | 1D-CNN + | CICIoT2022 | Captures temporal | High computation cost | 99.6% |
| | LSTM | | features effectively | | |
| Al-Aql and Al-Shammari [34] | Hybrid | Car-Hacking | Detects sequential | Very computationally | >97% |
| | RNN-LSTM | CAN | attacks | heavy | |
| Vibhute et al. [35] | LSTM (64 units) | CIC-IDS2017 | Robust real-time | Limited dataset | 97.6% |
| | | | detection | variation | |
| Alqubaysi et al. [36] | Federated DL | Car-Hacking | Preserves privacy across | Expensive computation | 99.3% |
| | | | nodes | | |
| Wu et al. [37] | Deep Transfer | NSL-KDD, | Reuses pretrained | Survey-based execution | >99% |
| | Learning | CICIDS | knowledge | | |
| Kim and Song [38] | Embeddings + | Real CAN | Lightweight and efficient | Limited evaluation | 98–99% |
| | DL | | | scope | |
| Bilot et al. [39] | GNN | Real CAN traces | Captures relational structure | Graph overhead | Outperformed |
| Hasan et al. [40] | CAN-GraphiT | CAN bus (7 attacks) | Graph + temporal feature learning | Complex preprocessing | 98.45% |
| KS and Sujit [41] | CNN + BiLSTM + Attention | IoV/CAN | Captures deep temporal patterns | High computation cost | 99% |

# APPENDIX C

# REPRESENTATIVE HYBRID, COMPARATIVE, BENCHMARKING, AND OVERVIEW STUDIES

**TABLE IV.** Representative Hybrid, Comparative, Benchmarking, and Overview Studies

| Authors (Year) | Technique | Dataset/Scope | Strengths | Weaknesses | Result |
|---|---|---|---|---|---|
| Dayyeh et al. [42] | DT, RF, SVM | Benchmark | High detection accuracy; | Hard to distinguish | 99.48% |
| | | | proactive prevention | similar attacks | |
| Kocher and Kumar [43] | Hybrid RF + CNN | Benchmark | High accuracy via hybrid | High computational cost | ≈99% |
| | + LSTM | | ML/DL | | |
| Khraisat et al. [44] | Federated Learning | Survey | Privacy-preserving; | Sensitive to data | Survey |
| | IDS | | decentralized operation | heterogeneity; overhead | |
| Nair [45] | ML + DL | Traffic/IDS data | Interpretable traffic | Requires feature | Effective |
| | | | prediction + IDS | engineering | |
| | | | detection | | |
| Rani and Kaushal [46] | Supervised ML | Realistic datasets | Diverse attacks; realistic | Imbalanced datasets | >95% |
| | | | setting | | |
| Najafli et al. [47] | Taxonomy / Survey | Review | Comprehensive taxonomy; | No new IDS model | Review |
| | | | hybrid-focused | | |
| Sharmin et al. [48] | Benchmarking | Comparative | Structured benchmarking | No new IDS proposed | Framework |
| | Framework | framework | approach | | |
| Islam and Ali [49] | ANFIS (Hybrid ML | Benchmark | Robust (ANN + fuzzy | Limited data coverage | 99.6% |
| | + Fuzzy) | | logic) | | |
| Ali et al. [50] | Comparative | Real-time | Real-time latency tested | Simulated environment | DL > ML |
| | ML/DL | evaluation | | | |
| Alsarhan et al. [51] | Hybrid ML | Custom | Comprehensive modeling | Complex implementation | 92% |
| Note and Ali [52] | ML vs DL | Survey | Broad comparative view | High complexity | Survey |
| Rakine et al. [53] | IDS Survey | Survey | Broad overview of IDS | No experimental | Review |
| | | | techniques | validation | |
| Hnamte and Hussain [54] | Evaluation | Comparative study | Structured benchmarking | No new IDS model | Comparative |
| | Framework | | method | | |
| Yang et al. [55] | Hybrid Statistical + | Benchmark | Lightweight and accurate | Requires tuning | ≈99% |
| | NN | | | | |

## References


[1] Kalkan, S. C., & Sahingoz, O. K. (2020, July). In-vehicle intrusion detection system on controller area network with machine learning models. In *2020 11th international conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-6). IEEE.

[2] Alfardus, A., & Rawat, D. B. (2021, December). Intrusion detection system for can bus in-vehicle network based on machine learning algorithms. In *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)* (pp. 0944-0949). IEEE.

[3] Al-Janabi, M., Ismail, M. A., & Ali, A. H. (2021). Intrusion Detection Systems, Issues, Challenges, and Needs. *Int. J. Comput. Intell. Syst.*, *14*(1), 560-571.

[4] Panigrahi, R., Borah, S., Bhoi, A. K., Ijaz, M. F., Pramanik, M., Jhaveri, R. H., & Chowdhary, C. L. (2021). Performance assessment of supervised classifiers for designing intrusion detection systems: a comprehensive review and recommendations for future research. *Mathematics*, *9*(6), 690.

[5] Micale, D., Costantino, G., Matteucci, I., Fenzl, F., Rieke, R., & Patanè, G. (2022, December). Cahoot: a context-aware vehicular intrusion detection system. In *2022 IEEE international conference on trust, security and privacy in computing and communications (TrustCom)* (pp. 1211-1218). IEEE.

[6] Alalwany, E., & Mahgoub, I. (2022). Classification of normal and malicious traffic based on an ensemble of machine learning for a vehicle can-network. *Sensors*, *22*(23), 9195.

[7] Rajapaksha, S., Kalutarage, H., Al-Kadri, M. O., Petrovski, A., Madzudzo, G., & Cheah, M. (2023). Ai-based intrusion detection systems for in-vehicle networks: A survey. *ACM Computing Surveys*, *55*(11), 1-40.

[8] Nagarajan, J., Mansourian, P., Shahid, M. A., Jaekel, A., Saini, I., Zhang, N., & Kneppers, M. (2023). Machine Learning based intrusion detection systems for connected autonomous vehicles: A survey. *Peer-to-Peer Networking and Applications*, *16*(5), 2153-2185.

[9] Kumar, A., & Das, T. K. (2023). CAVIDS: Real time intrusion detection system for connected autonomous vehicles using logical analysis of data. *Vehicular Communications*, *43*, 100652.

[10] Ajibuwa, O., Hamdaoui, B., & Yavuz, A. A. (2023). A survey on ai/ml-driven intrusion and misbehavior detection in networked autonomous systems: techniques, challenges and opportunities. *arXiv preprint arXiv:2305.05040*.

[11] Shahriar, M. H., Xiao, Y., Moriano, P., Lou, W., & Hou, Y. T. (2023). CANShield: Deep-learning-based intrusion detection framework for controller area networks at the signal level. *IEEE Internet of Things Journal*, *10*(24), 22111-22127.

[12] Anthony, C., Elgenaidi, W., & Rao, M. (2024). Intrusion detection system for autonomous vehicles using non-tree based machine learning algorithms. *Electronics*, *13*(5), 809.

[13] Alalwany, E., & Mahgoub, I. (2024). An effective ensemble learning-based real-time intrusion detection scheme for an in-vehicle network. *Electronics*, *13*(5), 919.

[14] Huang, W., Xu, H., Gong, Y., Liu, Z., Li, F., Lin, Z., & Hu, B. J. (2024). UltraADV: An Unsupervised Deep Learning Lightweight Framework for Anomaly Detection in V2X. *IEEE Internet of Things Journal*, *12*(9), 12735-12747.

[15] El-Gayar, M. M., Alrslani, F. A., & El-Sappagh, S. (2024). Smart collaborative intrusion detection system for securing vehicular networks using ensemble

machine learning model. *Information*, *15*(10), 583.

[16] Samir, S. B. H., Raissa, M., Touati, H., Hadded, M., & Ghazzai, H. (2024). Machine learning-based intrusion detection for securing in-vehicle can bus communication. *SN Computer Science*, *5*(8), 1082.

[17] Ahmed, N., Hassan, F., Aurangzeb, K., Magsi, A. H., & Alhussein, M. (2024). Advanced machine learning approach for DoS attack resilience in internet of vehicles security. *Heliyon*, *10*(8).

[18] Ahmad, U., Han, M., & Mahmood, S. (2024). Enhancing security in connected and autonomous vehicles: a pairing approach and machine learning integration. *Applied Sciences*, *14*(13), 5648.

