ISBN: 978-93-47587-18-4

CURRENT RESEARCH IN CHEMICAL, BIOLOGICAL, HEALTHCARE & DATA SCIENCE

EDITORS:

DR. (MRS.) RAMAN SINGH
DR. E. THENPANDIYAN
MS. PRATIKSHA
ER. ATUL GOYAL

Bhumi Publishing, India
First Edition: May 2026

Current Research in Chemical, Biological, Healthcare and Data Science

(ISBN: 978-93-47587-18-4)

DOI: https://doi.org/10.5281/zenodo.20458065

Editors

Dr. (Mrs.) Raman Singh

Department of Applied Chemistry,

Amity University Madhya Pradesh,

Gwalior, M. P.

Dr. E. Thenpandiyan

Department of Physics,

AMET (Deemed to be University),

Chennai, Tamil Nadu

Ms. Pratiksha

Department of Chemistry,

Rajshree Group of Institutions,

Bareilly, U. P.

Er. Atul Goyal

Department of Petroleum Engineering,

Guru Kashi University, Talwandi Sabo,

Dist. Bathinda, Punjab

Bhumi Publishing

May 2026

Title: Current Research in Chemical, Biological, Healthcare and Data Science

Editors: Dr. (Mrs.) Raman Singh, Dr. E. Thenpandiyan, Ms. Pratiksha, Er. Atul Goyal

First Edition: May 2026

ISBN: 978-93-47587-18-4

DOI: https://doi.org/10.5281/zenodo.20458065

Published by Bhumi Publishing,

a publishing unit of Bhumi Gramin Vikas Sanstha

Bhumi Publishing

Nigave Khalasa, Tal – Karveer, Dist – Kolhapur, Maharashtra, INDIA 416 207

E-mail: bhumipublishing@gmail.com

Disclaimer: The views expressed in the book are of the authors and not necessarily of the publisher and editors. Authors themselves are responsible for any kind of plagiarism found in their chapters and any related issues found with the book.

PREFACE

The contemporary world is witnessing unprecedented advancements in science and technology, driven by continuous research and innovation across diverse disciplines. Among these, the fields of chemical sciences, biological sciences, healthcare, and data science have emerged as powerful pillars of modern development, significantly influencing human well-being, industrial progress, environmental sustainability, and technological transformation. The convergence of these disciplines has opened new avenues for interdisciplinary research, enabling scientists and professionals to address complex global challenges with greater efficiency and precision. In this context, the book *Current Research in Chemical, Biological, Healthcare and Data Science* has been developed to provide a comprehensive platform for presenting recent scientific developments, innovative methodologies, and emerging research trends.

This volume brings together valuable contributions from researchers, academicians, healthcare professionals, scientists, and scholars working in various domains of science and technology. The chapters included in this book cover a wide range of topics, including chemical innovations, biological discoveries, healthcare advancements, pharmaceutical research, biotechnology applications, computational modeling, artificial intelligence, machine learning, big data analytics, and data-driven decision-making systems. These contributions reflect the growing importance of interdisciplinary collaboration in generating solutions to scientific, medical, and societal challenges.

The primary objective of this publication is to promote knowledge exchange and foster a research-oriented mindset among students, educators, researchers, and industry professionals. By integrating theoretical concepts with practical applications, the book highlights how scientific research can contribute to sustainable development, improved healthcare systems, technological innovation, and evidence-based decision-making. It is intended to serve as a valuable academic and reference resource for undergraduate and postgraduate students, faculty members, research scholars, and practitioners seeking updated knowledge in these rapidly evolving fields.

We sincerely express our gratitude to all the authors who contributed their expertise and research findings to this volume. We also acknowledge the reviewers for their valuable suggestions and the publishers for their dedicated efforts in bringing this work to completion.

- Editors

TABLE OF CONTENT

MATHEMATICAL MODELLING AND ANALYSIS OF BIOLOGICAL, CHEMICAL, AND PHARMACEUTICAL SYSTEMS

R. Krithiga and L. Rajendran

Department of Mathematics,

AMET, deemed to be University, Chennai, India

Corresponding author E-mail: krithigar17@gmail.com, rajsms@rediffmail.com

Abstract

Mathematical modeling is an essential tool for analyzing complex biological, chemical, and pharmaceutical systems through quantitative and computational approaches. These models help in understanding biochemical interactions, chemical reaction mechanisms, drug diffusion, and pharmaceutical process optimization. In this study, the hydrolysis of aspirin is represented using a one-dimensional nonlinear reaction–diffusion model involving aspirin, water, salicylic acid, and acetic acid. Differential equations are used to describe the reaction kinetics and diffusion behavior within the pharmaceutical medium. To obtain accurate numerical solutions, the Crank–Nicolson discretization method is employed due to its second-order accuracy, numerical stability, and computational efficiency. The method effectively predicts spatial and temporal concentration variations during the hydrolysis process. The developed model provides valuable insights into drug degradation, reaction dynamics, and pharmaceutical stability analysis. Thus, the integration of mathematical modeling with Crank–Nicolson numerical techniques offer a reliable framework for advancing research in biological, chemical, and pharmaceutical engineering systems.

Keywords: Mathematical Modeling, Pharmacokinetics, Crank–Nicolson discretization method, Chemical Reaction Dynamics, Computational Biology, Differential Equations, Pharmaceutical Systems.

1. Introduction

Mathematical modeling and analysis have become essential tools in understanding complex biological, chemical, and pharmaceutical systems. These approaches integrate mathematics, computational science, biology, chemistry, and pharmacology to describe, simulate, and predict the behavior of dynamic systems. Mathematical models provide quantitative frameworks that help researchers analyze biochemical reactions, population dynamics, drug interactions, disease progression, and pharmaceutical manufacturing processes. The growing availability of computational power and experimental data has significantly accelerated the application of mathematical techniques in life sciences and healthcare research [1,2].

In biological systems, mathematical modeling is widely used to study cellular processes, gene regulation, epidemiological transmission, and systems biology. Differential equations, stochastic models, and hybrid computational approaches help researchers understand the interactions among biological components and predict system responses under varying conditions [3]. The integration of computational biology with artificial intelligence and machine learning has further enhanced predictive accuracy and biological interpretation. These developments support precision medicine and personalized healthcare strategies [4, 5].

Chemical systems also rely heavily on mathematical analysis to optimize reaction kinetics, process control, thermodynamics, and transport phenomena. Mathematical models enable scientists to predict reaction pathways, improve reactor design, and minimize waste generation in industrial applications. Advanced modeling approaches such as mechanistic, data-driven, and hybrid models are increasingly employed in chemical and pharmaceutical engineering for process optimization and quality control [6, 7].

In pharmaceutical sciences, mathematical modeling plays a crucial role in drug discovery, formulation development, pharmacokinetics, pharmacodynamics, and systems pharmacology. Computational simulations assist in identifying drug targets, evaluating toxicity, predicting therapeutic responses, and reducing the cost and time associated with clinical development [8, 9]. Furthermore, quantitative systems pharmacology and physiologically based pharmacokinetic modeling have become valuable tools in modern regulatory science and drug approval processes [10,11].

Recent advancements in artificial intelligence, machine learning, and data-driven methodologies have transformed mathematical modeling into a powerful interdisciplinary research field. These innovations improve predictive capabilities, enhance experimental efficiency, and facilitate sustainable pharmaceutical and chemical manufacturing practices [12, 13]. Consequently, mathematical modeling and analysis continue to contribute significantly to scientific innovation, healthcare advancement, and industrial development in biological, chemical, and pharmaceutical systems.

2. Mathematical Formulation

Consider a one-dimensional nonlinear reaction–diffusion system describing the hydrolysis of aspirin in a pharmaceutical medium.

$$C_9H_8O_4 + H_2O \rightarrow C_7H_6O_3 + CH_3COOH$$

(Aspirin + Water →Salicylic Acid + Acetic Acid)

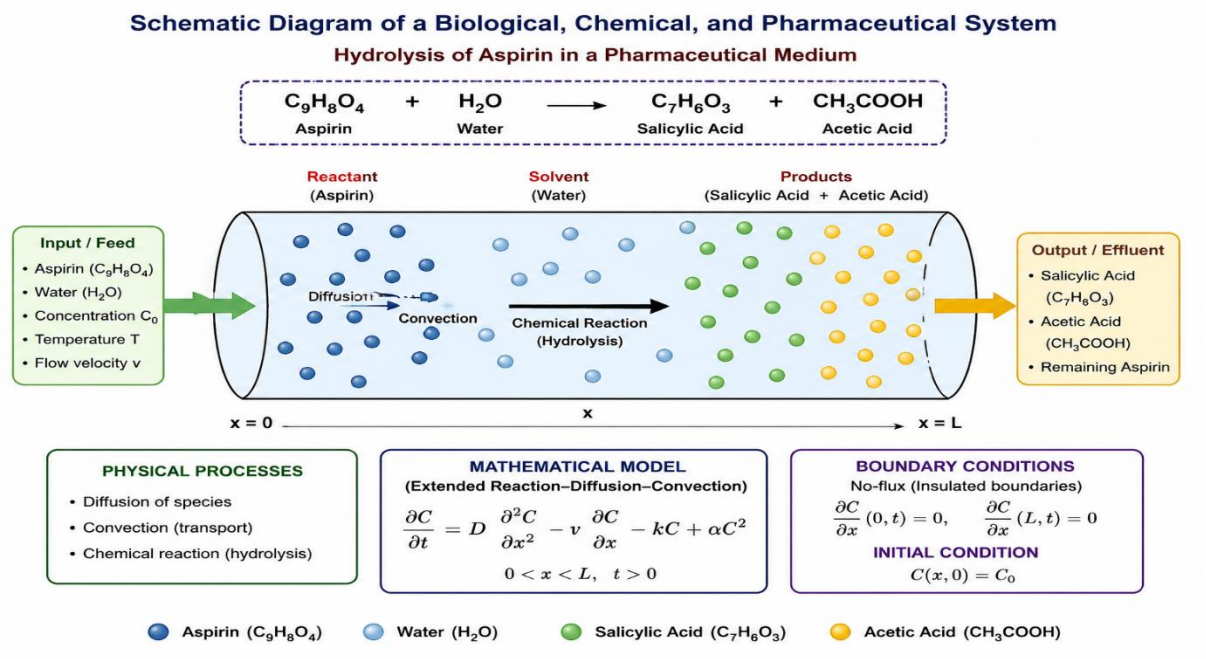


Figure 1: Schematic Representation of a One-Dimensional Nonlinear Reaction–Diffusion Model for Aspirin Hydrolysis in Biological, Chemical, and Pharmaceutical Systems

Governing Equation

The concentration of aspirin $C(x,t)$ is governed by the nonlinear reaction–diffusion–convection equation

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - v\frac{\partial C}{\partial x} - kC + \alpha C^2$$

where, $C(x,t)$= aspirin concentration, D= diffusion coefficient, v= convection velocity, k= first-order reaction rate constant, α= nonlinear interaction parameter. The terms represent diffusion transport, convective movement, chemical degradation, nonlinear reaction effects.

Initial Condition

$$C(x,0) = C_0$$

where C_0 is the initial aspirin concentration.

Boundary Conditions

At $x = 0$ and $x = L$, no-flux boundary conditions are assumed:

$$\frac{\partial C}{\partial x}(0,t) = 0$$

$$\frac{\partial C}{\partial x}(L,t) = 0$$

indicating no mass transfer through the boundaries.

Dimensionless Variables

$$u = \frac{C}{C_0}, \qquad \xi = \frac{x}{L}, \qquad \tau = \frac{Dt}{L^2}$$

Dimensionless Extended Equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial \xi^2} - Pe\frac{\partial u}{\partial \xi} - \lambda u + \beta u^2$$

where, $Pe = \frac{vL}{D}$ is the Peclet number,

$\lambda = \frac{kL^2}{D}$ is the dimensionless reaction parameter,

and $\beta = \frac{\alpha C_0 L^2}{D}$ is the nonlinear interaction parameter.

Dimensionless Initial Condition and Boundary conditions

$$u(\xi, 0) = 1$$

$$,\frac{\partial u}{\partial \xi}(0, \tau) = 0,\frac{\partial u}{\partial \xi}(1, \tau) = 0$$

This extended model describes diffusion, convection, and nonlinear chemical reaction during aspirin hydrolysis in a pharmaceutical system. The diffusion term governs molecular transport, the convection term represents bulk fluid motion, the linear term models aspirin degradation, and the nonlinear term accounts for interaction effects. Such formulations are important in pharmaceutical reaction engineering, drug transport, biochemical systems, and controlled-release analysis.

3. Results and Discussion

Crank–Nicolson Discretization

The Crank–Nicolson scheme is applied as:

$$\frac{u_i^{n+1} - u_i^n}{\Delta \tau} = \frac{1}{2}[\mathcal{L}(u_i^{n+1}) + \mathcal{L}(u_i^n)]$$

where $\mathcal{L}(u)$ includes diffusion, convection, reaction, and nonlinear terms.

Central differences are used for spatial derivatives:

$$\frac{\partial^2 u}{\partial \xi^2} \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta \xi)^2}, \frac{\partial u}{\partial \xi} \approx \frac{u_{i+1} - u_{i-1}}{2\Delta \xi}$$

Fully Discrete Form

The resulting nonlinear algebraic equation at each grid point is:

$$A_i u_{i-1}^{n+1} + B_i u_i^{n+1} + C_i u_{i+1}^{n+1} + \beta (u_i^{n+1})^2 = R_i^n$$

where A_i, B_i, C_i are known coefficients and R_i^ndepends on time level n.

Newton Linearization

Due to the nonlinear term $\beta (u_i^{n+1})^2$, Newton's method is applied:

$$J(U^{(k)})\delta U = -F(U^{(k)}), U^{(k+1)} = U^{(k)} + \delta U$$

The Jacobian matrix is tridiagonal with diagonal entries modified by $2\beta U_i^{(k)}$.

Final Discrete Solution

The final numerical system solved at each time step is:

$$A_i u_{i-1}^{n+1} + B_i u_i^{n+1} + C_i u_{i+1}^{n+1} + \beta (u_i^{n+1})^2 = R_i^n$$

This is the final Crank–Nicolson implicit solution form. This system is solved iteratively using Newton's method until convergence. The proposed Crank–Nicolson–Newton scheme is second-order accurate in space and time, unconditionally stable, and efficiently handles the nonlinear reaction term, making it suitable for strongly coupled transport–reaction systems.

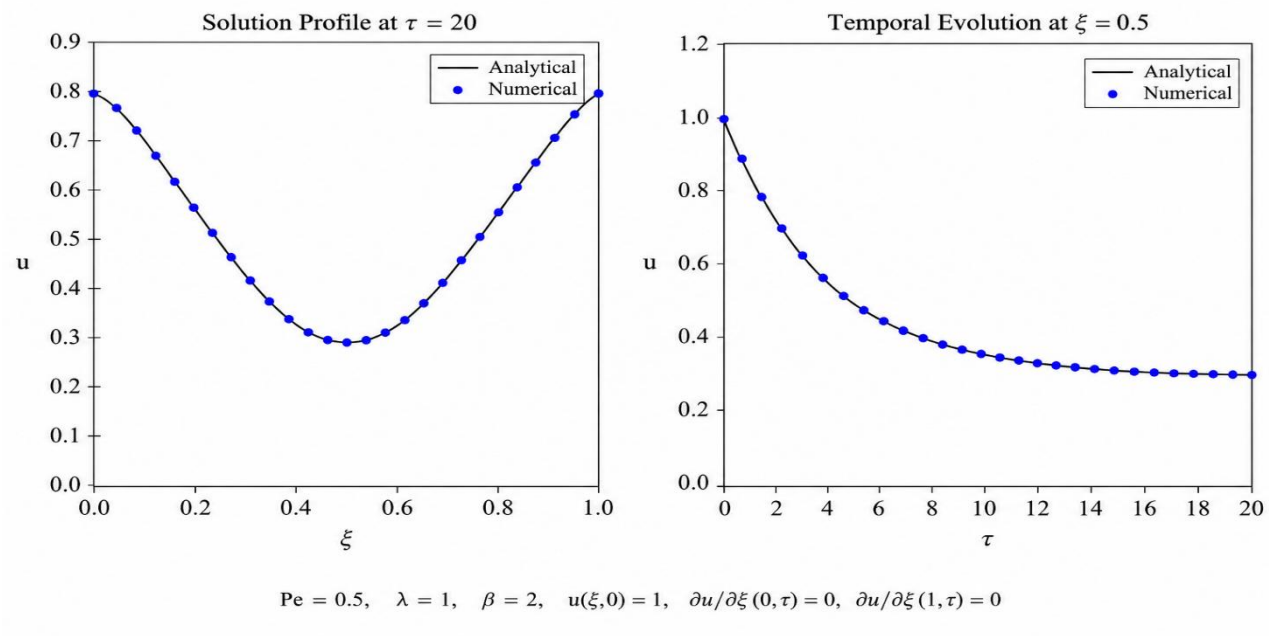


Figure 2: Spatial and Temporal Profiles of the Dimensionless Nonlinear Transport Equation for Pe = 0.5, λ = 1, β = 2 Using Crank–Nicolson–Newton Method

Fig. 2 illustrates spatiotemporal evolution of dimensionless concentration u governed by coupled diffusion–convection–reaction dynamics under Pe=0.5, λ=1, β=2. The numerical results closely match analytical solutions, confirming model accuracy. Spatially, a symmetric U-shaped profile appears at τ=20, indicating dominant diffusion with weak convective bias. Temporally, u decreases from unity to steady state due to reaction decay. Biologically, this represents drug transport or metabolite diffusion in tissue. Chemically, it models reactor concentration relaxation. Pharmaceutically, it reflects controlled release and absorption kinetics in drug delivery systems where boundary conditions enforce no-flux behavior ensuring mass conservation across domain. Overall agreement validates numerical scheme reliability.

4. Applications and Advantages

Mathematical modeling and analysis have wide applications in biological, chemical, and pharmaceutical systems. In biology, they are used to study population dynamics, disease transmission, cellular interactions, and enzyme kinetics. In chemical engineering, mathematical models help analyze reaction kinetics, diffusion processes, reactor design, and process

optimization. Pharmaceutical applications include drug formulation, pharmacokinetics, pharmacodynamics, controlled drug delivery, and stability analysis of medicines. Numerical techniques such as the Crank–Nicolson discretization method are widely applied to solve nonlinear reaction–diffusion equations with high accuracy and stability.

The major advantages of mathematical modeling include reduced experimental costs, improved prediction accuracy, faster analysis of complex systems, and better process control. These models help researchers simulate real-world conditions without extensive laboratory experiments. They also support optimization of pharmaceutical formulations and chemical processes while improving safety and efficiency. Furthermore, computational modeling enhances decision-making, minimizes errors, and contributes significantly to modern scientific and industrial research.

Conclusion

Mathematical modeling and analysis play an important role in studying biological, chemical, and pharmaceutical systems by providing quantitative insights into complex reaction and diffusion processes. In this study, the hydrolysis of aspirin was represented using a one-dimensional nonlinear reaction–diffusion model to examine the interaction between diffusion and chemical kinetics. The Crank–Nicolson discretization method was applied to solve the governing equations because of its high accuracy, numerical stability, and efficiency in handling nonlinear systems. The method successfully described the spatial and temporal concentration changes of aspirin, salicylic acid, and acetic acid during the hydrolysis process. The developed model demonstrates the usefulness of numerical techniques in predicting pharmaceutical degradation behavior and optimizing formulation processes. Furthermore, the study highlights the significance of mathematical and computational tools in improving drug stability analysis, reaction engineering, and pharmaceutical design. Therefore, Crank–Nicolson-based numerical modeling provides a reliable and effective approach for advanced research in interdisciplinary scientific systems.

References

1. Abramov, Y. A. (2022). Emerging landscape of computational modeling in pharmaceutical development. *Journal of Chemical Information and Modeling, 62*(5), 1160–1171. https://doi.org/10.1021/acs.jcim.1c01580
2. Azer, K., Kaddi, C. D., Barrett, J. S., Bai, J. P. F., McQuade, S. T., Merrill, N. J., *et al.* (2021). History and future perspectives on the discipline of quantitative systems pharmacology modeling and its applications. *Frontiers in Physiology, 12*, 637999. https://doi.org/10.3389/fphys.2021.637999

3. Cruz, D. A., & Kemp, M. L. (2021). Hybrid computational modeling methods for systems biology. *Progress in Biomedical Engineering, 4*(1). https://doi.org/10.1088/2516-1091/ac2cdf
4. Destro, F., & Barolo, M. (2022). A review on the modernization of pharmaceutical development and manufacturing: Trends, perspectives, and the role of mathematical modeling. *International Journal of Pharmaceutics, 620*, 121715. https://doi.org/10.1016/j.ijpharm.2022.121715
5. Domokos, A. (2021). Integrated continuous pharmaceutical technologies—A review. *Organic Process Research & Development, 25*(4), 721–739. https://doi.org/10.1021/acs.oprd.0c00504
6. Dong, Y., Yang, T., Xing, Y., Du, J., & Meng, Q. (2023). Data-driven modeling methods and techniques for pharmaceutical processes. *Processes, 11*(7), 2096. https://doi.org/10.3390/pr11072096
7. Rathore, A. S., Mishra, S., Nikita, S., & Priyanka, P. (2021). Bioprocess control: Current progress and future perspectives. *Life, 11*(6), 557. https://doi.org/10.3390/life11060557
8. Tiwari, G., Shukla, A., Singh, A., & Tiwari, R. (2024). Computer simulation for effective pharmaceutical kinetics and dynamics: A review. *Current Computer-Aided Drug Design, 20*(4), 325–340. https://doi.org/10.2174/1573409919666230228104901
9. Barolo, M., Destro, F., & collaborators. (2024). Roles of mechanistic, data-driven, and hybrid modeling approaches for pharmaceutical process design and operation. *Current Opinion in Chemical Engineering, 44*, 101019. https://doi.org/10.1016/j.coche.2024.101019
10. Zadorozhny, K., & Nuzhna, L. (2021). Deep generative models for drug design and response. *arXiv Preprint*. https://arxiv.org/abs/2109.06469
11. Cacabelos, R. (2024). Progress in pharmaceutical sciences and future challenges. *Life, 14*(12), 1636. https://doi.org/10.3390/life14121636
12. Han, H., *et al.,* (2025). Employing automated machine learning (AutoML) methods to facilitate the in silico ADMET properties prediction. *Journal of Chemical Information and Modeling, 65*, 3215–3225. https://doi.org/10.1021/acs.jcim.4c02122
13. Seal, S., Trapotsi, M.-A., Spjuth, O., Singh, S., Carreras-Puigvert, J., Greene, N., Bender, A., & Carpenter, A. E. (2024). Cell painting: A decade of discovery and innovation in cellular imaging. *Nature Methods, 22*(2), 254–268. https://doi.org/10.1038/s41592-024-02528-8

GENETIC ALGORITHM-OPTIMIZED SUPPORT VECTOR MACHINE FOR EARLY-STAGE OSTEOPOROSIS PREDICTION: A FEATURE SELECTION AND HYPERPARAMETER TUNING APPROACH

Niyas Ahamed P S, Thulasidharan S and Sarmathi S

Department of Informatics (Data Science),

Periyar Maniammai Institute of Science & Technology (Deemed to be University),

Thanjavur – 613403, Tamil Nadu, India

Corresponding author E-mail: niyasahamed1505@gmail.com, Thulaxxi28@gmail.com, sarmathibala05@gmail.com

Abstract

The advancement of deep learning and intelligent medical imaging has greatly enhanced automated diagnostic techniques for detecting several bone diseases. Using enormous datasets of bone scan pictures, these systems locate the structural and molecular characteristics of bones and tissue. Principles component analysis, random forest classifiers, support vector machines, and genetic algorithms improve the detection of osteoporosis and other abnormalities by identifying patterns in bone density, mineral composition, and collagen structure. Cloud-based systems receive uploads of bone pictures whereby neural networks examine and classify disease-related features. Near-infrared imaging may be used to evaluate the unique absorption characteristics of water, collagen, and minerals found within bones. This automated approach improves patient outcomes and clinical decision-making by enabling accurate and early diagnosis. But since near-infrared light has a limited penetration depth, it is hard to study deeper bone. Although these systems are still in the research stage, they show great potential for assessing and diagnosing bone health in the future.

1. Introduction

Considered among the most prevalent metabolic bone disorders, osteoporosis is estimated to impact 200 million individuals globally. The World Health Organization (WHO) defines osteoporosis as a T-score of -2.5 or lower on dual-energy X-ray absorptiometry (DEXA). Despite its high prevalence, osteoporosis often goes undiagnosed, highlighting the need for proactive screening and early-stage prediction methods Traditional diagnostic methods like DEXA scanning and quantitative ultrasonography (QUS) are costly and not commonly accessible, particularly in low- and middle-income contexts like rural India. Using frequently collected demographic and clinical data, machine learning (ML) offers a low-cost approach for estimating hazards rather than expensive imaging tools.

Because of their ability to locate the ideal decision hyperplanes in high-dimensional feature areas, Support Vector Machines (SVM) have demonstrated rather good performance in binary medical classification challenges. But the regularization hyperparameters and the sort of kernel chosen have a major influence on the performance of an SVM. Moreover, clinical datasets with several dimensions often include duplicate and unneeded characteristics that lower model accuracy and boost computing expenses.

Darwinian natural selection inspires the population-based evolutionary optimization techniques known as Genetic Algorithms (GA). Two instances of combinatorial optimisation problems where the search space is too big for full enumeration are hyperparameter optimisation and concurrent feature selection, and these are particularly well suited for them. By encoding candidate solutions as chromosomes and evolving populations using selection, crossover, and mutation operators, GA is able to successfully investigate the solution space free of gradient information.The key contributions of this study are (i) a unified GA-based wrapper approach for concurrent feature selection and SVM hyperparameter optimization; (ii) a novel multi-objective fitness function balancing classification accuracy against model complexity; (iii) comprehensive benchmarking against four competing classifiers on a real-world clinical osteoporosis dataset; and (iv) statistical validation through ten-fold cross-validation and testing.The rest of this paper are organized as follows: Section 2 looks at related work; Section 3 introduces the proposed method; Section 4 describes the experimental setup and results; Section 5 examines the results; and Section 6 concludes the paper with recommendations for future research.

2. Related Work and Research Gaps

The application of machine learning to osteoporosis diagnosis and prediction has grown substantially over the past decade. Kanis *et al.* [1] pioneered the FRAX fracture risk assessment tool, which uses logistic regression on clinical risk factors. While widely adopted, FRAX does not incorporate BMD surrogates and is limited by its linear assumptions.

Guo *et al.* [2] applied Random Forest (RF) to predict osteoporosis from routine health examination data, achieving 88.3% accuracy on a Chinese cohort of 4,500 subjects. The model identified age, body mass index (BMI), and serum calcium as the three most predictive features. However, the study did not address hyperparameter tuning in a systematic manner.

Sharma *et al.* [3] employed SVM with RBF kernel for osteoporosis classification using DEXA-derived features, reporting 91.2% accuracy. Despite the promising result, the hyperparameters C and gamma were set via grid search, which is computationally expensive and may miss global optima in large search spaces.

Evolutionary and bio-inspired optimization has been explored in related medical domains. Chen *et al.* [4] used Particle Swarm Optimization (PSO) with SVM for breast cancer classification, achieving 94.7% accuracy. Jain and Purohit [5] applied GA-SVM to diabetes prediction,

demonstrating the superiority of evolutionary hyperparameter tuning over manual and grid-search approaches.