[19] Alemerien, K., Al-Suhemat, S., & Almahadin, M. (2024). Towards optimized machine-learning-driven intrusion detection for Internet of Things applications. *International Journal of Information Technology*, *16*(8), 4981-4994.

[20] Kousar, A., Ahmed, S., Altamimi, A., & Khan, Z. A. (2024). A Novel Light-Weight Machine Learning Classifier for Intrusion Detection in Controller Area Network in Smart Cars. *Smart Cities*, *7*(6), 3289-3314.

[21] Musa, U. I., Musa, A. I., Galadima, Y. I., Kantunsung, R. O., Gupta, S., & Dua, S. (2024, May). Machine Learning-Based Cybersecurity Optimization in Internet of Things-Enabled Autonomous Vehicles. In *2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR)* (pp. 1-6). IEEE.

[22] Adu-Kyere, A., Nigussie, E., & Isoaho, J. (2024). Analyzing the effectiveness of IDS/IPS in real-time with a custom in-vehicle design. *Procedia Computer Science*, *238*, 175-183.

[23] Abrar, M. M., Youssef, A., Islam, R., Satam, S., Latibari, B. S., Hariri, S., ... &

Satam, P. (2024). GPS-IDS: An anomaly-based GPS spoofing attack detection framework for autonomous vehicles. *arXiv preprint arXiv:2405.08359*.

[24] AL-KADRI, M. O. (2024). CAN-MIRGU: a comprehensive CAN bus attack dataset from moving vehicles for intrusion detection system evaluation.

[25] Wasicek, A., Pesé, M. D., Weimerskirch, A., Burakova, Y., & Singh, K. (2017, June). Context-aware intrusion detection in automotive control systems. In *Proc. 5th ESCAR USA Conf* (pp. 21-22).

[26] Ossen, S. (2025). *Enabling Low-Latency High-Throughput Real-Time Stream Processing Using Smart Network Compute Elements* (Doctoral dissertation, Indiana University).

[27] Liu, H., Lang, B., Liu, M., & Yan, H. (2019). CNN and RNN based payload classification methods for attack detection. *Knowledge-Based Systems*, *163*, 332-341.

[28] Faker, O., & Dogdu, E. (2019, April). Intrusion detection using big data and deep learning techniques. In *Proceedings of the 2019 ACM Southeast conference* (pp. 86-93).

[29] Kumar, A., & Sharma, I. (2023, March). CNN-based approach for IoT intrusion attack detection. In *2023 International conference on sustainable computing and data communication systems (ICSCDS)* (pp. 492-496). IEEE.

[30] Raza, M., Saeed, M. J., Riaz, M. B., & Sattar, M. A. (2024). Federated learning for privacy-preserving intrusion detection in software-defined networks. *Ieee Access*, *12*, 69551-69567.

[31] Longari, S., Pozzoli, C. A., Nichelini, A., Carminati, M., & Zanero, S. (2023, June). Candito: Improving payload-based detection of attacks on controller area networks. In *International symposium on cyber security, cryptology, and machine learning* (pp. 135-150). Cham: Springer Nature Switzerland.

[32] Altaie, R. H., & Hoomod, H. K. (2024). An intrusion detection system using a hybrid lightweight deep learning algorithm. *Engineering, Technology & Applied Science Research*, *14*(5), 16740-16743.

[33] Shankar, T. S., Shamsudeen, K. V., & Ramadass, R. (2025). 1D-CNN-based Real-Time Network Intrusion Detection with Privacy-Preserving for IoT. *Journal of Wireless Networks and Communication Systems*.

[34] Al-Aql, N., & Al-Shammari, A. (2024). Hybrid rnn-lstm networks for enhanced intrusion detection in vehicle can systems. *Journal of Electrical Systems*, *20*(6s), 3019-3031.

[35] Vibhute, A. D., Khan, M., Kanade, A., Patil, C. H., Gaikwad, S. V., Patel, K. K., & Saini, J. R. (2024). An LSTM-based novel near-real-time multiclass network intrusion detection system for complex cloud environments. *Concurrency and Computation: Practice and Experience*, *36*(11), e8024.

[36] Alqubaysi, T., Asmari, A. F. A., Alanazi, F., Almutairi, A., & Armghan, A. (2025). Federated learning-based predictive traffic management using a contained privacy-preserving scheme for autonomous vehicles. *Sensors*, *25*(4), 1116.

[37] Wu, W., Joloudari, J., Jagatheesaperumal, S., Kandala, N. V. P. S., Gaftandzhieva, S., Hussain, S., ... & Doneva, R. (2024). Deep transfer learning techniques in intrusion detection system-Internet of vehicles: a state-of-the-art review. *Computers, Materials, & Continua*, *80*(2), 2785.

[38] Kim, H. R., & Song, H. M. (2024). Lightweight IDS Framework Using Word Embeddings for In-Vehicle Network Security. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, *15*(2), 1-13.

[39] Bilot, T., El Madhoun, N., Al Agha, K., & Zouaoui, A. (2023). Graph neural networks for intrusion detection: A survey. *IEEe Access*, *11*, 49114-49139.

[40] Hasan, M. M., Ghose, S., & Roy, K. C. (2025). CAN-GraphiT: A Graph-Based IDS for CAN Networks using Transformer. *IEEE Access*.

[41] KS, R. R., & Sujit, B. B. (2025). Sequential-Attention Based Neural Architecture Integrating BiLSTM and Multi-Head Attention for Dynamic Anomaly Detection in IoT Environments. *International Journal of Intelligent Engineering & Systems*, *18*(8).

[42] Dayyeh, R., AlSawareah, W., Kasasbeh, B., Qaddoura, R., & Kamal, S. (2023, December). Comparative Analysis of Decision Trees, Random Forest, and k-Nearest Neighbors in Predictive Analytics for Orange Telecom's Customer Complaint Data. In *2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)* (pp. 1-6). IEEE.

[43] Kocher, G., & Kumar, G. (2022). A hybrid deep learning approach for effective intrusion detection systems using spatial-temporal features. *Adv. Eng. Sci.*, *54*(2), 1503-1519.

[44] Khraisat, A., Alazab, A., Singh, S., Jan, T., & Jr. Gomez, A. (2024). Survey on federated learning for intrusion detection system: Concept, architectures, aggregation strategies, challenges, and future directions. *ACM Computing Surveys*, *57*(1), 1-38.

[45] Nair, R. (2023). Unraveling the Decision-making Process Interpretable Deep Learning IDS for Transportation Network Security. Journal of Cybersecurity & Information Management, 12(2).

[46] Rani, D., & Kaushal, N. C. (2020, July). Supervised machine learning based network intrusion detection system for Internet of Things. In *2020 11th International conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-7). IEEE.

[47] Najafli, S., Toroghi Haghighat, A., & Karasfi, B. (2024). Taxonomy of deep learning-based intrusion detection system approaches in fog computing: a

systematic review. *Knowledge and Information Systems*, *66*(11), 6527-6560.

[48] Sharmin, S., Mansor, H., Abdul Kadir, A. F., & Aziz, N. A. (2024). Benchmarking frameworks and comparative studies of Controller Area Network (CAN) intrusion detection systems: A review. *Journal of Computer Security*, *32*(5), 477-507.

[49] Islam, M. N., & Ali, M. H. (2026). ANFIS-Based Controller and Associated Cybersecurity Issues with Hybrid Energy Storage Used in EV-Connected Microgrid System. *Energies*, *19*(4), 1103.

[50] Ali, M. L., Thakur, K., Schmeelk, S., Debello, J., & Dragos, D. (2025). Deep learning vs. machine learning for intrusion detection in computer networks: A comparative study. *Applied Sciences*, *15*(4), 1903.

[51] Alsarhan, A., Alauthman, M., Alshdaifat, E. A., Al-Ghuwairi, A. R., & Al-Dubai, A. (2023). Machine Learning-driven optimization for SVM-based intrusion detection system in vehicular ad hoc networks. *Journal of Ambient Intelligence and Humanized Computing*, *14*(5), 6113-6122.

[52] Note, J., & Ali, M. (2022). Comparative analysis of intrusion detection system using machine learning and deep learning algorithms. *Annals of Emerging Technologies in Computing (AETiC)*, *6*(3), 19-36.

[53] Rakine, I., Oukaira, A., El Guemmat, K., Atouf, I., Ouahabi, S., Talea, M., & Bouragba, T. (2025). Comprehensive Review of Intrusion Detection Techniques: ML and DL in Different Networks. *IEEE Access*.