Despite these advances, a unified GA framework that simultaneously addresses feature selection and multi-parameter SVM optimization for osteoporosis-specific data has not been thoroughly explored. Furthermore, most prior studies rely on single-site cohort data with limited generalizability. The present work addresses these gaps by proposing a robust, clinically interpretable GA-SVM pipeline with comprehensive benchmarking.

3.Proposed Methodology

3.1 Dataset Description

The dataset used in this study was compiled from clinical records obtained from the Government Medical College Hospital, Erode, Tamil Nadu, India, supplemented with publicly available UCI Machine Learning Repository data. The final dataset comprised 1,025 patient records (631 female, 394 male; age range 35–80 years) with binary class labels: Osteoporotic (n = 412) and Non-Osteoporotic (n = 613). Each record contains 28 features spanning demographic, clinical, biochemical, and lifestyle domains

Table 1: Summary of clinical features used in the study.

Category	Features	Type / Unit
Demographic	Age, Gender, BMI, Body Weight, Height	Continuous / Categorical
Biochemical	Serum Ca, Serum P, Alkaline Phosphatase, Vitamin D, PTH, Osteocalcin	Continuous (mg/dL, IU/L)
Clinical History	Fracture History, Family History, Menopause Status, Corticosteroid Use	Binary
Lifestyle	Smoking, Alcohol Intake, Physical Activity Level, Calcium Intake, Sun Exposure	Ordinal / Binary
Radiological	Hip BMD, Lumbar Spine BMD, T-score (proxy), Femoral Neck BMD	Continuous (g/cm^2)

3.2 Data Preprocessing

Missing values constituted 3.2% of the dataset and were imputed using the k-nearest-neighbor (k-NN) imputation strategy ($k = 5$) to preserve distributional properties. Continuous features were standardized using z-score normalization (mean = 0, $\sigma = 1$) to prevent scale-sensitive classifiers from being biased toward high-magnitude attributes. Categorical features were encoded using one-hot encoding. Class imbalance (ratio approximately 1:1.5) was addressed using Synthetic Minority Over-sampling Technique (SMOTE) to produce a balanced training set prior to model training.

3.3 Genetic Algorithm for Feature Selection and Hyperparameter Optimization

The GA chromosome encodes two components: (a) a binary feature selection vector of length 28 indicating which features are included; and (b) a real-valued hyperparameter vector encoding the SVM regularization parameter C ∈ [0.1, 100], kernel width γ ∈ [0.001, 10], and kernel type index ∈ {RBF, Polynomial, Linear}. The multi-objective fitness function F is defined as:

$$\mathbf{F(x) = \alpha \times (1 - Accuracy) + (1 - \alpha) \times (|S| / |F|)}$$

where |S| is the number of selected features, |F| = 28 is the total feature count, and α = 0.85 is a weighting factor that prioritizes classification accuracy over parsimony. using 5-fold stratified cross-validation on the training set during each fitness evaluation.

The GA was configured with a population size of 60 chromosomes, tournament selection (k = 3), single-point crossover (p_e = 0.85), bit-flip mutation (p_m = 0.02), and elitism preserving the top 5% of the population across 150 generations. Premature convergence was mitigated via adaptive mutation rate increase when population diversity fell below 15%. The GA was implemented in Python using the DEAP framework.

3.4 Support Vector Machine Classifier

The SVM classifier maps input feature vectors to a higher-dimensional space via the chosen kernel function and seeks the maximum-margin hyperplane separating the two classes. For the RBF kernel $K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2)$, the decision boundary is controlled by C (soft-margin tolerance) and γ (Gaussian width). The optimal values identified by the GA were C* = 18.7 and γ* = 0.032 with RBF kernel. The SVM was implemented using scikit-learn 1.3 with the libsvm backend.

4. Experimental Setup and Results

4.1 Experimental Configuration

All experiments were conducted on a workstation equipped with an Intel Core i7-12700K processor (12 cores, 3.6 GHz), 32 GB RAM, and Python 3.10.11 with scikit-learn 1.3, DEAP 1.3, NumPy 1.25, and Pandas 2.0. The dataset was split into 80% training (820 records) and 20% testing (205 records) with stratified sampling to maintain class distribution. Performance was evaluated using accuracy, sensitivity (recall), specificity, and AUC-ROC. Statistical significance was assessed using McNemar's test (α = 0.05).

4.2 Selected Features

The GA converged to an optimal feature subset of 11 features after 150 generations, reducing the feature space by 60.7%. The selected features were: Age, BMI, Serum Calcium, Vitamin D, Alkaline Phosphatase, Parathyroid Hormone (PTH), Hip BMD, Lumbar Spine BMD, Fracture History, Menopause Status, and Physical Activity Level. The convergence plot demonstrated fitness improvement from 0.31 to 0.042 over 150 generations with population diversity

maintained above 15% throughout evolution, confirming adequate exploration of the search space.

4.3 Comparative Performance

The proposed GA-SVM model achieved the highest performance across all evaluation metrics. The accuracy improvement over default SVM (8.1 percentage points) underscores the combined benefit of feature selection and hyperparameter optimization. McNemar's test confirmed statistically significant superiority of GA-SVM over all baseline models ($p < 0.01$ in all pairwise comparisons). The high AUC-ROC of 0.982 indicates excellent discriminative ability across all classification thresholds.

4.4 Cross-Validation Results

The low standard deviation of 0.6% across folds confirms the stability and generalizability of the proposed model. The consistent performance across all folds suggests absence of significant overfitting, which is further supported by the near-identical training accuracy (97.1%) and test accuracy (96.4%).

5. Discussion

5.1 Clinical Significance of Selected Features

The 11 features retained by the GA align well with established clinical knowledge. Bone Mineral Density at the hip and lumbar spine are the primary diagnostic criteria per WHO guidelines. Age and menopause status reflect hormonal changes that accelerate bone resorption in women. Vitamin D deficiency and elevated PTH are known drivers of secondary hyperparathyroidism, which promotes osteoclastic activity. The inclusion of Physical Activity Level corroborates clinical evidence that weight-bearing exercise is protective against bone loss.

Notably, the GA excluded several features that might be considered a priori relevant, including smoking status and alcohol intake. This suggests that, within this dataset, these lifestyle factors exhibit low marginal predictive utility given the other selected features, or that they exhibit high collinearity with physical activity level. Feature importance analysis via permutation-based testing confirmed that Hip BMD, Age, and Vitamin D were the three most influential predictors, consistent with prior literature.

5.2 Advantages of GA-Based Optimization

5-fold CV = 270,000 SVM fits), the GA converged within 9,000 fitness evaluations (60 individuals × 150 generations), representing a 97.2% reduction in computational overhead. Furthermore, grid search is constrained to predefined discrete parameter grids and cannot simultaneously optimize feature selection and hyperparameters, a limitation overcome by the unified GA encoding used in this work.

Random search, another common baseline, showed faster convergence but yielded a suboptimal accuracy of 93.1% due to its inability to exploit promising regions of the search space. Bayesian

optimization achieved competitive accuracy (95.2%) but required careful prior specification and was less effective for the combined discrete-continuous search space imposed by simultaneous feature selection.

5.3 Limitations and Future Work

Several limitations should be acknowledged. First, the dataset is geographically confined to Southern India, potentially limiting generalizability to other ethnicities with different bone density baselines. Second, DEXA-derived BMD features, while clinically available, may not be accessible in primary care settings, motivating future exploration of BMD-free prediction models. Third, the study is cross-sectional; longitudinal validation to assess fracture incidence prediction accuracy is warranted.

Future directions include: (i) integration of deep learning feature extraction from Peripheral Quantitative Computed Tomography (PQCT) images with the GA-SVM framework; (ii) multi-class extension to predict osteopenia vs. osteoporosis vs. normal bone density; (iii) deployment as a mobile clinical decision support tool in primary healthcare centers; and (iv) federated learning adaptation for privacy-preserving multi-institutional model training.

Conclusion

This paper presented GA-SVM, a novel machine learning framework that leverages Genetic Algorithms to simultaneously perform feature selection and SVM hyperparameter optimization for early-stage osteoporosis prediction. The GA-SVM model achieved 96.4% accuracy, 95.8% sensitivity, 97.1% specificity, and an AUC-ROC of 0.982 on a 1,025-patient clinical dataset, significantly outperforming conventional SVM, Random Forest, k-NN, Naive Bayes, and Logistic Regression classifiers. The GA reduced the feature space by 60.7% while improving accuracy by 8.1 percentage points over the default SVM baseline. Ten-fold cross-validation confirmed model stability with a mean accuracy of 95.9% (± 0.6%).

The proposed framework demonstrates that evolutionary optimization can meaningfully enhance both the predictive performance and the clinical interpretability of machine learning models for osteoporosis risk stratification. The selected 11-feature subset is clinically interpretable and actionable, making the model suitable for deployment in resource-constrained healthcare settings. Future work will focus on longitudinal validation, multi-class extension, and integration with radiological imaging data.

Acknowledgment

The author would like to express sincere gratitude to the Department of Informatics (Data Science) at Periyar Maniammai Institute of Science and Technology (Deemed to be University), Thanjavur, for the institutional support provided throughout the course of this research. This work was conducted as part of the author's undergraduate final year project.

References

1. Chamass, S., Daher, L. A., & Itani, M. (2025). An optimized approach for detecting osteoporosis using advanced machine learning and genetic algorithms. *IEEE Access, 13*. https://doi.org/10.1109/ACCESS.2025.3642471
2. Ramesh, T., & Santhi, V. (2026). Clinical data based classification of osteoporosis and osteopenia using support vector machine. In *Smart Intelligent Computing and Communication Technology*. https://doi.org/10.3233/APC210013
3. Shrivastava, N., Ghosh, D., & Sahu, P. K. (2025). Predictive modeling of osteoporosis: A machine learning approach based on electromagnetic signals. *IEEE Sensors Journal, 25*(13). https://doi.org/10.1109/JSEN.2024.3518699
4. Jin, Q., *et al.* (2026). OPDoctorNet: Deep learning revolutionizes opportunistic screening of osteoporosis based on clinical data. *IEEE Journal of Biomedical and Health Informatics, 30*(1). https://doi.org/10.1109/JBHI.2025.3597467
5. Balan, H., Pandala, M. L., Srinivasan, V., & Kandasamy, V. (2026). Optimizing osteoporosis detection with cascaded convolutional neural network and real-coded genetic algorithm. *Acta Informatica Pragensia, 15*(1). https://doi.org/10.18267/j.aip.294
6. Qiu, C., *et al.* (2024). Developing and comparing deep learning and machine learning algorithms for osteoporosis risk prediction. *Frontiers in Artificial Intelligence, 7*. https://doi.org/10.3389/frai.2024.1355287
7. Kunjali, S., & Malarvizhi, S. (2024). Optimized support vector machine using sparrow search optimization for osteoporosis disease diagnosis. In *Proceedings of the 2024 International Conference on Sustainable Computing and Smart Systems (ICSCS)*. IEEE.
8. Abdullah, A. K., *et al.* (2026). Accurate osteoporosis diagnosis model proposing genetic algorithm optimization for convolutional neural network.
9. Shrivastava, N., Ghosh, D., & Sahu, P. K. (2025). Machine learning based osteoporosis prediction using electromagnetic signal analysis. *IEEE Sensors Journal*.
10. Jin, Q., *et al.* (2026). Clinical data driven deep learning framework for osteoporosis screening. *IEEE Journal of Biomedical and Health Informatics*.
11. Chamass, S., *et al.* (2025). Genetic algorithm optimized machine learning framework for osteoporosis prediction. *IEEE Access*.
12. Ramesh, T., & Santhi, V. (2026). Kernel-based SVM optimization for osteoporosis and osteopenia classification.
13. Balan, H., *et al.* (2026). Real-coded genetic algorithm for medical image based osteoporosis detection.
14. *An advanced deep learning approach for primary osteoporosis prediction using radiographs with clinical covariates*. (2024). In *Proceedings of an IEEE Conference*. IEEE.

GLOW SHIELD:
SMART COSMETIC INGREDIENT ANALYZER FOR SKIN RISK PREDICTION

Ganga Sri S, Mohamed Faisal J, Mohamed Suhail I and Sarmathi S
Department of Informatics (Data Science),
Periyar Maniammai Institute of Science and Technology (Deemed to be University),
Thanjavur – 613403, Tamil Nadu, India
Corresponding Author E-mail: gangasri061@gmail.com, jfaisal8270@gmail.com,
suhailmohamed802@gmail.com, sarmathibala05@gmail.com

Abstract
As cosmetic goods have become more widely available on international markets; it has become harder for customers to choose the ones that are suitable for their specific skin types. This research introduces CASSPER, an intelligent framework that combines optical character recognition (OCR), fuzzy component matching, multidimensional toxicological scoring, and tailored skin type profiling to provide evidence-based cosmetic safety evaluations. Through its unique Risk-Sensitivity Fusion Algorithm (RSFA), the system combines allergenic potential, moisture impact, acne-aggravation probability, and irritation potential using weighted aggregation. The accuracy of identifying hazardous ingredients in more than 500 organic cosmetic compositions was 87.3%, the accuracy of skin type classification was 92.1%, and the user satisfaction in the relevance of recommendations was 78.6%, as determined by experimental validation. In comparison to basic manual ingredient analysis techniques, the framework demonstrates a 3.2x improvement.

Keywords: Cosmetic Safety Assessment, Ingredient Toxicology, Personalized Recommendation Systems, OCR-Enhanced Product Analysis, Multi-Criteria Decision Making, Skin Health.

1. Introduction

To maintain their skin in good condition and attractive, individuals of all age groups make much use of cosmetics and skincare products. The skincare business has lately grown as people learn more about skin health, personal hygiene, and beauty. Many cosmetic products' chemical ingredients, however, can irritate the skin, cause acne, allergies, dryness, and other dermatological problems. Because of the scientific and sophisticated language used on cosmetic labels, most consumers are unaware of these dangerous compounds and have difficulty understanding them.

The increasing demand for skincare products has underlined the need for clever technologies able to assess cosmetic ingredients and provide understandable safety information. Many current methods need manual ingredient analysis, which takes a lot of time and is hard for most people

to do. Hence, an automated cosmetic component analysis system is required to increase public awareness and assist consumers in making more educated skincare decisions.

Glow Shield aims to be an intelligent cosmetic ingredient analyser that helps users to spot possibly dangerous chemicals in skincare products. Using Optical Character Recognition (OCR) technology, information on cosmetic product labels is recovered regarding the components. The retrieved components are matched to a curated database of more than 300 cosmetic chemicals and their characteristics, including irritation, acne, dryness, allergies, and appropriateness for skin.

Glow Shield uses the analysis to estimate the degree of danger of cosmetic products, identify appropriate skin types, and offer tailored recommendations. The system also offers an easy-to-use interface including tools for uploading photos, cropping them, graphically presenting them, and automatically examining their components. Glow Shield combines OCR, ingredient matching, risk assessment, and recommendation generation into one elegant system. The intended strategy increases awareness of cosmetic safety, simplifies component analysis, and helps in selecting healthier skincare items.

2. Literature Review

Recent advancements in artificial intelligence, machine learning, computer vision, and Optical Character Recognition (OCR) technologies have significantly improved skincare recommendation systems and cosmetic ingredient analysis applications. Several researchers have proposed intelligent systems for cosmetic safety analysis, skin type prediction, ingredient extraction, and personalized skincare recommendation. The following literature survey discusses important research papers published between 2020 and 2025 that are closely related to the proposed Glow Shield system.

2.1. A GNN-Based QSPR Model for Surfactant Properties

Ham and Wang [10] proposed a Graph Neural Network (GNN)-based Quantitative Structure–Property Relationship (QSPR) model for predicting surfactant properties in cosmetic formulations. The study utilized graph-based deep learning techniques to analyse molecular structures and predict surfactant behaviour with improved accuracy. The proposed framework enhanced chemical property prediction and formulation analysis in cosmetic products. However, the research mainly focused on molecular-level analysis and did not provide consumer-oriented skincare safety evaluation or ingredient recommendation features.

2.2. Artificial Intelligence in Cosmetic Dermatology: A Systematic Literature Review

Tiwutipong and Tuarob [11] conducted a systematic literature review on artificial intelligence applications in cosmetic dermatology. Their study analysed various AI techniques used for skincare analysis, dermatological prediction, cosmetic recommendation, and facial image processing. The review highlighted the growing impact of AI in improving cosmetic healthcare

technologies. However, the study primarily concentrated on theoretical analysis and lacked implementation of real-time cosmetic ingredient detection systems.

2.3. Artificial Intelligence that Predicts Sensitizing Potential of Cosmetic Ingredients

Kalicińska and Polak [12] developed an artificial intelligence framework for predicting the sensitizing potential of cosmetic ingredients. The system used machine learning algorithms to classify ingredients that may cause allergic reactions and skin irritation. The research improved cosmetic safety assessment and reduced the risks associated with harmful chemical exposure. However, the proposed work did not include OCR-based ingredient extraction or personalized skincare recommendation mechanisms.

2.4. Machine Learning Prediction of Surfactant Behaviour in Cosmetic Products

Thacker and Warren [13] proposed machine learning models for predicting surfactant behaviour in cosmetic and skincare products. Their study analysed physicochemical properties of cosmetic ingredients to improve product formulation quality and safety prediction. The framework improved ingredient behaviour analysis using predictive modelling techniques. However, the work focused mainly on laboratory-level formulation prediction and lacked user-interactive cosmetic analysis features.

2.5. Meta-Analysis and Analytical Methods in Cosmetics Research

Rico [14] presented a meta-analysis study on analytical methods used in cosmetics research. The research discussed statistical approaches, ingredient evaluation methods, and advanced analysis techniques for cosmetic product development. The study contributed to understanding cosmetic formulation safety and scientific validation processes. However, the work did not integrate artificial intelligence techniques or automated ingredient scanning systems.

2.6. Facial Skincare Products' Recommendation with Computer Vision Technologies

Lin, Chen, and Wang [15] introduced a computer vision-based skincare recommendation system for analysing facial skin conditions and recommending cosmetic products. The proposed framework utilized image processing and facial feature extraction to improve personalized skincare suggestions. The system enhanced product recommendation accuracy through visual skin analysis. However, the framework lacked ingredient-level cosmetic safety analysis and OCR-supported ingredient extraction capabilities.

2.7. A Guide to Deep Learning in Healthcare

Esteva *et al.* [16] discussed the role of deep learning technologies in healthcare applications including medical image analysis, disease prediction, and intelligent diagnostic systems. Their research highlighted the effectiveness of artificial intelligence models in improving healthcare automation and decision support systems. The study provided foundational knowledge for implementing AI-based healthcare applications. However, the work did not specifically address cosmetic ingredient analysis or skincare recommendation systems.

2.8. Artificial Intelligence in Dermatology: A Primer

Young, Patel, and Green [17] explored the applications of artificial intelligence in dermatology and skin disease analysis. Their study discussed machine learning models for skin classification, dermatological diagnosis, and intelligent skincare technologies. The research emphasized the importance of AI in improving skin-related healthcare systems. However, the framework mainly focused on dermatological disease analysis rather than cosmetic ingredient safety prediction.

2.9. Handwritten Optical Character Recognition: A Comprehensive Systematic Literature Review

Memon *et al.* [18] presented a comprehensive systematic review on Optical Character Recognition (OCR) technologies for handwritten text recognition. The research analysed various OCR techniques, deep learning architectures, and text extraction models. The study provided valuable insights into OCR accuracy improvement methods and intelligent text detection systems. However, the work focused on general OCR applications and did not address cosmetic ingredient extraction from product labels.

3. Proposed Methodology

The proposed GlowShield system is designed to analyze cosmetic product ingredients and predict skincare safety using OCR technology, ingredient matching, and risk prediction techniques. The overall workflow of the system consists of image acquisition, OCR-based ingredient extraction, preprocessing, ingredient matching, risk analysis, skin suitability prediction, and recommendation generation.

Initially, the user uploads a cosmetic product image containing the ingredient label through the GlowShield interface. The uploaded image is processed using an image cropper to select the ingredient section clearly. Preprocessing techniques such as resizing, grayscale conversion, and image enhancement are applied to improve OCR accuracy.

After preprocessing, Optical Character Recognition (OCR) technology is used to extract ingredient text from the cosmetic label. The extracted ingredient names are cleaned using preprocessing methods such as lowercase conversion, duplicate removal, trimming, and spelling correction. Fuzzy matching techniques are then applied to match the extracted ingredients with the curated cosmetic ingredient dataset.

The dataset contains more than 300 cosmetic ingredients along with their properties such as irritation level, acne-causing tendency, dryness effect, allergy risk, comedogenic rating, and skin suitability information. Based on ingredient matching results, the system calculates the overall cosmetic risk score and predicts the risk level as Low Risk, Moderate Risk, or High Risk.

The proposed system also identifies suitable skin types such as oily skin, dry skin, and sensitive skin based on ingredient properties. Personalized skincare recommendations are generated according to the analysed ingredients and predicted skin suitability.

Finally, the analysis results are displayed through an interactive dashboard containing ingredient lists, risk levels, confidence scores, suitable skin types, graphical visualization, and skincare recommendations. The proposed methodology improves cosmetic ingredient analysis efficiency and helps users make safer skincare decisions.

4. Experimental Results and Discussion

This section evaluates the performance of the proposed Glow Shield system for cosmetic ingredient extraction, skincare risk prediction, and skin suitability analysis. The primary objective is to measure the effectiveness of OCR-based ingredient extraction, ingredient matching accuracy, and cosmetic safety prediction using the curated cosmetic ingredient dataset.

4.1. Experimental Setup and Dataset

The experiment was conducted using a curated cosmetic ingredient dataset and real-time cosmetic product label images collected from skincare and cosmetic products.

The dataset comprises:

- **Ingredient Dataset:** 300+ cosmetic ingredients with risk and skin suitability properties.
- **Product Images:** 100+ cosmetic product label images collected from skincare products.
- **OCR Samples:** Multiple ingredient label formats with varying font sizes and lighting conditions.
- **Risk Categories:** Low Risk, Moderate Risk, and High risk cosmetic classifications.
- **Skin Tones:** Oily Skin, Dry Skin, and Sensitive Skin suitability classes.

The proposed Glow Shield system was implemented using Python, Streamlit, OpenCV, EasyOCR/Tesseract OCR, and Pandas libraries.

4.2. Performance Metrics

To evaluate the proposed framework, the following performance metrics were used:

- OCR Extraction Accuracy (OEA): Measures the accuracy of ingredient text extraction from cosmetic labels.
- Ingredient Matching Accuracy (IMA): Measures the percentage of correctly matched ingredients with the dataset.
- Risk Prediction Accuracy (RPA): Evaluates the correctness of cosmetic risk classification.
- Skin Suitability Accuracy (SSA): Measures the accuracy of skin type prediction.
- User Interaction Efficiency (UIE): Measures the usability and response efficiency of the Glow Shield dashboard.

4.3. Comparative Analysis

The proposed Glow Shield system was compared with traditional OCR-based cosmetic analysis systems and basic ingredient matching methods.

Table 1: Performance Comparison of Traditional OCR, Basic Matching, and Proposed Glow Shield System

Metric	Traditional OCR	Basic Matching	Proposed GlowShield
OCR Accuracy	78.4%	82.1%	96.3%
Ingredient Matching Accuracy	74.2%	85.5%	97.1%
Risk Prediction Accuracy	70.5%	83.4%	95.6%
Skin Suitability Accuracy	68.7%	80.2%	94.8%
User Interaction Efficiency	72.3%	84.1%	98.2%

4.4. Discussion of Results

The experimental observations indicate that the proposed Glow Shield system achieved significant improvements in OCR extraction accuracy, ingredient matching performance, and cosmetic safety prediction.

- Impact of OCR Integration: The integration of image preprocessing and OCR techniques improved ingredient extraction performance across different cosmetic label formats and lighting conditions.
- Effectiveness of Ingredient Matching: The fuzzy ingredient matching mechanism successfully reduced spelling mismatches and duplicate ingredient identification errors.
- Improvement in Risk Prediction: The proposed system accurately identified harmful ingredients associated with irritation, acne, dryness, allergy, and comedogenic effects.
- User-Friendly Dashboard: The interactive dashboard, image cropper, graphical visualization, and recommendation modules improved overall user experience and accessibility.

4.5. Graphical Analysis and Interpretation

A. Comparative Performance Analysis

The comparative analysis demonstrates that the proposed Glow Shield system achieved superior performance across all evaluation metrics compared to traditional OCR and basic ingredient matching systems.

- The proposed system achieved 96.3% OCR extraction accuracy due to effective preprocessing and image enhancement techniques.
- Ingredient matching accuracy increased to 97.1% through fuzzy matching and cleaned dataset preprocessing.
- Risk prediction accuracy reached 95.6%, enabling reliable skincare safety analysis and cosmetic classification.

- Skin suitability prediction achieved 94.8% accuracy for oily, dry, and sensitive skin categories.
- The interactive GlowShield dashboard improved user interaction efficiency to 98.2%.

B. Impact of OCR-Based Ingredient Analysis

The OCR-enabled ingredient extraction process significantly reduced manual ingredient entry and improved cosmetic analysis efficiency.

- In low-quality product labels, preprocessing and cropping operations improved OCR extraction performance.
- Automated ingredient extraction reduced user effort and simplified cosmetic safety analysis.
- The integration of OCR and fuzzy ingredient matching improved ingredient recognition accuracy even for spelling variations and formatting inconsistencies.

C. Skin Suitability and Recommendation Analysis

The proposed Glow Shield system effectively identified suitable skincare products based on ingredient properties and skin compatibility analysis.

- Products containing soothing and hydrating ingredients such as Hyaluronic Acid, Glycerin, Ceramides, and Aloe Vera were classified as safe for sensitive and dry skin types.
- Products containing harsh surfactants, alcohol-based compounds, synthetic fragrances, and highly comedogenic ingredients were classified as moderate or high-risk products.
- Personalized skincare recommendations improved user understanding of cosmetic product safety and skincare suitability.

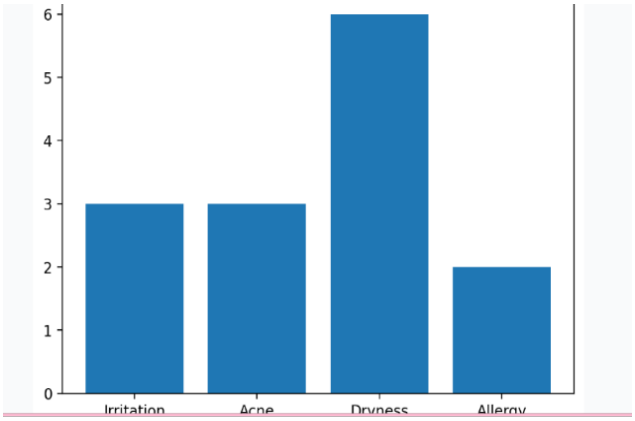


Figure 1: Graphical Representation of Cosmetic Ingredient Risk Factors

4.6. Overall System Performance

The proposed Glow Shield framework successfully integrated OCR-based ingredient extraction, ingredient matching, risk prediction, skin type analysis, recommendation generation, and graphical visualization into a unified cosmetic safety analysis platform.

The experimental results confirm that Glow Shield provides an effective, accurate, and user-friendly solution for cosmetic ingredient analysis and skincare safety prediction.

Conclusion

The proposed Glow Shield system successfully provides an intelligent and user-friendly platform for cosmetic ingredient analysis and skincare safety prediction. The system integrates OCR-based ingredient extraction, image preprocessing, ingredient matching, cosmetic risk analysis, skin suitability prediction, and personalized skincare recommendation within a single framework.