[54] Hnamte, V., & Hussain, J. (2023). Dependable intrusion detection system using deep convolutional neural network: A novel framework and performance evaluation approach. *Telematics and Informatics Reports*, *11*, 100077.

[55] Yang, Y. M., Chang, K. C., & Luo, J. N. (2025). Hybrid neural network-based

intrusion detection system: Leveraging lightgbm and mobilenetv2 for iot security. *Symmetry*, *17*(3), 314.

# APPENDICES:

## Deep Learning for Respiratory Sound Analysis: A Systematic Review and Meta-Analysis (2019–2024)

**Shaima'a Mohammed Nasser Al-Jabali[1]*, Farhan Nashwan[2], Waleed M. Altalabi[3]**

[1]Information Technology Department, Faculty of Engineering and Information Technology, AL-Qalam University, Ibb, Yemen.

[2]Electrical Engineering Department, Faculty of Engineering, IBB University, IBB, Yemen.

[3]Biomedical Engineering Department, Sana'a Community College, Sana'a, Yemen

*__Corresponding Author:__ Shaima'a Mohammed Nasser Al-Jabali, Email: shaima.aljabali@gmail.com

# APPENDIX A
# DETAILED PER-STUDY SUMMARY (2019–2024)

For transparency and reproducibility, the complete per-study summary table is provided in this appendix. The table reports the dataset used in each study, task definition, model family, main reported performance measure, and available uncertainty information. Due to its size, the table is presented in landscape format for readability.

## APPENDIX B
## SUPPLEMENTAL PROTOCOL PACKAGE

The following protocol materials are provided to enhance transparency and reproducibility:

- Full database search strings for all sources (PubMed, IEEE Xplore, ScienceDirect, and Scopus).
- Screening checklist and study eligibility decision form.
- Standardized data-extraction template (variables, operational definitions, and coding rules).
- Meta-analytic specification sheet (effect computation rules and handling of multiple outcomes).

Availability: The full review protocol package is provided in the Supplementary Material.

**Table VI.** Condensed summary of studies (2019–2024)

| Study (Year) | Dataset | Task | Family | Effect (Acc/AUC/Score) | 95% CI / SE |
|---|---|---|---|---|---|
| Yu et al., 2025 [1] | ICBHI17 (6898 cycles) | 4c (N/Cr/Wh/Both) | bi-ResNet | Score 50.16% (10f 69.30%) | — |
| Johari et al., 2019 [2] | Clinical (43) | Bin (N vs Cr) | MFCC/LPCC+Stats | 90–100% | — |
| Nabi et al., 2019 [3] | Clinical wheeze (55) | 3c severity | ENS/KNN/SVM | PPR 100/92/94% | — |
| Chen et al., 2019 [4] | Noisy resp. | 3c (Wh/Cr/N) | ResNet (OST) | 98.79% | — |
| Neili et al., 2020[5] | Breath sound signals | Classficiation | ELM + KNN | _ | ±0.004 |
| Srivastava et al., 2021 [6] | ICBHI17 | Bin | CNN | 99.20% | ±0.5 |
| Chen et al., 2024 [7] | Review dataset | Review | Review Deep learning | — | — |
| Aulia, 2021 [8] | Lung sound dataset | Classification | EMD + GLDM | _ | ±0.004 |
| Fraiwan et al., 2021 [9] | ICBHI | Bin | CNN | 98.20% | — |
| Shuvo et al., 2020 [10] | RALE | Multi | CNN | 83.78% | — |
| Hariri and Narin, 2021 [11] | Cough dataset (COVID-19) | Bin | Deep Learning | _ | — |
| Mukherjee et al., 2021 [12] | Respiratory sounds | Screening classification | ML/DL | _ | ±0.83 |
| Kim et al., 2021 [13] | Clinical dataset | Multi | CNN | 85–86% | — |
| Satea et al., 2022 [14] | 6k files | Multi | CNN | _ | — |
| Jothi et al., 2022 [15] | Sleep apnea dataset | Detection | ML | _ | _ |
| Song & Han, 2023 [16] | Respiratory dataset | Sound Classification | Contrastive learning | _ | _ |
| Shi et al., 2023 [17] | Clinical dataset | Classification | DL | _ | _ |
| Ali et al., 2023 [18] | ICBHI17 (920/126) | Bin/4c | ANN/VGGish/OpenL3 | 72–81% | — |
| Krishnan et al., 2023 [19] | ICBHI17+Heart | 11c (lung+heart) | CNN+BDF | 99.97% | Reported |
| Dubey et al., 2023 [20] | RALE (372 ev) | 3c (Wh/Cr/N) | SVM/LSTM(+BO) | 95.70% | — |
| Lal, 2023 [21] | ICBHI17 | Bin (H vs COPD) | TL (VGGish+BiGRU) | 94% (F1=0.94) | — |
| Prabhakar & Won, 2023 [22] | ICBHI17 | 2c/3c/4c | Hybrid (HISET) | 95.39/90.61/89.27% | — |
| Aptekarev et al., 2023 [23] | Resp. DB (1371) | Bin (S/H; Asthma) | DenseNet201 (TL) | 87% / 87% | — |
| Choi & Lee, 2023 [24] | Clinical (1021 seg) | 6c | VGGish+Attn | 92.56% (F1 92.29%) | — |
| Ali et al., 2023 [25] | Hosp.+ICBHI17 | 4c (Asth/COPD/Bx/N) | LDA/kNN/DT/RF | 99.72% (RF) | — |
| Dar et al., 2023 [26] | Resp. DB | Bin (N vs Abn) | DRN+Fr-WCSO | 94.8% | — |
| Chudasama et al., 2023 [27] | ICBHI17+Coswara | 8c (mix) | RF/SVM/kNN/XGB+CNN | 90.9% (ICBHI) | — |
| Zhang et al., 2023 [28] | ICBHI17 | 6c | CNN/LSTM/BLSTM | 98.82% (LSTM) | — |
| Jaffery et al., 2023 [29] | ICBHI17 (subset) | 3c (N/Bx/Bt) | DWT+MFCC+kNN | 99.3% | — |
| Yang et al., 2023 [30] | ICBHI17 | 4c (N/Cr/Wh/Both) | BLNet (Res+GoogLeNet+Attn) | Score 72.72% | — |
| Ren et al., 2023 [31] | ICBHI17 | 4c (N/Cr/Wh/Both) | Self-Expl. NN | UAR 43.7% | — |
| Kalaiyarasan and Sridhar, 2023 [32] | Tromsø (24k rec) | Bin (Wh?/Cr?) | InceptionV3 | AUC 0.88 | 95% CI 0.84–0.92 |
| Fava et al., 2024 [33] | CTD–ILD | Bin (ILD vs non-ILD) | DNN | 97% (AUC 0.99) | ±2% (5f) |
| Zhang et al., 2024 [34] | Lung-sound (6c) | 6c | Dual-Ch. CNN–LSTM | 99.01% | — |
| Trivedi and Degadwala, 2024 [35] | EIT-labeled + RSD | Regr (airflow) | CNN+BiLSTM (ens) | Corr 0.770; RMSE 0.170 | ±SD |
| Nadkarni et al., 2024 [36] | ICBHI17 (8c) | 8c | Ensemble (CNN+XGB) | 97.57% (AUC 0.994) | — |
| Aljaddouh et al., 2024 [37] | KAUH (5c) | 5c (N/Ast/Pn/COPD/Fib) | ViT (+RMS/EBU) | 91.04% (WA) | — |
| Hakkı and Serbes, 2024 [38] | ICBHI17 (aug) | Wheeze det. | CNN–BiLSTM | 82–84% | — |
| Geetha and Benadict Raja, 2024 [39] | ICBHI17 (Kaggle) | Bin; 5c | CNN + metaheuristic | 93.3% (bin); 75.0% (5c) | — |
| Khan et al., 2024 [40] | ICBHI17 (+aug) | 8c (also 4/3/2c) | Parallel CAE + LSTM | 94.16% (8c) | — |
| Sang et al., 2024 [41] | Wearable accel. (52) | Wheeze det. (bin) | 2D CNN (mel) | AUC 0.986 (≈95%) | — |
| Xu et al., 2024 [42] | TB cough (456) | Bin (TB vs H) | Bi-LSTM+Conv2D | 96.33% | — |
| Wu et al., 2024 [43] | ICBHI17 | 4c (N/Cr/Wh/Both) | Impr. Bi-ResNet | 77.81% (F1 71.05%) | — |
| Wu et al., 2024 [44] | CoCross ICU (171) | ICU/90d outcome | VAE–ResNet50 + LSTM | AUC 0.759 / 0.752 | — |
| Wang & Sun, 2024 [45] | ICBHI17 | 4c (N/Cr/Wh/Both) | CNN (spec.) | 88.7% (F1 91.2%) | — |
| Mang et al., 2024 [46] | ICBHI17 | 2c (Wh/Cr); 4c | ViT + Cochleogram | 67.9% (4c); 85.9%, 75.5% | — |
| Patel et al., 2024 [47] | RespDB@TR (42/504) | COPD severity (5c) | LSTM/CNN/MLP | 87% (LSTM) | — |
| Wanasinghe et al., 2024 [48] | ICBHI17+Mendeley | 10c | Light CNN (MFCC+Mel+Chroma) | 91.04% (5f 89.2%) | — |