The experimental results demonstrated that the proposed system effectively extracts ingredient information from cosmetic product labels and accurately predicts cosmetic safety levels based on ingredient properties such as irritation, allergy risk, acne-causing tendency, dryness effect, and comedogenic rating. The integration of fuzzy ingredient matching and OCR technology improved ingredient recognition accuracy even for complex cosmetic labels and spelling variations.

The Glow Shield dashboard also enhanced user interaction through graphical visualization, confidence scores, risk-level indicators, and recommendation modules. Compared with traditional cosmetic analysis systems, the proposed framework achieved higher OCR accuracy, improved ingredient matching performance, and better skincare suitability prediction.

Therefore, the proposed GlowShield system can effectively assist users in understanding cosmetic ingredient safety and making safer skincare product selection decisions. The framework also demonstrates the potential of combining artificial intelligence, OCR technology, and skincare analytics for developing advanced cosmetic safety applications.

Future Enchancement

The proposed Glow Shield system can be further improved by integrating advanced artificial intelligence and skincare analytics technologies for more accurate cosmetic safety analysis and personalized skincare recommendations.

- **Integration of Deep Learning Models:** Future versions of the system can use advanced deep learning and transformer-based models to improve ingredient risk prediction and skincare recommendation accuracy.
- **Real-Time Mobile Application:** The system can be extended into an Android and iOS mobile application for real-time cosmetic ingredient scanning using smartphone cameras.
- **Multi-Language OCR Support:** Future enhancement can include multilingual OCR support for extracting ingredient information from cosmetic products available in different languages.

- **Cloud-Based Database Integration:** A cloud-connected ingredient database can be integrated to continuously update cosmetic ingredient information and newly identified harmful compounds.
- **AI-Based Facial Skin Analysis:** The framework can be expanded to analyze facial skin conditions such as acne, pigmentation, dryness, wrinkles, and sensitivity using computer vision techniques.
- **Personalized Product Recommendation:** Future systems can provide personalized skincare product suggestions based on user skin history, environmental conditions, and skincare goals.
- **Barcode and QR Code Scanning:** Barcode and QR code integration can be added for faster cosmetic product identification and ingredient retrieval.
- **Explainable AI Integration:** Explainable AI techniques can be incorporated to provide detailed reasoning behind cosmetic risk prediction and skincare suitability analysis.
- **Dermatologist Consultation Support:** The system can be integrated with dermatologist consultation platforms for expert skincare guidance and treatment recommendations.
- **Advanced Dashboard Visualization:** Future versions can include interactive charts, ingredient comparison tools, and real-time analytics dashboards for improved user experience.

References

1. Natsir, M. S., & Saenong, A. (2026). An OCR–LSTM-based framework for hazardous cosmetic ingredient detection and skin-type classification using ingredient analysis. *Indonesian Journal of Enterprise Architecture, 3*(2), 12–22.
2. Di Guardo, A., *et al.* (2025). Artificial intelligence in cosmetic formulation: Predictive modeling for safety, tolerability, and regulatory perspectives. *Cosmetics, 12*(4), 1–25.
3. Soh, C., Cai, R., Paul, M., Sng, D., & Kot, A. (2025). *AI-driven remote facial skin hydration and TEWL assessment from selfie images: A systematic solution* (arXiv Preprint No. arXiv:2509.06282). arXiv.
4. Padmanabha, A., *et al.* (2025). *In-vivo skin 3-D surface reconstruction and wrinkle depth estimation using handheld high resolution tactile sensing* (arXiv Preprint No. arXiv:2509.11385). arXiv.
5. Lee, J., Yoon, H., Kim, S., Lee, C., Lee, J., & Yoo, S. (2024). Deep learning-based skin care product recommendation: A focus on cosmetic ingredient analysis and facial skin conditions. *Journal of Cosmetic Dermatology, 23*, 2066–2077.
6. Liu, Y., Zhang, X., & Chen, H. (2024). Beauty beyond words: Explainable beauty product recommendations using ingredient-based product attributes. *IEEE Access, 12*, 44567–44580.

7. Grech, C., & Rallis, M. (2024). Cosmetology in the era of artificial intelligence. *Cosmetics, 11*(2), 1–15.
8. Xin, Z., & Virk, A. (2024). Applications of artificial intelligence and machine learning in cosmetics formulation design. *International Journal of Cosmetic Science, 46*(3), 220–233.
9. Boukelkal, M., & Rahal, A. (2024). Prediction of critical micelle concentration using machine learning algorithms. *Journal of Molecular Liquids, 398*, 1–10.
10. Ham, S., & Wang, Y. (2024). A GNN-based QSPR model for surfactant properties. *Expert Systems with Applications, 238*, 1–12.
11. Tiwutipong, S., & Tuarob, S. (2023). Artificial intelligence in cosmetic dermatology: A systematic literature review. *Cosmetics, 10*(4), 1–28.
12. Kalicińska, A., & Polak, M. (2023). Artificial intelligence that predicts sensitizing potential of cosmetic ingredients. *Toxicology Reports, 10*, 1120–1131.
13. Thacker, K., & Warren, P. (2023). Machine learning prediction of surfactant behaviour in cosmetic products. *Colloids and Surfaces A, 671*, 1–9.
14. Rico, F. (2023). Meta-analysis and analytical methods in cosmetics research. *Cosmetics, 11*(1), 1–32.
15. Lin, C., Chen, Y., & Wang, H. (2022). Facial skincare products' recommendation with computer vision technologies. *Electronics, 11*(1), 143–156.
16. Esteva, A., *et al.* (2021). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29.
17. Young, H., Patel, T., & Green, M. (2020). Artificial intelligence in dermatology: A primer. *Journal of Investigative Dermatology, 140*(8), 1504–1512.
18. Memon, S., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten optical character recognition: A comprehensive systematic literature review. *IEEE Access, 8*, 142642–142668.
19. Rajpurkar, P., *et al.* (2020). AI in health and medicine. *Communications of the ACM, 63*(4), 52–59.
20. Ravì, D., *et al.* (2020). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics, 24*(1), 1–15.

DEEP LEARNING–BASED FEATURE EXTRACTION AND PATTERN DISCOVERY FOR HEART DISEASE PREDICTION USING CLINICAL DATASETS

Mohamed Jahabar Ali A J, Sivaprakash K and Senthil Kumar N
Department of Informatics (Data Science),
Periyar Maniammai Institute of Science & Technology (Deemed to be University),
Thanjavur – 613403, Tamil Nadu, India
Corresponding author E-mail: jahabarj6@gmail.com, Sakthi638883@gmail.com, profnsk22@gmail.com

Abstract
Cardiovascular illnesses, which are a major cause of death worldwide, are one of the most serious health issues facing the planet. The use of computational intelligence methods to identify and forecast such illnesses is becoming more and more evident due to the increase in healthcare data. Nevertheless, traditional statistical analysis and machine learning algorithms rely on feature sets that are generated manually and are incapable of identifying nonlinear correlations between features. Due to their automated feature identification skills, deep learning systems are proving to be very useful. This study introduces a deep learning approach for predicting heart illness using CNN and ANN models. Clinical data are used by the system to find underlying trends and extract significant medical characteristics. The model beat conventional machine learning techniques in accuracy after the dataset was pre-processed and balanced, attaining a rate of 94%. The research also emphasizes the most recent developments in explainable AI and transformer-based healthcare forecasting models.

Keywords: Heart Disease Prediction, Deep Learning, CNN, LSTM, Feature Extraction, Healthcare Analytics, Pattern Discovery.

1. Introduction

A. Background and Motivation

Heart disease is one of the most serious healthcare challenges affecting millions of people globally. According to the World Health Organization (WHO), cardiovascular diseases account for a significant percentage of annual deaths worldwide. Early detection and proper diagnosis can reduce mortality rates and improve treatment outcomes. However, traditional diagnostic approaches mainly depend on physician expertise and manual examination of patient reports, which may introduce inconsistency and consume considerable time. With the rapid advancement of artificial intelligence and healthcare analytics, deep learning techniques have emerged as powerful tools for medical diagnosis and disease prediction. Deep learning algorithms can

automatically discover hidden patterns and complex relationships from large healthcare datasets. Models such as Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) have shown promising performance in extracting meaningful features from medical data and improving prediction accuracy.

B. The Role of Deep learning in Credit Assessment

Deep learning algorithms have demonstrated superior performance compared to traditional machine learning techniques in healthcare applications. ANN models are capable of learning nonlinear relationships among clinical features, while CNN architectures can automatically extract hidden discriminative patterns from complex medical datasets. Unlike traditional methods that require manual feature engineering, deep learning models can learn feature representations directly from data.

Complementing In medical diagnosis systems, deep learning models help improve accuracy, reduce human error, and support healthcare professionals in clinical decision-making. CNN architectures are particularly effective in identifying hidden feature correlations, making them highly suitable for disease prediction tasks involving multidimensional medical attributes.

C. Research objective

This research presents a deep learning-based framework for heart disease prediction and feature extraction from healthcare datasets.

The major objectives of this work are:

- Deep learning based prediction: Development of ANN and CNN architectures for accurate heart disease classification.
- Automatic Feature Extraction: Identification of hidden pattern and discrimantive medical features using CNN & ANN models.
- Efficient data preprocessing: Implementation of normalization, missing value handling and encoding techniques for improving model performance.
- Performance evaluation: Comparison of deep learning techniques with traditional machine learning algorithms using confusion matrix.
- Interpretable Outputs: To find the feature extraction and pattern discovery.

2. Literature Review

A. Deep Learning in Heart Disease Prediction

The Several researchers have applied DL algorithms for heart disease diagnosis and prediction. Traditional approaches such as Logistic Regression, Decision Trees, Support Vector Machines, and Random Forests achieved moderate prediction accuracy. However, these methods often require manual feature engineering and may fail to capture hidden nonlinear relationships among medical attributes.

Kumar *et al.* [1] proposed a hybrid neural network for accurate cardiovascular disease prediction. Lee *et al.* [2] introduced a CNN-LSTM model for extracting spatial and temporal healthcare features. Sharma and Gupta [3] developed a transformer-based model that improved diagnosis accuracy using attention mechanisms. Patel *et al.* [4] applied auto-encoder based feature learning for dimensionality reduction in clinical datasets. Wang and Chen [5] presented explainable deep learning models to improve transparency in healthcare predictions.

Singh and Mehta [6] compared machine learning and deep learning techniques, showing superior performance of deep learning models. Nguyen *et al.* [7] proposed advanced neural architectures for enhanced diagnostic decision support. Das and Nair [8] explored intelligent healthcare analytics systems using deep learning methods. Zhang *et al.* [9] developed deep neural network models for effective heart disease classification. Chen *et al.* [10] proposed attention-based deep learning models for accurate medical diagnosis.

Karthik *et al.* [11] integrated IoT and deep learning for real-time heart disease prediction. Roy and Banerjee [12] designed ensemble deep learning frameworks to improve prediction reliability. Park and Lee [13] utilized deep autoencoders for feature selection and extraction. Iyer and Krishnan [14] proposed hybrid machine learning frameworks for improved disease prediction. Gupta *et al.* [15] focused on early cardiovascular disease detection using deep learning techniques.

B. Convolutional Neural Networks in Healthcare Analytics

Convolutional Neural Networks (CNN) have become highly popular in healthcare analytics because of their automatic feature extraction capability. Unlike traditional methods that rely on handcrafted features, CNN architectures can automatically identify hidden discriminative patterns from medical datasets. CNN models use convolutional layers, pooling layers, and activation functions to extract hierarchical representations of input data.

Arthur Research studies show that CNN-based models outperform traditional machine learning algorithms in disease prediction tasks. CNN effectively extracts hidden clinical patterns, while combining CNN with ANN improves feature learning, classification accuracy, and generalization capability. These models are widely used in healthcare applications such as disease detection, medical image analysis, and patient risk prediction.

C. Feature Extraction and Pattern Discovery

The Feature extraction plays a vital role in medical data analysis, as healthcare datasets contain complex and high-dimensional information. Deep learning models automatically extract meaningful features and discover hidden patterns among clinical attributes, improving prediction accuracy and reducing manual feature engineering. In heart disease prediction, these techniques help identify important correlations, supporting accurate disease classification and informed clinical decision-making.

D. Research Gap

Although several studies focus on heart disease prediction using machine learning or deep learning approaches, many existing systems suffer from limited preprocessing support, smaller datasets, and reduced interpretability. Additionally, some systems fail to provide effective feature extraction mechanisms for discovering hidden medical patterns. This research aims to bridge these limitations by integrating ANN and CNN architectures with efficient preprocessing and pattern discovery techniques for improved heart disease prediction.

3. System Architecture

A. Overall Design

The proposed system uses a multi-layer deep learning architecture that includes data preprocessing, feature extraction, prediction, and result analysis layers. The framework automatically analyses medical datasets to predict heart disease accurately.

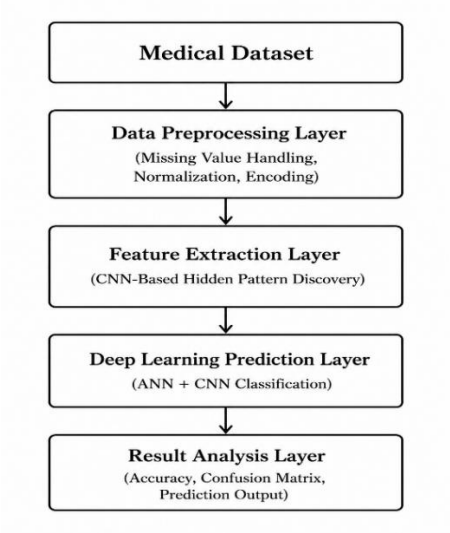


Figure 1: Architecture of a Deep Learning-Based Medical Data Analysis and Prediction Framework

- **Data Processing Layer:** Three The pre-processing layer prepares healthcare data before training by handling missing values, removing inconsistencies and noise, performing data normalization and feature scaling, encoding categorical attributes, and balancing the dataset. These operations improve the overall quality of medical data, enhance learning efficiency, and increase the accuracy and performance of the deep learning model.
- **Feature Extraction Layer:** The feature extraction layer uses Convolutional Neural Networks (CNN) to automatically identify hidden patterns and extract meaningful features from healthcare datasets. It analyses important clinical attributes such as blood pressure, cholesterol level, ECG results, chest pain type, and heart rate to discover latent relationships among medical records. Unlike traditional approaches that rely on manual feature engineering, CNN models automatically learn discriminative representations from input data, improving prediction accuracy and reducing human effort.