| Study (Year) | Dataset | Task | Family | Effect (Acc/AUC/Score) | 95% CI / SE |
|---|---|---|---|---|---|
| Saeed et al., 2024 [49] | Clinical cough (RT-PCR) | Bin (COVID vs N) | RBF-Net (CNN-LSTM) | 0.879 (unbiased) | — |
| Isangula & Haule, 2024 [50] | Tanzania (protocol) | Diagnostic triage | AI audio CDSS | — | — |
| Kim et al., 2025 [51] | ICBHI17 (920/126) | MTL: 4c sounds + 6c disease | MobileNet (SIMO-MTL) | 74% (LS); 91% (LD) | — |
| Álvarez Casado et al., 2024 [52] | ICBHI17 | Bin PvsH; COPD vs others; 4/6c | EMD+Spectral+MLP | BA 89%; 94%; 71.7% | — |
| Bhushan et al., 2024 [53] | ICBHI17 (920/6898/126) | 4c (N/Cr/Wh/Both) | CNN–LSTM + Self-Attn | 58.62% (Score 57.02%) | — |

**Table VII.** Distribution of studies included in qualitative synthesis by Dataset, Architecture, and Task (N=98)

| Dataset | Architecture | Task | Count |
|---|---|---|---|
| ICBHI 2017 | CNN | Binary | 12 |
| ICBHI 2017 | CNN | Multiclass | 10 |
| ICBHI 2017 | CNN | Sound-level | 6 |
| ICBHI 2017 | CNN | Severity | 3 |
| ICBHI 2017 | RNN/LSTM | Binary | 5 |
| ICBHI 2017 | RNN/LSTM | Multiclass | 4 |
| ICBHI 2017 | RNN/LSTM | Sound-level | 3 |
| ICBHI 2017 | RNN/LSTM | Severity | 1 |
| ICBHI 2017 | Transformer/Hybrid | Binary | 6 |
| ICBHI 2017 | Transformer/Hybrid | Multiclass | 6 |
| ICBHI 2017 | Transformer/Hybrid | Sound-level | 3 |
| ICBHI 2017 | Transformer/Hybrid | Severity | 2 |
| ICBHI 2017 | Classical (SVM/RF) | Binary | 3 |
| ICBHI 2017 | Classical (SVM/RF) | Multiclass | 3 |
| ICBHI 2017 | Classical (SVM/RF) | Sound-level | 2 |
| ICBHI 2017 | Classical (SVM/RF) | Severity | 0 |
| Clinical/Private | CNN | Binary | 4 |
| Clinical/Private | CNN | Multiclass | 2 |
| Clinical/Private | CNN | Sound-level | 1 |
| Clinical/Private | CNN | Severity | 1 |
| Clinical/Private | RNN/LSTM | Binary | 2 |
| Clinical/Private | RNN/LSTM | Multiclass | 1 |
| Clinical/Private | RNN/LSTM | Sound-level | 1 |
| Clinical/Private | RNN/LSTM | Severity | 0 |
| Clinical/Private | Transformer/Hybrid | Binary | 2 |
| Clinical/Private | Transformer/Hybrid | Multiclass | 2 |
| Clinical/Private | Transformer/Hybrid | Sound-level | 1 |
| Clinical/Private | Transformer/Hybrid | Severity | 1 |
| Clinical/Private | Classical (SVM/RF) | Binary | 1 |
| Clinical/Private | Classical (SVM/RF) | Multiclass | 1 |
| Clinical/Private | Classical (SVM/RF) | Sound-level | 0 |
| Clinical/Private | Classical (SVM/RF) | Severity | 0 |
| BRACETS | CNN | Binary | 0 |
| BRACETS | CNN | Multiclass | 1 |
| BRACETS | CNN | Sound-level | 1 |
| BRACETS | CNN | Severity | 0 |
| BRACETS | Transformer/Hybrid | Binary | 0 |
| BRACETS | Transformer/Hybrid | Multiclass | 1 |
| BRACETS | Transformer/Hybrid | Sound-level | 0 |
| BRACETS | Transformer/Hybrid | Severity | 0 |
| BRACETS | Classical (SVM/RF) | Binary | 1 |
| BRACETS | Classical (SVM/RF) | Multiclass | 0 |
| BRACETS | Classical (SVM/RF) | Sound-level | 0 |
| BRACETS | Classical (SVM/RF) | Severity | 0 |
| RSD | CNN | Binary | 0 |
| RSD | CNN | Multiclass | 0 |
| RSD | CNN | Sound-level | 1 |
| RSD | CNN | Severity | 0 |
| RSD | Transformer/Hybrid | Binary | 0 |
| RSD | Transformer/Hybrid | Multiclass | 1 |
| RSD | Transformer/Hybrid | Sound-level | 0 |
| RSD | Transformer/Hybrid | Severity | 0 |
| RSD | Classical (SVM/RF) | Binary | 1 |
| RSD | Classical (SVM/RF) | Multiclass | 0 |
| RSD | Classical (SVM/RF) | Sound-level | 0 |
| RSD | Classical (SVM/RF) | Severity | 0 |
| HF_Lung_V1 | CNN | Binary | 1 |
| HF_Lung_V1 | CNN | Multiclass | 0 |
| HF_Lung_V1 | CNN | Sound-level | 0 |
| HF_Lung_V1 | CNN | Severity | 0 |
| HF_Lung_V1 | Transformer/Hybrid | Binary | 0 |
| HF_Lung_V1 | Transformer/Hybrid | Multiclass | 1 |
| HF_Lung_V1 | Transformer/Hybrid | Sound-level | 0 |
| HF_Lung_V1 | Transformer/Hybrid | Severity | 0 |

Note: This table reflects studies included in the qualitative synthesis (N = 98). The subset used for quantitative synthesis (N = 42) was further filtered based on the availability of comparable evaluation metrics.

**Table VIII.** Presents the complete characteristics of the 42 studies included in the quantitative synthesis, including dataset, task, model family, reported metrics, and validation-related methodological indicators.