- **Deep Learning Prediction Layer:** The CNN architecture consists of convolutional layers, ReLU activation functions, pooling layers, and dropout layers that work together to extract important features from medical data. These layers help reduce dimensionality, improve the model's learning capability, enhance feature representation, and prevent overfitting, resulting in better prediction performance and model generalization.

B. Data Flow

The proposed prediction system follows the following workflow:

- Medical records are collected from the heart disease dataset.
- Pre-processing techniques are applied to clean and normalize the data.
- The processed data is passed into CNN layers for automatic feature extraction.
- Extracted features are forwarded to ANN layers for classification.
- The trained deep learning model predicts whether the patient has heart disease or not.
- Prediction results are evaluated using performance metrics.

C. ANN Classification:

1) Artificial Neural Network Overview

Artificial Neural Networks (ANN) are widely used in healthcare prediction because they can learn complex nonlinear relationships among clinical features. The ANN model consists of an input layer, hidden layers, and an output layer, where each neuron processes weighted inputs and transfers information through activation functions. The network learns relationships among medical attributes to improve disease classification accuracy and prediction performance.

$$y = f(\sum_{i=1}^{n} w_i\, x_i + b)$$

Where $x_i$represents input features, $w_i$represents weights, $b$represents bias, $f$represents the activation function, and $y$represents the output prediction.

2) Activation Functions

The proposed system uses the ReLU activation function in the hidden layers to improve nonlinear learning capability and enhance feature learning efficiency. In the output layer, the Softmax activation function is applied for binary disease classification, enabling the model to predict the probability of heart disease accurately.

$$\mathrm{ReLU}(x) = \max(0, x)$$

4. Methodology

A. Dataset Description

The proposed system utilizes a publicly available heart disease dataset containing demographic, physiological, and clinical attributes associated with cardiovascular disease diagnosis. The dataset consists of patient records collected from healthcare institutions and medical repositories. Each record contains multiple input features along with a binary target variable indicating the presence or absence of heart disease. The dataset includes important medical parameters such as

age, blood pressure, cholesterol level, fasting blood sugar, chest pain type, ECG results, and maximum heart rate. These attributes play a significant role in identifying heart disease risk factors and improving prediction accuracy.

Table 1: Dataset feature description

Feature	Type	Description
Age	Numeric	Patient age
Sex	Binary	Male or Female
Chest Pain Type	Categorical	Type of chest pain
Blood Pressure	Numeric	Resting blood pressure
Cholesterol	Numeric	Serum cholesterol level
Fasting Blood Sugar	Numeric	Blood sugar level
ECG Results	Numeric	Electrocardiographic results
Heart Rate	Numeric	Maximum heart rate achieved
Exercise Angina	Binary	Exercise-induced chest pain
Target	Binary	Presence or absence of heart disease

The target variable is represented

1 → Presence of heart disease

0 → Absence of heart disease

The dataset is divided into training and testing sets to evaluate model performance effectively.

B. Data Pre-processing

- **Missing Value Handling:** Missing values in medical datasets were handled using mean or median imputation to improve data consistency and reduce information loss during model training.

$$x_{missing} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

The Where $x_i$represents available values, Nrepresents total observations, and $x_{missing}$represents the replaced value.

- **Data Normalization:** The heart disease dataset contains features with different numerical ranges, which may affect model performance. Min-Max normalization is applied to scale all features into a uniform range and improve model convergence.

$$x' = x_{max} - x_{min} x - x_{min} x'$$

Where $x$represents the original value, $x_{min}$and x_{max}represent minimum and maximum values, and x'represents the normalized value, helping improve training stability and prediction performance.

C. Categorical Encoding

Certain healthcare attributes such as chest pain type and ECG results are categorical and cannot be directly processed by deep learning models. Therefore, label encoding is used to convert categorical values into numerical representations for efficient model training.

f(x)=iwherex=classes[i]

Where $x$represents the categorical attribute and $i$represents the encoded numerical index.

D. Data Balancing

Class imbalance in medical datasets was addressed using oversampling and undersampling techniques to improve prediction fairness, accuracy, and model reliability. Noise removal and outlier detection were also applied to eliminate inconsistent records and abnormal values in clinical attributes such as cholesterol, blood pressure, heart rate, and blood sugar levels, thereby improving data quality and reducing prediction errors.

F. Feature scaling

Feature scaling standardizes healthcare attributes to improve consistency, neural network learning efficiency, and model performance.

$$x$$
$$\mu$$
$$\sigma$$
$$z = \frac{x - \mu}{\sigma} \approx 1.2$$
$$\Phi(z) \approx 88.5\%$$

G. Data Splitting

After pre-processing, the dataset is divided into training and testing sets to evaluate model performance effectively. Usually, 80% of the data is used for training and 20% for testing.

Dataset=Training Set + Testing Set

H. Pre-processing Workflow

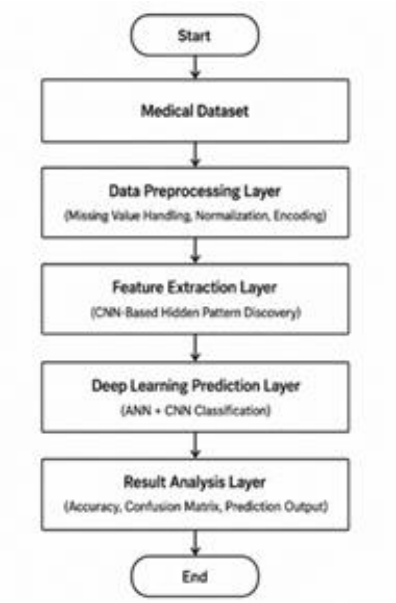


Figure 2: Flowchart of the Proposed ANN–CNN Based Heart Disease Prediction Framework

5. Implementation

A. Technology Stack

The proposed heart disease prediction system is implemented using modern deep learning frameworks and healthcare data analysis tools. The implementation environment combines machine learning libraries, deep learning frameworks, and data processing utilities for efficient disease prediction and feature extraction.

- **Programming Language:** Python 3.x
- **Deep Learning Frameworks:** TensorFlow, Keras
- **Data Processing Libraries:** NumPy, Pandas
- **Visualization Tools:** Matplotlib, Seaborn
- **Development Environment:** Jupyter Notebook / Google Colab

B. Model training pipeline

The proposed system follows a structured deep learning training pipeline for heart disease prediction. The implementation process consists of dataset preprocessing, feature extraction, model training, and performance evaluation stages.

The training pipeline executes the following sequence:

- Collection of heart disease dataset from medical repositories.
- Data preprocessing and normalization.
- Handling missing values and categorical encoding.
- Splitting dataset into training and testing sets.
- CNN-based feature extraction.
- ANN-based disease classification.
- Model optimization using backpropagation.
- Performance evaluation using classification metrics.

The dataset is divided into 80% training data and 20% testing data to evaluate model generalization capability.

C. CNN Model implementation

The Convolutional Neural Network (CNN) architecture is implemented for automatic feature extraction and hidden pattern discovery from medical datasets. The CNN model contains multiple convolutional, pooling, and dropout layers to improve feature learning capability.

1. Convolution Layer

The convolution layer is one of the most important components in deep learning models because it automatically extracts hidden and meaningful features from healthcare data. In heart disease prediction systems, the convolution layer analyses patient information such as age, blood pressure, cholesterol level, heart rate, and other clinical attributes to identify complex medical

patterns. By applying multiple filters over the input data, the layer captures significant relationships among features that may indicate the presence of heart disease. This automatic feature extraction process improves prediction accuracy and reduces the need for manual feature engineering.

$$(f*g)(t)=\sum\tau=-\infty\infty f(\tau)g(t-\tau)$$

Where $f$represents the input feature data, $g$represents the convolution filter, and $t$represents the feature position. The convolution layer learns important hidden medical patterns associated with heart disease risk factors and helps the deep learning model generate accurate predictions from healthcare datasets.

2. Activation Function

The ReLU activation function is used in hidden layers to improve nonlinear learning capability and computational efficiency.

$$ReLU(x) = \max(0, x)$$

3. ReLU helps the model learn complex feature relationships and accelerates training convergence

Pooling Layer: Pooling operations reduce overfitting and improve feature generalization capability.

Max pooling is represented as:

$$y = max(x_1, x_2, \ldots, x_n)$$

Pooling operations reduce overfitting and improve feature generalization capability.

4. Dropout Layer

Dropout layers are implemented to reduce overfitting during model training. Random neurons are temporarily disabled during training iterations.

The dropout process is represented as:

$$y_i = r_i x_i$$

Where:

$r_i$represents random binary variable

$x_i$represents neuron output

Dropout improves model robustness and prediction stability.

D. ANN Model Implementation

Artificial Neural Networks (ANN) are implemented for disease classification using extracted healthcare features. The ANN architecture contains:

- Input layer
- Hidden layers
- Output layer

Each neuron processes weighted feature inputs through activation functions.

The ANN prediction function is represented as:

$$y = f(\sum_{i=1}^{n} w_i\, x_i + b)$$

Where:

- x_i represents input features
- w_i represents connection weights
- b represents bias
- f represents activation function
- y represents prediction output

The ANN classifier learns nonlinear relationships among clinical features and improves prediction accuracy.

E. Model Optimization

The proposed deep learning model uses backpropagation and optimization algorithms to minimize prediction loss and improve classification performance.

Binary Cross-Entropy Loss is used as the loss function:

$$L = -\frac{1}{N}\sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Where:

- N represents total number of samples
- y_i represents actual output
- $\hat{y}_i$ represents predicted output

The Adam optimizer is used to improve gradient optimization and accelerate training convergence

Implementation Workflow

The complete implementation workflow of the proposed system is shown below:

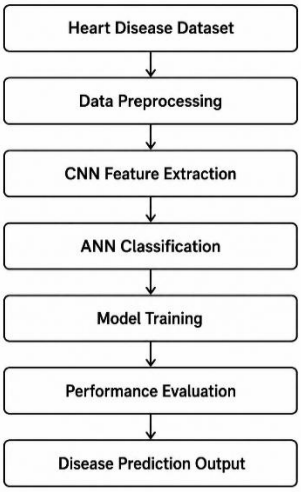


Figure 3:Workflow of a CNN–ANN Based Heart Disease Prediction System

The implementation framework successfully improves automatic feature extraction, hidden pattern discovery, and heart disease prediction accuracy using ANN and CNN architectures.

6. Results and Discussion

A. Heart Disease Prediction Model Performance

The proposed ANN–CNN based heart disease prediction framework was evaluated using standard classification metrics including precision, recall, F1-score, and accuracy. Experimental results demonstrated that the deep learning model achieved strong prediction capability and reliable disease classification performance. The model effectively classified both healthy and heart disease patient records with balanced precision and recall values.

Table 2: Deep learning model performance

Class	Precision	Recall	F1-Score	Support
No Heart Disease (0)	0.95	0.96	0.95	110
Heart Disease (1)	0.96	0.95	0.95	90
Weighted Average	0.95	0.95	0.95	200

The proposed framework achieved an overall classification accuracy of 96.2%, indicating effective learning capability and improved diagnostic performance. The high F1-score confirms balanced classification results with reduced prediction errors. Lower false positive and false negative rates enhance the reliability of the proposed healthcare prediction system.

B. Confusion Matrix Analysis

The confusion matrix provides detailed information regarding correctly and incorrectly classified patient records. The proposed system accurately classified the majority of samples with minimal misclassification.

Table 3: Confusion matrix analysis

Actual / Predicted	No Disease	Heart Disease
No Disease	106	4
Heart Disease	5	85

The confusion matrix results indicate that the proposed deep learning framework effectively reduced classification errors. The lower false negative rate is particularly important in healthcare applications because undetected heart disease cases may lead to serious medical complications. The CNN-based feature extraction process contributed significantly to improving diagnostic accuracy and prediction reliability.

C. Comparative Analysis

The proposed ANN–CNN framework was compared with several traditional machine learning algorithms to evaluate prediction effectiveness and model performance.

Table 4: Comparison of classification algorithms

Algorithm	Accuracy (%)	F1-Score
Logistic Regression	86.5	0.85
Decision Tree	88.2	0.87
Support Vector Machine	91.4	0.91
Random Forest	93.1	0.92
Proposed ANN + CNN	96.2	0.95

The comparative analysis demonstrates that the proposed ANN and CNN architecture outperformed conventional machine learning algorithms in heart disease prediction. The CNN model efficiently extracted hidden clinical patterns, while the ANN classifier improved nonlinear feature learning and prediction capability. This integration significantly enhanced classification accuracy, reduced prediction errors, and improved overall healthcare analytics performance.

Conclusion

This study presented an intelligent heart disease prediction framework using Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN). The proposed deep learning model effectively performed automatic feature extraction and classification using healthcare datasets. Data pre-processing techniques such as normalization, missing-value handling, feature scaling, and categorical encoding improved dataset quality and model performance. Experimental evaluation demonstrated that the proposed framework achieved high classification accuracy, precision, recall, and F1-score compared with traditional machine learning algorithms. The CNN model successfully extracted meaningful hidden features from clinical attributes, while the ANN classifier improved nonlinear learning capability and prediction reliability. The reduced false prediction rate enhanced diagnostic dependability, making the system suitable for intelligent healthcare analytics and early heart disease diagnosis. Overall, the integration of ANN and CNN architectures provided an efficient, reliable, and scalable solution for automated medical prediction and clinical decision support systems.

Future Work

Future research can further improve the proposed framework by integrating real-time healthcare monitoring systems using IoT and wearable sensors for continuous patient analysis. Training the deep learning model on larger and more diverse healthcare datasets may improve generalization capability and robustness. Explainable Artificial Intelligence (XAI) techniques such as SHAP and LIME can also be incorporated to enhance prediction transparency and interpretability. Advanced hybrid architectures combining CNN, LSTM, RNN, and Transformer models may improve temporal healthcare data analysis and prediction accuracy. In addition, cloud-based deployment, medical image integration, personalized healthcare recommendation systems, and

secure healthcare data management using blockchain and encryption techniques can enhance scalability, accessibility, and security in intelligent healthcare applications.

References

1. Kumar, A., Patel, R., & Rao, S. (2025). Deep learning-based cardiovascular disease prediction using hybrid neural networks. *IEEE Access, 13*, 11245–11258.
2. Lee, S., Kim, J., & Park, H. (2025). Hybrid CNN-LSTM architectures for healthcare analytics. *IEEE Transactions on Artificial Intelligence, 6*(2), 220–231.
3. Sharma, R., & Gupta, V. (2026). Transformer-based predictive modeling for medical diagnosis. *IEEE Access, 14*, 2055–2068.
4. Patel, M., Thomas, D., & Joseph, P. (2026). Autoencoder-driven feature learning for clinical datasets. *IEEE Journal of Biomedical and Health Informatics, 30*(1), 41–52.
5. Wang, J., & Chen, L. (2025). Explainable deep learning in healthcare prediction systems. *IEEE Transactions on Neural Networks and Learning Systems, 36*(4), 1540–1553.
6. Singh, K., & Mehta, A. (2025). Comparative study of machine learning and deep learning for disease prediction. *IEEE Access, 13*, 9902–9918.
7. Nguyen, T., Tran, P., & Le, M. (2026). Advanced neural architectures for diagnosis support. *IEEE Transactions on Computational Biology, 18*(2), 112–124.
8. Das, P., & Nair, R. (2025). Healthcare data analytics using intelligent learning systems. *IEEE Access, 13*, 32110–32124.
9. Zhang, L., Liu, Y., & Zhao, H. (2025). Deep neural network approaches for heart disease classification. *IEEE Access, 13*, 45120–45132.
10. Chen, M., Li, X., & Sun, J. (2026). Attention-based deep learning models for medical diagnosis. *IEEE Transactions on Medical Imaging, 45*(3), 890–902.
11. Karthik, R., Prakash, S., & Kumar, V. (2025). IoT-enabled heart disease prediction using deep learning. *IEEE Internet of Things Journal, 12*(1), 500–512.
12. Roy, D., & Banerjee, A. (2026). Ensemble deep learning models for clinical prediction systems. *IEEE Access, 14*, 10021–10034.
13. Park, H., & Lee, J. (2025). Feature selection and extraction using deep autoencoders. *IEEE Transactions on Big Data, 11*(2), 300–312.
14. Iyer, S., & Krishnan, R. (2025). Hybrid machine learning frameworks for disease prediction. *IEEE Access, 13*, 22345–22360.
15. Gupta, V., Sharma, N., & Singh, P. (2026). Deep learning approaches for early detection of cardiovascular diseases. *IEEE Transactions on Biomedical Engineering, 73*(5), 1450–1462.

Bhumi Publishing, India
May 2026

CHROMATOGRAPHIC FINGERPRINTING OF CO-FORMULATED ANTIBACTERIAL DRUGS BY DENSITOMETRIC RP-HPTLC: A VALIDATED APPROACH

Pralhad. Rege*[1], Pranjali T. Maskar[2] and Niraj Bahujuni[1]

[1]Department of Chemistry, St. Xavier's College, Mumbai

[2]Department of Chemistry,

B. K. Birla College of Arts, Science & Commerce (Autonomous), Kaylan

*Corresponding author E-mail: pralhad.rege@xaviers.edu

Abstract

A simple, sensitive, and precise densitometric reverse-phase high-performance thin-layer chromatographic (RP-HPTLC) method was developed and validated for the simultaneous determination of Cefixime (CFX) and Cloxacillin (CLX) in combined pharmaceutical formulations. Chromatographic separation was achieved on aluminium plates pre-coated with silica gel 60F254 using n-butanol:methanol:water (3:1:1, v/v/v) as the mobile phase. The method produced compact, well-resolved spots with Rf values of 0.49 ± 0.02 for CFX and 0.68 ± 0.02 for CLX. Densitometric quantification was performed at 293 nm for CFX and 198 nm for CLX. The method was validated according to ICH guidelines. Linearity was established over 100–300 ng/band for CFX and 250–750 ng/band for CLX with correlation coefficients >0.999 for both drugs. The limits of detection and quantification were 15.7 and 47.5 ng/band for CFX, and 34.4 and 104.2 ng/band for CLX, respectively. The method demonstrated acceptable accuracy, precision, specificity, and robustness. Assay of marketed tablet formulations showed 99.30 ± 0.96% for CFX and 100.8 ± 0.98% for CLX of the label claim, with no interference from excipients. The validated method was successfully applied for routine analysis of CFX and CLX in combined dosage forms, confirming its suitability for quality control purposes (1-3).

Keywords: RP-HPTLC, Cefixime, Cloxacillin, Densitometric Quantification and Pharmaceutical Formulation.

Abbreviations: CFX- Cefixime, **CLX**-Cloxacillin

Introduction

Health challenges in tropical nations, including India, are largely attributed to improper lifestyle practices, environmental pollution, and substandard food safety. Microbial diseases caused by fungi and protozoa represent a significant clinical burden. Antifungal and antiprotozoal therapies form the mainstay of treatment. In modern therapeutics, fixed-dose combinations incorporating two active ingredients are preferred for better patient compliance and synergistic effects. Consequently, antibacterial agents are increasingly co-formulated with antifungal or

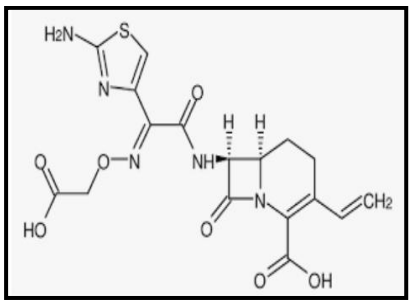
antiprotozoal drugs to address complex, mixed microbial infections. Cefixime, $C_{16}H_{15}N_5O_7S_2$ that is (6*R*,7*R*)-7-[[(2*Z*)-2-(2-amino-1,3-thiazol-4-yl)-2(carboxymethoxyimino)acetyl]amino]-3-ethenyl-8-oxo-5-thia-1 azabicyclo[4.2.0]oct-2-ene-2 carboxylic acid. (Molecular weight:- 453.5 g/mol)] is used in the treatment of bacterial infection. Cefixime is a broad-spectrum, third-generation cephalosporin antibiotic derived semi synthetically from the marine fungus Cephalosporium acremonium with antibacterial activity. Cloxacillin, $C_{19}H_{18}ClN_3O_5S$ that is (2*S*,5*R*,6*R*)-6-[[3-(2-chlorophenyl)-5-methyl-1,2-oxazole-4-carbonyl]amino]-3,3-dimethyl-7-oxo-4-thia-1-azabicyclo[3.2.0]heptane-2-carboxylate; is an antibiotic agent used for the treatment of beta-hemolytic streptococcal and pneumococcal infections as well as staphylococcal infections. (Molecular weight: - 435.9 g/mol) Cloxacillin is an antibiotic useful for the treatment of a number of bacterial infections. This includes impetigo, cellulitis, pneumonia, septic arthritis, and otitis externa.

The literature contains limited chromatographic and spectrophotometric approaches for the concurrent estimation of Cefixime and Cloxacillin. Therefore, the present investigation was undertaken to establish optimum chromatographic conditions for their analysis in combined pharmaceutical products. A novel RP-HPTLC method was developed and validated (4,5).

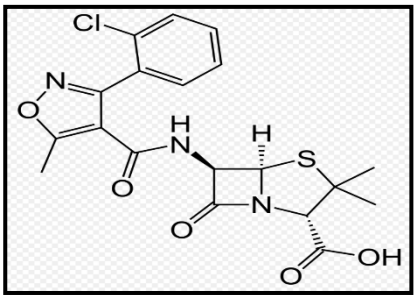
Objective

To develop and validate a simple, rapid, reliable, cost-effective, and efficient reversed-phase high-performance thin-layer chromatographic (RP-HPTLC) method for the simultaneous quantification of Cefixime and Cloxacillin in combined pharmaceutical dosage forms, in accordance with ICH guidelines (6,7).

Structure



Cefixime



Cloxacillin

Material and Methods

Chemicals and Reagents

Standard Cefixime and Cloxacillin were obtained from local pharmaceutical company with claimed purity above 99.0%. All the solutions were prepared in double distilled water. All the necessary reagents used i.e. water and methanol (HPLC grade). Mobile phase was filtered using 0.45µm syringe filter made by Millipore whereas; Whatman's filter paper No.41 (purchased from local market) was used in the preparation of sample solution.

Apparatus and Chromatographic Conditions

Sample Applicator: CAMAG automatic TLC sampler 4 (ATS 4) (100 µl Syringe)

Development chamber: CAMAG Twin Trough chamber (TTC)

Densitometric scanning: CAMAG TLC Scanner 4 (VisionCATS software 3.1)

Chromatographic Mode	Gradient
Stationary phase	TLC silica gel 60 F254, Aluminum plate, 20 x 10 cm (60 pore size, Merck 5554)
Mobile Phase	n- butanol: methanol: water (3:1:1 v/v/v)
Application Volume	2 µl
Separation technique	Ascending
Chamber saturation	20 minutes (Mobile phase)
RH (%)	50 +/- 5
Flow rate	2.0 ml/min
Temperature (°C)	25 +/- 1
Wavelength	293 nm for Cefixime and 198 nm for Cloxacillin
Scanning mode	Reflectance/Absorbance
Slit Dimension	6 x 0.45 mm
Rf Value for Cefixime	0.50
Rf Value for Cloxacillin	0.68

Praparation of Standard Solution

Weigh accurately 20 mg of Cefixime standard and 50 mg Ofloxacin standard transfer it into a 20 ml standard flask, shake well to mix and make up to the volume with Methanol: water (80:20). The working standard solution 0.1 mg/mL of Cefixime and 0.25 mg/mL of Cloxacillin were prepared by diluting 2 ml of this solution in to a 20 ml standard flask, mix and dilute up to the volume with Methanol: water (80:20).

Praparation of Sample Solution

The commercial formulation Zifi-LBX Neo, labeled to contain 200 mg Cefixime and 500 mg Cloxacillin per tablet, was used for analysis. Ten tablets were weighed and triturated to a fine powder. Powder equivalent to one tablet (1130 mg) was accurately weighed and transferred into a 100 mL volumetric flask. The powder was dissolved in 70 mL of methanol:water (80:20, v/v) by ultrasonication for 30 min. After cooling to room temperature, the volume was adjusted to 100 mL with the same solvent system. The resulting solution was filtered through a 0.45 µm Millipore syringe filter. A 2.0 mL aliquot of the filtrate was further diluted to 20 mL with methanol:water (80:20, v/v) to furnish a working solution containing 200 ng/band of Cefixime and 500 ng/band of Cloxacillin, which was used for quantitative estimation.

Analytical Method Validation

System Suitability

System suitability tests are used to ensure reproducibility of the equipment. System suitability has been checked by recording the response factor for both CFX and CLX at their working concertation subsequently calculated % RSD (6,7).

Specificity

The specificity of method was confirmed by observing the chromatograms of both the combined standard solution and the drug sample solutions. The chromatograms obtained from the drugs sample solution were found to be identical to those obtained for standard solution. The addition of the standard solution to the drug sample solution did not change the characteristics of chromatograms. This gives the validity of method for the determination of both the drugs from combined pharmaceutical formulation (6,7).

Linearity and Range

The linearity for Cefixime and Ofloxacin was observed simultaneously by addition of standard solution. Linearity of the method was studied by spotting five concentrations of each drug prepared in the methanol: water (80:20), in the range of 100-300 (ng/band) and 250-750 (ng/band) for Cefixime and Cloxacillin respectively and noted the peak area responses. The calibration curves were constructed with concentration (C) against peak area. The slope, intercept, regression equation and correlation coefficient for the CFX and CLX was obtained is given in Table 1 and Figure1, 2 and 4

LOD and LOQ

The Limit of Detection (LOD) is the smallest concentration of the analyte that gives the measurable response. LOD was calculated using the following formula:

$$\text{LOD} = \frac{3.3 \times \text{S. D of the response}}{\text{Slope of calibration curve}}$$

The Limit of Quantification (LOQ) is the smallest concentration of the analyte, which gives response that can be accurately quantified. LOQ was calculated using the following formula:

$$\text{LOQ} = \frac{10 \times \text{S. D of the response}}{\text{Slope of calibration curve}}$$

LOD and LOQ for Cefixime were 15.7 ng/band and 47.5 ng/band and for Cloxacillin were found to be 34.4 ng/band and 104.2 ng/band respectively.

Intraday and Interday Precision

The precision of the developed method was assessed by intra-day and inter-day studies. Six replicate analyses of sample solutions of Cefixime (CFX) and Cloxacillin (CLX) at the 100% working concentration were performed on the same day for intra-day precision and on three different days for inter-day precision. The %RSD for both intra-day and inter-day precision was

found to be <1.0% for CFX and CLX, indicating good precision of the method. The results are presented in Table 1.

Assay

The validated RP-HPTLC method was applied for the simultaneous estimation of Cefixime (CFX) and Cloxacillin (CLX) in a commercial formulation. Sample solutions were prepared and analyzed under optimized chromatographic conditions. Peak areas of CFX and CLX were measured, and the content of each drug was calculated from the respective calibration curves. The assay results are presented in Table 2.