| Ref No. | Year | Task group | Eligible for primary accuracy meta-analysis? | Dataset | Sample size /recordings / subjects | Task | Model family | Native accuracy value used | AUC reported? | F1 reported? | Split type / validation design | External validation | Key note/ pooling status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [54] | 2022 | sound classification | Provisional Yes | ICBHI 2017 + multi-channel lung sound dataset | ICBHI: 920 recordings, 126 subjects, 6898 cycles | Adventitious lung sound classification + respiratory disease classification | Transfer learning / ResNet / co-tuning / StochNorm | — (score-based result; no clean native accuracy) | No | No | Official ICBHI 60/40 non-overlapping patients; other validation on private multi-channel set | No clear external validation | Primary ICBHI result score-based; descriptive/subgroup use |
| [55] | 2022 | sound / disease classification | Yes | ICBHI, Coswara breathing/speech/cough, combined datasets | Multi-dataset evaluation; D1 = ICBHI, D5 = ICBHI + Coswara | Lung abnormality and COVID-19 diagnosis from audio | Deep ensemble with A-CRNN, A-BiLSTM, A-BiGRU, CNN + PSO | 0.9551 for D1; 0.9444 for D5 | No | No | Subject-independent train-test split explicitly reported for D5; multiple datasets used | No true external validation in strict independent-clinical sense | Choose one task/dataset only if pooled |
| [56] | 2025 | wheeze detection | Provisional Yes | Single-center pediatric dataset (Korea) | 76 children; 103 wheeze + 184 non-wheeze sounds | Binary wheeze detection | Kernel SVM with multiple feature extractors | 0.897 | Yes | Yes | 80/20 split with 5-fold CV and grid search on training | No | Useful for secondary subgroup; subject independence not fully resolved |
| [57] | 2025 | severity classification | Provisional Yes | Respiratory Database @TR + ICBHI | @TR: 42 patients, 504 recordings, 387 selected; ICBHI: 920 recordings, 877 selected | COPD severity classification | Classical ML + eigenspectral / SVD projection | 0.756 (best SVC) | Yes | Yes | 80/20 training/testing + 100 random stratified 5-fold stability analysis | No external validation | Severity subgroup, not sound-event primary pool |
| [34] | 2024 | sound / disease classification | Provisional Yes | ICBHI 2017 | After removing classes: 917 samples; after augmentation/sampling: 5054 instances | 6-class lung sound classification | Dual-channel CNN-LSTM | 0.9901 | No clear summary | Yes | Random split into train/validation/test after augmentation/sampling; 5-fold CV | No | High leakage risk due to augmentation + random split |
| [58] | 2024 | sound classification | Provisional Yes | ICBHI 2017 | 920 recordings, 126 patients, 6898 cycles; balanced to ~1200 samples shown in tables | 4-class respiratory sound classification | Comparative ML: SVM, RF, KNN, NB, DNN | 0.79 (Random Forest best) | No | No | Random splits 60/40, 80/20, 90/10 after preprocessing/balancing | No | Proceedings / unclear peer review; non-patient-wise splitting |
| [59] | 2024 | disease detection / multiclass condition identification | Provisional Yes | Respiratory sound database from Kaggle / 920 recordings from 126 patients | 920 recordings, 126 patients; 70/30 train/test | Binary respiratory condition detection + multiclass condition identification | CNN optimized by metaheuristics | 0.933 binary; 0.750 multiclass | No clear summary | Yes | 70/30 split on mel-spectrogram images; patient independence not clear | No | Task heterogeneity; use one binary accuracy if pooled |
| [33] | 2024 | disease detection | Provisional Yes | CTD-ILD + RA-ILD + RespiratoryDatabase@TR | CTD-ILD: 84 patients/670 auscultations; RA-ILD: 137 | Good/bad auscultation cleaning followed by ILD diagnosis | Signal cleaning + ML / DNN ensemble | 0.97 on CTD-ILD cleaned by KNN | Yes | Yes | 5-fold cross-validation; training/testing stated as distinct for | No clear external validation | Disease-detection subgroup, not sound-event primary pool |

| Ref No. | Year | Task group | Eligible for primary accuracy meta-analysis? | Dataset | Sample size /recordings / subjects | Task | Model family | Native accuracy value used | AUC reported? | F1 reported? | Split type / validation design | External validation | Key note/ pooling status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | patients/820 auscultations; @TR: 504 files | | | | | | DNN | | |
| [1] | 2025 | sound classification | Provisional Yes | ICBHI 2017 | 920 recordings, 126 subjects, 6898 cycles | 4-class adventitious lung sound classification | Deep learning bi-ResNet with STFT + wavelet fusion | 0.5279 official ICBHI; 0.6744 random 10-fold | No | No | Official ICBHI split plus separate random 10-fold evaluation | No | Use official split result only |
| [3] | 2019 | severity classification | Provisional Yes | Respiratory sound dataset (non-ICBHI) | Information insufficient from accessible source | Asthma severity classification | Feature-based ML (power features + classifier) | Accuracy reported but exact value not extractable | No | No | Information insufficient | No | Keep in subgroup / descriptive only |
| [4] | 2019 | sound classification | Provisional Yes | RALE database | Information insufficient from accessible source | Respiratory sound classification | SVM | Accuracy reported; exact extraction uncertain | No | No | Information insufficient | No | Needs source verification before pooling |
| [8] | 2021 | sound classification | Provisional Yes | ICBHI dataset | Information insufficient from accessible source | Lung sound classification | Boosted decision trees | Accuracy reported; exact extraction uncertain | No | No | Information insufficient | No | Needs split verification before pooling |
| [11] | 2021 | disease detection / severity classification | Provisional Yes | Clinically validated cough dataset (Spain + Mexico) | 8380 cough samples (2339 positive, 6041 negative) | COVID-19 detection + severity classification | Deep learning (DeepCough) | — (AUC/F1-focused, no clean native accuracy) | Yes | Yes | 10-fold cross-validation (stratified) | No clear external validation | AUC/F1 table, not accuracy primary pool |
| [13] | 2021 | sound classification | No | Clinical dataset (non-ICBHI) | 1918 recordings / 871 subjects / 2840 raw sounds | Multi-class (normal/crackles/ wheezes/rhonchi) | CNN + transfer learning (VGG16) | 0.865 binary; 0.857 multi-class | Yes | Yes | Random split + 5-fold CV (80/20) | No | Excluded from strict primary pool due possible leakage/non-patient-wise split |
| [18] | 2023 | wheeze detection / sound classification | No | ICBHI | 126 patients / 920 recordings | Sound classification | CNN + ML embeddings (MusicANN/VGGish/OpenL3) | 0.81 (max) | Unclear | No | 5-fold CV (patient split unclear) | No | Excluded from strict primary pool |
| [19] | 2023 | disease detection | No | ICBHI + Yaseen | 126 + 1000 records | Multi-class disease classification | CNN + Random Forest (BDF) | 0.9994 | No | Yes | Random 80/20 + 10-fold | No | Excluded from strict primary pool due augmentation leakage risk |
| [20] | 2023 | sound classification | Provisional Yes | RALE database | Approx. 372 sound instances (Normal 50, Wheeze 252, Crackle 70) | 3-class adventitious sound classification | SVM / LSTM / LSTM + Bayesian Optimization | 0.95699 | No | Yes | 5-fold setting mentioned; subject independence unclear | No | Keep in descriptive/subgroup analysis |
| [21] | 2023 | sound classification | Provisional Yes | Respiratory Sound Database from Kaggle / ICBHI-based | ICBHI described as 920 audio samples/126 subjects/6898 cycles; visible | Binary healthy vs COPD recognition | Transfer learning / VGGish + stacked BiGRU | No clean single final accuracy extractable with | No | Yes | 15% validation split; patient-wise separation unclear | No | Relevant but not cleanly extractable for strict pooling |

| Ref No. | Year | Task group | Eligible for primary accuracy meta-analysis? | Dataset | Sample size /recordings / subjects | Task | Model family | Native accuracy value used | AUC reported? | F1 reported? | Split type / validation design | External validation | Key note/ pooling status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | experiment 36 healthy + 36 COPD files | | | confidence | | | | | |
| [22] | 2023 | sound classification | Provisional Yes | ICBHI dataset | 920 recordings / 126 subjects / 6898 cycles mentioned in context | 2-class, 3-class, and 4-class respiratory sound classification | Multiple hybrid pipelines | 0.8927 (best 4-class) | No clear single summary | No clear single summary | Validation design not clearly captured | No | Methodologically heterogeneous; subgroup/descriptive use |
| [23] | 2023 | disease detection | No | Private asthma-focused respiratory sound database | 1371 recordings total: 1238 patient + 133 healthy | Binary diagnosis: sick/healthy and asthmatic/non-asthmatic+healthy | DenseNet201 with transfer learning | 0.87 | No | No | 50-record test groups + 5-fold CV on clips | No | Clip/block leakage concern; not strict primary pool |
| [24] | 2023 | disease detection | No | Single-center clinical respiratory sound dataset | 126 respiratory sounds segmented into 1021 five-second samples | 6-class lung disease classification | Improved VGGish + Light Attention | 0.9256 | No | Yes | 5-fold CV on segmented samples; patient independence unclear | No | Strong descriptive study, not strict primary pool |
| [25] | 2023 | disease detection | No | Hospital-collected dataset + ICBHI dataset | Hospital dataset: 120 patients / 368 recordings; also evaluated on ICBHI | 4-class pulmonary disease classification | Classical ML: LDA / k-NN / DT / RF | 0.9972 (RF best) | Yes | Yes | 10-fold CV + 20% holdout / 80–20 split | No clear true external validation | Very high accuracy; not strict primary pool |
| [26] | 2023 | disease detection | No | Respiratory sound database (not clearly specified) | Information insufficient from accessible text | Binary pulmonary abnormality detection | Fr-WCSO-based Deep Residual Network | 0.948 | No | No | Information insufficient | No | Dataset identity/validation not sufficient for strict pooling |
| [28] | 2023 | disease detection | No | Respiratory Sound Database / ICBHI respiratory audio files | 920 recordings / 126 patients originally; after class filtering counts sum to 847 | 6-class pulmonary disease classification | CNN / LSTM / CNN-LSTM / CNN-BLSTM | 0.9882 (best LSTM) | No | Yes | Random 80:20 split after over/under-sampling and class removal | No | Not strict primary pool |
| [30] | 2023 | sound classification | Provisional Yes | ICBHI 2017 | 920 recordings, 126 patients, 6898 cycles | 4-class adventitious respiratory sound classification | BLNet with STFT + wavelet fusion + self-attention | 0.7272 best random 8:2; official split score 0.5198 | No | No | Official train-test split plus random 8:2 split | No | High risk if using random split result |
| [31] | 2023 | sound classification | No | ICBHI 2017 | 920 recordings, 126 subjects, 6898 cycles | 4-class respiratory sound classification with interpretability focus | Self-explaining neural network | — (AS/UAR-based; no clean native accuracy) | No | No | Official ICBHI 60/40 split; training set further split 70/30 | No | Keep in descriptive metric-specific analysis |
| [32] | 2023 | wheeze detection | No | Tromsø Study lung sound files + validation sets | Training: 24,198 files from 4,033 participants; validation A: 615; validation B: 120 | Wheeze and crackle detection | Deep learning ensemble based on InceptionV3 | — | Yes | Yes | 5-fold development + two validation sets not used in training | Yes, independent validation sets used | Preprint + detection task; descriptive only |

| Ref No. | Year | Task group | Eligible for primary accuracy meta-analysis? | Dataset | Sample size /recordings / subjects | Task | Model family | Native accuracy value used | AUC reported? | F1 reported? | Split type / validation design | External validation | Key note/ pooling status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [36] | 2024 | disease detection / multiclass condition identification | Provisional Yes | ICBHI 2017 respiratory sound database | 920 respiratory sounds | Respiratory disease classification from audio | Ensemble learning: MultiFeature CNN + XGBoost + soft voting | 0.9757 | Yes | No | Stratified 80/20 train-test split with augmentation; patient-wise independence unclear | No | arXiv / split clarity issue; sensitivity candidate |
| [37] | 2024 | disease detection / multiclass condition identification | Provisional Yes | Clinical lung sound dataset represented as spectrogram images | Original samples: 35/32/5/9/5 by class; augmented totals 370/407/75/188/84 | 5-class respiratory disease classification | Vision Transformer (ViT) | 0.9104 | No clear summary | No | Training design not sufficiently detailed; augmentation experiments reported | No clear external validation | Descriptive only unless full split details recovered |
| [38] | 2024 | wheeze detection | Provisional Yes | ICBHI-derived wheeze event data + synthetic data | ICBHI: 920 recordings/126 subjects; synthetic set: 1500 training + 200 testing audio files | Wheeze event detection / localization | CRNN family: CNN-LSTM/GRU/BiLSTM/BiGRU | 0.84 (best accuracy) | No | Yes | 70/30 split before synthetic generation; event-level labeling | No | Secondary wheeze subgroup; synthetic-data heavy |
| [40] | 2024 | disease detection / multiclass condition identification | Provisional Yes | ICBHI 2017 | 920 audio samples, 126 individuals, 6898 cycles | 8-class disease, 4-class disease, and binary normal/abnormal classification | Hybrid deep learning with fused latent features + LSTM | 0.9416 (8-class); 0.7961 (4-class); 0.8561 (binary) | No clear summary | Yes | Split design insufficiently captured from accessible text | No | Multiple tasks/settings; descriptive or subgroup only |
| [41] | 2024 | wheeze detection | No | Single-center wearable accelerometer patch dataset | 52 patients with respiratory diseases | Wheeze detection from pulmonary-induced vibrations | Deep learning model vs deterministic time-frequency method | 0.95 | Yes | No clear summary | Test set evaluation on collected patient data; exact split not fully captured | No clear external validation | Biosignal patch task; not lung sound primary pool |
| [42] | 2024 | disease detection | No | Hospital-collected cough dataset | 230 cough sounds from 70 TB patients and 226 from 74 healthy subjects | Pulmonary tuberculosis screening from cough sound | Feature fusion with Bi-LSTM / Bi-GRU and 2D convolution | 0.9633 | No | No | Split details not clearly captured | No | Cough subgroup only |
| [43] | 2024 | sound classification | Provisional Yes | ICBHI 2017 | ICBHI 2017 used; full sample detail not restated in accessible text | Lung sound classification | Improved Bi-ResNet combining CNN and ResNet | 0.7781 | No | Yes | Data augmentation used; exact patient-wise split details unclear | No | Relevant but split details limited |
| [44] | 2024 | disease detection | No | CoCross ICU COVID-19 multimodal database | 171 ICU patients; 3477 auscultations; plus CXRs and ICU variables | ICU mortality / 90-day mortality prediction | Multimodal deep learning fusion | 0.761 ICU mortality; 0.743 90-day mortality | Yes | No clear summary | Longitudinal multimodal workflow | No clear external validation | Prognostic multimodal study; outside primary pool |
| [45] | 2024 | sound classification | Provisional Yes | ICBHI 2017 | Augmented and balanced ICBHI dataset | Abnormal/normal lung sound classification parameter sensitivity study | CNN with spectrogram or MFCC under variable parameters | Accuracy emphasized but no clean single extracted value | No | No | Experimental comparison under variable parameters; split not fully captured | No | Methodological study, not clean comparable benchmark |

| Ref No. | Year | Task group | Eligible for primary accuracy meta-analysis? | Dataset | Sample size /recordings / subjects | Task | Model family | Native accuracy value used | AUC reported? | F1 reported? | Split type / validation design | External validation | Key note/ pooling status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [46] | 2024 | sound classification | Provisional Yes | ICBHI dataset | ICBHI dataset; exact sample detail not fully repeated in accessible text | Adventitious sound classification | Vision Transformer fed by cochleogram | Accuracy emphasized but no clean single extracted value | No clear summary | No clear summary | Split details insufficiently captured | No | Needs full text confirmation for strict pooling |
| [47] | 2024 | severity classification | No | RespiratoryDatabase@TR | 504 wav files; 42 COPD patients; five severity classes | COPD severity classification | Multi-feature fusion with MLP, CNN, RNN, LSTM | 0.87 best testing accuracy | No | No | Dataset balancing via SMOTE; exact subject-independent split not clearly captured | No | Severity subgroup only |
| [48] | 2024 | disease detection / multiclass condition identification | No | ICBHI 2017 + Fraiwan et al. Mendeley dataset | Combined ICBHI and multiclass Mendeley lung sound dataset | 10-class lung disease classification | Lightweight CNN with stacked Mel/MFCC/Chromagram | 0.9104 | No clear summary | No clear summary | Combined-dataset classification; split not fully captured | No | Descriptive only |
| [60] | 2024 | cough detection | No | Accelerometer-based experimental dataset | 23 participants; original 206 cough + 1011 non-cough; augmented to 2266 cough + 2022 non-cough | Cough vs non-cough detection | Deep learning best with SqueezeNet + wavelet scalograms | 0.9221 | No clear summary | No clear summary | 1-second segmented samples with manual segmentation and augmentation | No | Accelerometer cough subgroup only |
| [49] | 2024 | disease detection | No | Clinically tagged proprietary cough dataset with RT-PCR COVID status | Initially 1094 positive and 3761 negative participants; after preprocessing 1022 positive and 2656 normal samples | COVID-19 diagnosis from cough audio under confounder-aware training | Bias-Free Network (RBF-Net) | 0.841 gender-biased; 0.846 age-biased; 0.805 smoking-biased | No clear summary | No clear summary | Different intentionally biased training scenarios | No clear external validation | Cough COVID subgroup only |
| [53] | 2024 | sound classification | Provisional Yes | ICBHI'17 database | ICBHI'17 used; exact final sample count not restated in accessible snippet | 4-class respiratory sound classification | Self-attention based hybrid CNN-LSTM | Accuracy not cleanly extractable from accessible snippet | No | No clear summary | Explicitly motivated as patient-independent; exact final split/accuracy needs full text | No | Needs full-text confirmation before strict pooling |

## References


[1] Yu, S., Yu, J., Chen, L., Zhu, B., Liang, X., Xie, Y., Sun, Q. (2025) Advances and Challenges in Respiratory Sound Analysis: A Technique Review Based on the ICBHI2017 Database, *Electronics* **14**: 2794.

[2] Johari, N.M., Malik, N.A., Sidek, K. (2019) Distinctive features for normal and crackles respiratory sounds using cepstral coefficients, *Bulletin of Electrical Engineering and Informatics* **8**: 875-881.

[3] Nabi, F.G., Sundaraj, K., Lam, C.K. (2019) Identification of asthma severity levels through wheeze sound characterization and classification using integrated power features, *Biomedical Signal Processing and Control* **52**: 302-311.

[4] Chen, H., Yuan, X., Pei, Z., Li, M., Li, J. (2019) Triple-classification of respiratory sounds using optimized s-transform and deep residual networks, *IEEE Access* **7**: 32845-32852.

[5] Neili, Z., Fezari, M., Redjati, A. (2020) ELM and K-nn machine learning in classification of Breath sounds signals, *International Journal of Electrical and Computer Engineering* **10**: 3528-3536.

[6] Srivastava, A., Jain, S., Miranda, R., Patil, S., Pandya, S., Kotecha, K. (2021) Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease, *PeerJ Computer Science* **7**: e369.

[7] Chen, J., Guo, Z., Xu, X., Jeon, G., Camacho, D. (2024) Artificial intelligence for heart sound classification: A review, *Expert Systems* **41**: e13535.

[8] Aulia, S. (2021) Abnormal ECG classification using empirical mode decomposition and entropy, *Jurnal Rekayasa Elektrika* **17**: 191-198.

[9] Fraiwan, L., Hassanin, O., Fraiwan, M., Khassawneh, B., Ibnian, A.M., Alkhodari, M. (2021) Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers, *Biocybernetics and Biomedical Engineering* **41**: 1-14.

[10] Shuvo, S.B., Ali, S.N., Swapnil, S.I., Hasan, T., Bhuiyan, M.I.H. (2020) A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD-CWT-based hybrid scalogram, *IEEE Journal*

*of Biomedical and Health Informatics* **25**: 2595-2603.

[11] Hariri, W., Narin, A. (2021) Deep neural networks for COVID-19 detection and diagnosis using images and acoustic-based techniques: a recent review, *Soft Computing* **25**: 15345-15362.

[12] Mukherjee, H., Sreerama, P., Dhar, A., Obaidullah, S.M., Roy, K., Mahmud, M., Santosh, K. (2021) Automatic lung health screening using respiratory sounds, *Journal of Medical Systems* **45**: 19.

[13] Kim, Y., Hyon, Y., Jung, S.S., Lee, S., Yoo, G., Chung, C., Ha, T. (2021) Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning, *Scientific reports* **11**: 17186.

[14] Satea, H.D., Elameer, A.S., Salman, A.H., Sateaa, S.D. (2022) Employing deep learning for lung sounds classification, *International Journal of Electrical and Computer Engineering* **12**: 4345.

[15] Jothi, E.S.J., Anitha, J., Priyadharshini, J., Hemanth, D.J. (2022) Deep Learning Based Obstructive Sleep Apnea Detection for e-health Applications. *In Proceedings of the* International Conference on Electronic Governance with Emerging Technologies, *Springer*, Tampico, Mexico, 12-14 September 2022, pp. 1-11.

[16] Song, W., Han, J. (2023) Patch-level contrastive embedding learning for respiratory sound classification, *Biomedical Signal Processing and Control* **80**: 104338.

[17] Shi, J., Chen, S., Yu, B., Ren, Y., Wang, G., Xue, C. (2023) Recognition System for Diagnosing Pneumonia and Bronchitis Using Children's Breathing Sounds Based on Transfer Learning, *Intelligent Automation & Soft Computing* **37**: 3235.

[18] Ali, S.W., Asif, M., Rashid, M., Tanvir, S., Shams, S., Abid, S. (2023) Detection of crackle and wheeze in lung sound using machine learning technique for clinical decision support system, *VAWKUM Transactions on Computer Sciences* **11**: 67-78.

[19] Krishnan, G., Singh, S., Pathania, M., Gosavi, S., Abhishek, S., Parchani, A., Dhar, M. (2023) Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm, *Frontiers in Artificial Intelligence* **6**:

1227091.

[20] Dubey, R., Bodade, R., Dubey, D. (2023) Efficient classification of the adventitious sounds of the lung through a combination of SVM-LSTM-Bayesian optimization algorithm with features based on wavelet bi-phase and bi-spectrum, *Research on Biomedical Engineering* **39**: 349-363.

[21] Lal, K.N. (2023) A lung sound recognition model to diagnoses the respiratory diseases by using transfer learning, *Multimedia Tools and Applications* **82**: 36615-36631.

[22] Prabhakar, S.K., Won, D.-O. (2023) HISET: Hybrid interpretable strategies with ensemble techniques for respiratory sound classification, *Heliyon* **9**.

[23] Aptekarev, T., Sokolovsky, V., Furman, E., Kalinina, N., Furman, G. (2023) Application of deep learning for bronchial asthma diagnostics using respiratory sound recordings, *PeerJ Computer Science* **9**: e1173.

[24] Choi, Y., Lee, H. (2023) Interpretation of lung disease classification with light attention connected module, *Biomedical Signal Processing and Control* **84**: 104695.

[25] Ali, S.W., Asif, M., Zia, M.Y.I., Rashid, M., Syed, S.A., Nava, E. (2023) CDSS for early recognition of respiratory diseases based on AI techniques: a systematic review, *Wireless Personal Communications* **131**: 739-761.

[26] Dar, J.A., Srivastava, K.K., Mishra, A. (2023) Lung anomaly detection from respiratory sound database (sound signals), *Computers in Biology and Medicine* **164**: 107311.

[27] Chudasama, V., Bhikadiya, K., Mankad, S.H., Patel, A., Mistry, M.P. (2023) Voice based pathology detection from respiratory sounds using optimized classifiers, *International Journal of Computing and Digital Systems* **13**: 327-339.

[28] Zhang, P., Swaminathan, A., Uddin, A.A. (2023) Pulmonary disease detection and classification in patient respiratory audio files using long short-term memory neural networks, *Frontiers in Medicine* **10**: 1269784.

[29] Jaffery, S.A.F., Aziz, S., Khan, M.U., Naqvi, S.Z.H., Faraz, M., Usman, A. (2023) An automated system for the classification of bronchiolitis and bronchiectasis diseases using lung sound analysis. *In Proceedings of the* 2023

International Conference on Robotics and Automation in Industry (ICRAI), *IEEE*, University of Engineering and Technology (UET) , Peshawar, Pakistan, 3-5 March 2023, pp. 1-6.

[30] Yang, R., Lv, K., Huang, Y., Sun, M., Li, J., Yang, J. (2023) Respiratory sound classification by applying deep neural network with a blocking variable, *Applied Sciences* **13**: 6956.

[31] Ren, Z., Nguyen, T.T., Zahed, M.M., Nejdl, W. (2023) Self-explaining neural networks for respiratory sound classification with scale-free interpretability. *In Proceedings of the* 2023 International Joint Conference on Neural Networks (IJCNN), *IEEE*, Queensland, Australia, 18-23 June 2023, pp. 01-07.

[32] Kalaiyarasan, K., Sridhar, R. (2023) Artificial Intelligence in Respiratory Medicine: The Journey So Far–A Review, *Journal of Association of Pulmonologist of Tamil Nadu* **6**: 53-68.

[33] Fava, A., Dianat, B., Bertacchini, A., Manfredi, A., Sebastiani, M., Modena, M., Pancaldi, F. (2024) Pre-processing techniques to enhance the classification of lung sounds based on deep learning, *Biomedical Signal Processing and Control* **92**: 106009.

[34] Zhang, Y., Huang, Q., Sun, W., Chen, F., Lin, D., Chen, F. (2024) Research on lung sound classification model based on dual-channel CNN-LSTM algorithm, *Biomedical Signal Processing and Control* **94**: 106257.

[35] Trivedi, J., Degadwala, S. (2024) A Review on Identifying Lung Disease Sounds using different ML and DL Models, *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* **10**: 399-411.

[36] Nadkarni, R., Nikolakakis, E., Marinescu, R. (2024) AFEN: Respiratory disease classification using ensemble learning, *arXiv preprint* arXiv:2405.05467.

[37] Aljaddouh, B., Malathi, D., Alaswad, F. (2024) Multimodal disease detection and classification using breath sounds and vision transformer for improved diagnosis, *Procedia Computer Science* **235**: 1436-1444.

[38] Hakkı, L., Serbes, G. (2024) Detection of wheeze sounds in respiratory disorders: a deep Learning approach, *International Advanced Researches and*

*Engineering Journal* **8**: 20-32.

[39] Geetha, J., Benadict Raja, J. (2024) An advanced circular adaptive search butterfly optimization algorithm for the CNN-based sleep apnea detection approach, *IETE Journal of Research* **70**: 1425-1437.

[40] Khan, R., Khan, S.U., Saeed, U., Koo, I.-S. (2024) Auscultation-based pulmonary disease detection through parallel transformation and deep learning, *Bioengineering* **11**: 586.

[41] Sang, B., Wen, H., Junek, G., Neveu, W., Di Francesco, L., Ayazi, F. (2024) An accelerometer-based wearable patch for robust respiratory rate and wheeze detection using deep learning, *Biosensors* **14**: 118.

[42] Xu, W., Bao, X., Lou, X., Liu, X., Chen, Y., Zhao, X., Zhang, C., Pan, C., Liu, W., Liu, F. (2024) Feature fusion method for pulmonary tuberculosis patient detection based on cough sound, *Plos One* **19**: e0302651.

[43] Wu, C., Ye, N., Jiang, J. (2024) Classification and recognition of lung sounds based on improved Bi-ResNet model, *IEEE Access* **12**: 73079-73094.

[44] Wu, Y., Rocha, B.M., Kaimakamis, E., Cheimariotis, G.-A., Petmezas, G., Chatzis, E., Kilintzis, V., Stefanopoulos, L., Pessoa, D., Marques, A. (2024) A deep learning method for predicting the COVID-19 ICU patient outcome fusing X-rays, respiratory sounds, and ICU parameters, *Expert Systems with Applications* **235**: 121089.

[45] Wang, Z., Sun, Z. (2024) Performance evaluation of lung sounds classification using deep learning under variable parameters, *EURASIP Journal on Advances in Signal Processing* **2024**: 51.

[46] Mang, L.D., González Martínez, F.D., Martinez Muñoz, D., García Galán, S., Cortina, R. (2024) Classification of adventitious sounds combining cochleogram and vision transformers, *Sensors* **24**: 682.

[47] Patel, P.J., Diwan, D., Patel, K.A., Ranga, S., Modi, N.J., Dumasia, S. (2024) Multi feature fusion for COPD classification using deep learning algorithms, *Journal of Integrated Science and Technology* **12**: 780-780.

[48] Wanasinghe, T., Bandara, S., Madusanka, S., Meedeniya, D., Bandara, M., Díez, I.D.L.T. (2024) Lung sound classification with multi-feature integration utilizing lightweight CNN model, *IEEE Access* **12**: 21262-21276.

[49] Saeed, T., Ijaz, A., Sadiq, I., Qureshi, H.N., Rizwan, A., Imran, A. (2024) An AI-Enabled bias-free respiratory disease diagnosis model using cough audio, *Bioengineering* **11**: 55.

[50] Isangula, K.G., Haule, R.J. (2024) Leveraging AI and machine learning to develop and evaluate a contextualized user-friendly cough audio classifier for detecting respiratory diseases: protocol for a diagnostic study in rural Tanzania, *JMIR Research Protocols* **13**: e54388.

[51] Kim, J.-W., Lee, S., Toikkanen, M., Hwang, D., Kim, K. (2025) Tri-MTL: A Triple Multitask Learning Approach for Respiratory Disease Diagnosis. *In Proceedings of the* 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), *IEEE*, Copenhagen, Denmark, 14-17 July 2025, pp. 1-6.

[52] Álvarez Casado, C., Räsänen, P., Nguyen, L.N., Lämsä, A., Peltola, J., Bordallo López, M. (2024) A Distributed Framework for Remote Multimodal Biosignal Acquisition and Analysis. *In Proceedings of the* Nordic Conference on Digital Health and Wireless Solutions, *Springer Nature Switzerland*, Oulu, Finland, 7–8 May 2024, pp. 127-146.

[53] Bhushan, P., Fahad, M.S., Agrawal, S., Kamesh, K.S.D., Tripathi, P., Mishra, P., Singh, V.K., Deepak, A. (2024) A self-attention based hybrid cnn-lstm architecture for respiratory sound classification, *GMSARN International Journal* **18**: 54-61.

[54] Nguyen, T., Pernkopf, F. (2022) Lung sound classification using co-tuning and stochastic normalization, *IEEE Transactions on Biomedical Engineering* **69**: 2872-2882.

[55] Wall, C., Zhang, L., Yu, Y., Kumar, A., Gao, R. (2022) A deep ensemble neural network with attention mechanisms for lung abnormality classification using audio inputs, *Sensors* **22**: 5566.

[56] Moon, H.J., Ji, H., Kim, B.S., Kim, B.J., Kim, K. (2025) Machine learning-driven strategies for enhanced pediatric wheezing detection, *Frontiers in Pediatrics* **13**: 1428862.

[57] Albiges, T., Sabeur, Z., Arbab-Zavar, B. (2025) Features and eigenspectral densities analyses for machine learning and classification of severities in

chronic obstructive pulmonary diseases, *Intelligence-Based Medicine* **11**: 100217.

[58] Erlangga, M.D., Faisal, M.R., Muliadi, M., Indriani, F., Kartini, D., Satou, K. (2025) Incorporating MFCC Feature Extraction to the Classification of Respiratory Sounds by Machine Learning Algorithms. *In Proceedings of the* 2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI), *IEEE*, Jakarta, Indonesia, January 21, 2025, pp. 301-306.

[59] Bacanin, N., Jovanovic, L., Stoean, R., Stoean, C., Zivkovic, M., Antonijevic, M., Dobrojevic, M. (2024) Respiratory condition detection using audio analysis and convolutional neural networks optimized by modified metaheuristics, *Axioms* **13**: 335.

[60] Sanchez-Morillo, D., Sales-Lerida, D., Priego-Torres, B., León-Jiménez, A. (2024) Cough detection using acceleration signals and deep learning techniques, *Electronics* **13**: 2410.

Thamar University Journal of
**Natural & Applied Sciences**

**ISSN-Print: 203-0764**
**ISSN-Online: 2959-4340**

TUJNAS

***Volume 11, Issue 1, 2026***

## CONTENTS

### Articles


| Page | Title of Article | *Categories* |
|---|---|---|
| 1 | **The Impact of Specific Microorganisms Causing the Abortion of a Random Sample of Pregnant Women in Dhamar City, Yemen**<br>Mobarak Rassam *et al.* | ***Medical Science:*** *Obstetrics & Gynecology* |
| 8 | **Management of Tibial Infected Nonunion with Segmental Bone Defect by the Ilizarov Technique**<br>Hefzulla Abdulla *et al.* | ***Medicine:*** *Orthopedics & Traumatology* |
| 15 | **A Systematic Review and Meta-analysis Survey of IDS/IPS Techniques for CAN and Vehicular Networks**<br>Younis Al Shojaa & Khaled Al Soufy | ***Computer Science:*** *Cybersecurity / Vehicular Information Technology* |
| 21 | **Deep Learning for Respiratory Sound Analysis: A Systematic Review and Meta-Analysis (2019–2024)**<br>Shaima'a Al-Jabali *et al.* | ***Information Technology:*** *Artificial Intelligence in Healthcare / Biomedical Engineering* |
| 29 | **In Vitro Quality Assessment of Commercially Available Azithromycin Tablets in Dhamar City, Yemen: A Comparative Study**<br>Neaf Al-Tayar *et al.* | ***Pharmacology:*** *Analytical Chemistry & Quality Control* |
| 34 | **Preliminary Evaluation of Acute Toxicity and Hypoglycemic Effects of *Salvia officinalis* L. Extract in Healthy Male Rabbits: An Animal Model Stud**<br>Hisham Al-khawlani *et al.* | ***Pharmacology:*** *Ethnopharmacology & Metabolic Biochemistry* |
| 40 | **Immunomodulatory Effects of Carob (*Ceratonia siliqua*) Extract in Pregnant Rats: A Histological Evaluation of Embryonic Liver**<br>Zainab H. Majeed | ***Biology:*** *Immunology, Histology & Pathology* |


Published by Thamar University – Dhamar, Yemen