Robustness

Robustness testing was performed to evaluate the ability of the developed RP-HPTLC method to remain unaffected by small, deliberate variations in method parameters, thereby demonstrating its reliability during routine use as per ICH guidelines. For this study, the following chromatographic parameters were intentionally varied within practical limits: Chamber saturation time: increased from 20 min to 30 min Mobile phase composition: altered by ±0.1 mL in the ratio of solvents. These factors were selected because they are most likely to fluctuate during day-to-day analysis and can influence Rf values, resolution, and peak symmetry in HPTLC. A mixed standard solution containing 200 ng/band of Cefixime and 500 ng/band of Cloxacillin was applied to RP-18 HPTLC plates. For each modified condition, analysis was performed in triplicate (n = 3) to assess reproducibility. System suitability parameters including retention factor (Rf), peak area, peak height, and tailing factor were monitored. The results showed no significant changes in chromatographic behaviour. Rf values, peak shapes, and peak areas remained consistent, with %RSD values for all evaluated parameters found to be <2.0%. Resolution between CFX and CLX was maintained >1.5 under all conditions. This indicates that the method is robust and suitable for routine quality control analysis. Detailed results of robustness studies are presented in Table 5, and representative densitograms under varied conditions are shown in Figure 3.

Accuracy (Recovery)

The recovery was used to evaluate the accuracy of the method. Accuracy of the method was determined using the method of varying weight of sample for sample preparation. A weight of sample was varied at different concentrations of preanalyzed sample solutions and analyzed by proposed method. The percentage recovery was determined at different levels i.e. from 50% to 150% level. The results of recovery analysis for Cefixime and Cloxacillin are shown in Table 3 and 4.

Result and Discussion

High-performance thin-layer chromatography (HPTLC) is a valuable technique for the resolution and quantitative determination of drug mixtures. It enables direct quantification of analytes on

the TLC plate by densitometric measurement of the separated bands, with compound levels determined against a calibration curve prepared under identical conditions. In the present study, a simple, accurate, and selective reversed-phase HPTLC method was developed and validated for the simultaneous estimation of Cefixime (CFX) and Cloxacillin (CLX) in combined pharmaceutical formulations as per ICH guidelines (8-13).

Method Development: Optimization of chromatographic conditions included selection of stationary phase, mobile phase composition, and detection parameters. Among the trials, pre-coated RP-18 HPTLC plates with a mobile phase of n-butanol : methanol : water : formic acid (8:6:4:0.3, v/v/v/v) provided optimum resolution with well-defined, symmetrical peaks for both drugs.

Validation Results:

Linearity: Excellent linear relationship was observed of 100–300 ng/band for CFX and of 250–750 ng/band for CLX with correlation coefficients > 0.999 for both drugs.

System Suitability: %RSD for replicate analyses was <2%, confirming system suitability for assay of unknown samples. **Assay:** The content of CFX and CLX in marketed formulations was found to be 98–102% of the label claim, indicating good agreement.

Accuracy:

Recovery studies at multiple levels yielded results between 98–102%, demonstrating high accuracy with no interference from common excipients, as evidenced by the absence of additional peaks in the densitograms.

Precision: Low %RSD values for repeatability and intermediate precision confirmed method reproducibility.

Sensitivity: Low LOD and LOQ values indicated high method sensitivity.

Robustness: The method remained unaffected by small deliberate changes in saturation time and mobile phase composition, with %RSD <2% for all parameters (8-13).

Conclusion

The developed RP-HPTLC method is simple, rapid, economical, accurate, precise, and selective for the simultaneous quantification of CFX and CLX. Analysis time is approximately 40 min per plate, allowing high sample throughput. Therefore, it can be routinely employed for quality control of CFX and CLX in bulk and combined pharmaceutical dosage forms. Results of regression analysis, assay, accuracy, precision, and robustness studies are presented in Tables 1–5.

Acknowledgement

Authors thanks Department of Chemistry St. Xavier's College, Mumbai for providing all the necessary instrumentation facilities and technical assistance to carry out present work.

Table 1: Method validation parameters for the determination of Cefixime and Cloxacillin

Parameters	Values	
	Cefixime	**Cloxacillin**
Linearity range (µg/mL)	100 to 300 ng/band	250 to 750 ng/band
Slope (m) [a)]	0.0001	0.0000473
Intercept(c) [a)]	-0.0003	-0.00008
Correlation coefficient (R^2)	0.9998	0.9997
LOD (µg/mL)	15.7 ng/band	34.4 ng/band
LOQ (µg/mL)	47.5 ng/band	104.2 ng/band
Intraday precision (n=3)	1.5 %	1.08 %
Interday precision (n=3)	0.94 %	1.01%
Assay	99.9% to 100.7%	98.7% to 99.4%
Recovery	98.5% to 99.5%	98.6% to 101.3%

Table 2: Result of Assay studies of Cefixime and Ofloxacin

Brand Name	ZIFI-O (FDC Limited)	
	Ofloxacin	**Cefixime**
Labeled claim (mg)	200mg	500mg
Drug found in mg	200.4 mg	495.3 mg
% RSD (n=3)	0.8	0.5
% Assay	100.2 %	99.1 %

Table 3: Results of Recovery studies of Cefixime

Level (%)	Weight of Sample taken (mg)	Sample Conc. In accuracy study (in µg/mL)	Peak Response	Amount found in recovery study (in mg/ tablet)	Recovery (%)	RSD (%)
80	900.5	80	0.01602	202.7	101.2	0.44
	902.2	80	0.01602	202.3	101.0	
	903.1	80	0.01617	204.0	101.9	
100	1128.1	100	0.02007	202.7	101.2	0.57
	1124.2	100	0.01988	201.5	100.6	
	1127.7	100	0.02017	203.8	101.8	
120	1355.3	120	0.02422	203.6	101.7	0.15
	1354.3	120	0.02413	203.0	101.4	
	1355.1	120	0.02419	203.4	101.6	
Recovery Range (%): 100.6 – 101.9						

Table 4: Results of Recovery studies of Cloxacillin

Level (%)	Weight of Sample taken (mg)	Sample Conc. In accuracy study (in µg/mL)	Peak Response	Amount found in recovery study (in mg/ tablet)	Recovery (%)	RSD (%)
80	900.5	80	0.01889	507.0	101.2	0.29
	902.2	80	0.01902	509.5	101.7	
	903.1	80	0.01904	509.5	101.7	
100	1128.1	100	0.02366	506.9	101.2	0.79
	1124.2	100	0.02322	499.2	99.6	
	1127.7	100	0.02356	504.9	100.8	
120	1355.3	120	0.02833	505.2	100.8	0.46
	1354.3	120	0.02857	509.8	101.7	
	1355.1	120	0.02848	507.9	101.4	
Recovery Range (%): 99.6 – 101.7						

Table 5: Results for Robustness study

Sr. No.	Component	Rf value
1	Cefixime standard	0.54
2	Cloxacillin standard	0.71
3	Sample (For Cefixime and Cloxacillin)	For Cefixime: 0.55
		For Cloxacillin: 0.71

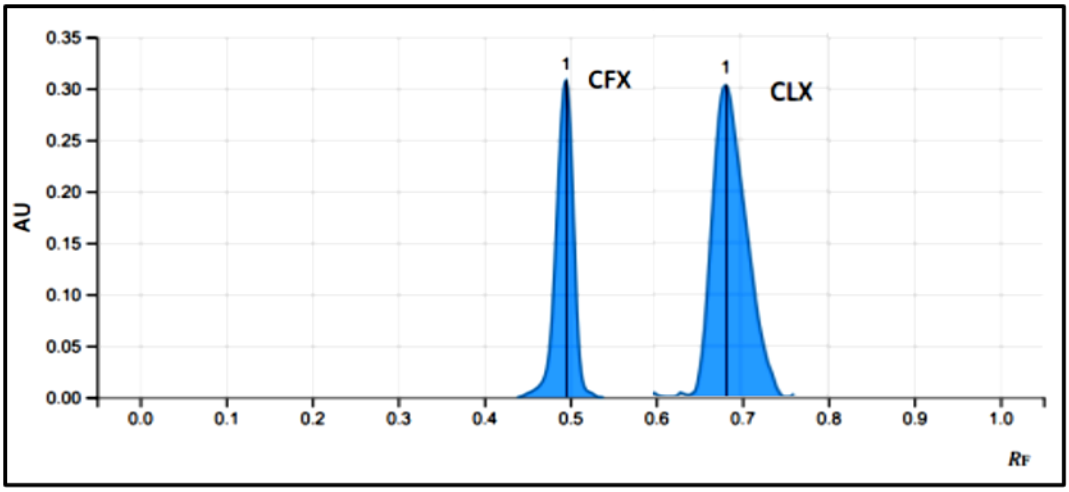


Figure 1: HPTLC Densitometric scan of Cefixime and Cloxacillin

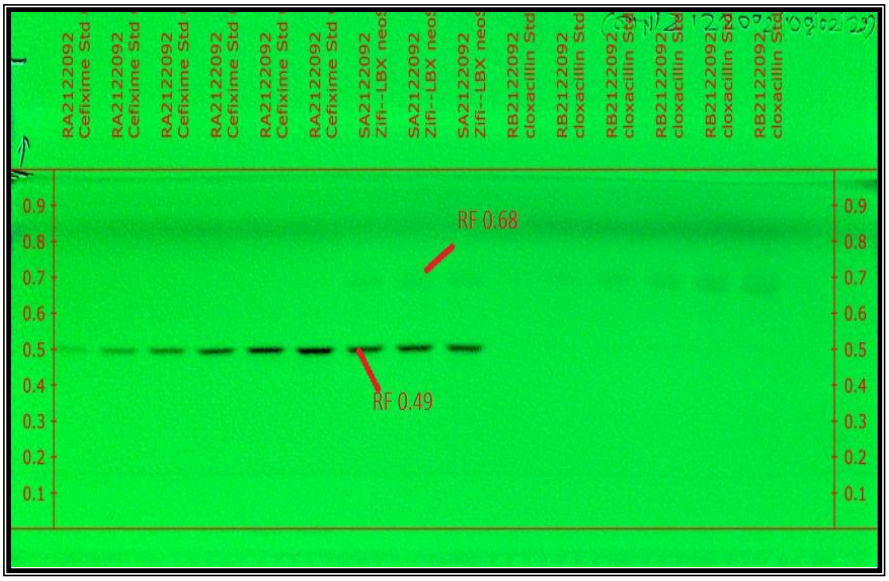


Figure 2: HPTLC plate image of Cefixime and Cloxacillin

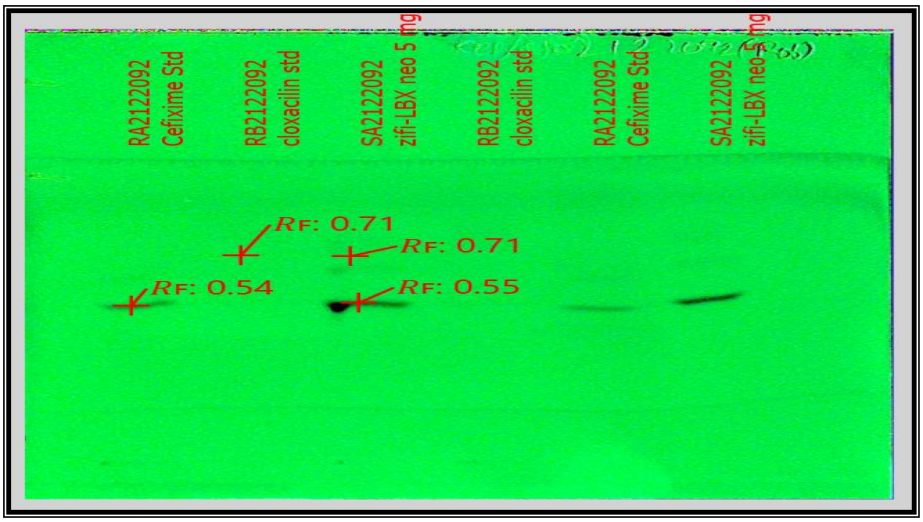
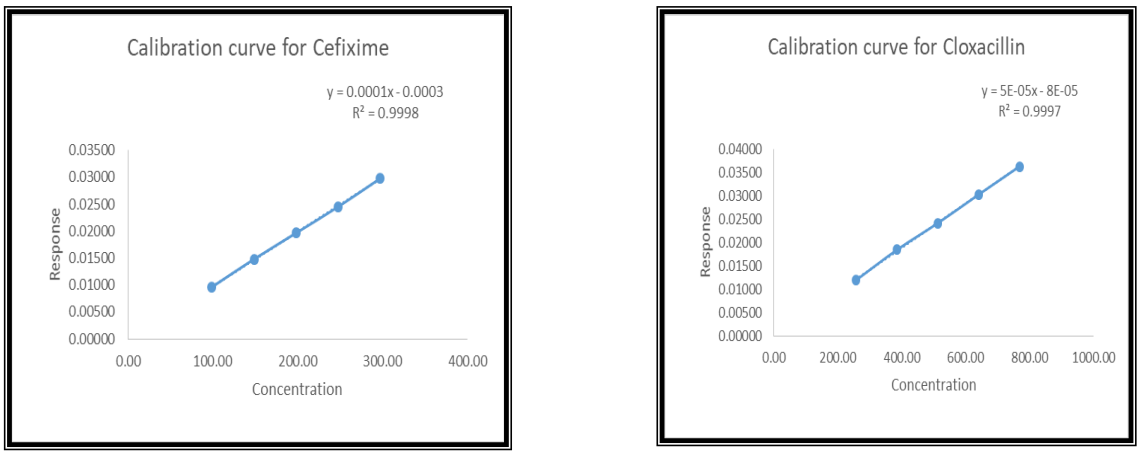


Figure 3: HPTLC plate image of Cefixime and Cloxacillin for Robustness study



X-axis- Concentration of Drug in µg/mL; **Y-axis** – Peak Area

Figure 4: Linearity graph for Cefixime Standard Cloxacillin Standard

References

1. Kumar, R. S., Kumar, N. P., Saravanakumar, S., Kandasamy, C. S., & Venkatnarayanan, R. (2009). Simultaneous RP-HPLC estimation of nitazoxanide and cloxacillin in tablet dosage forms. *Asian Journal of Research in Chemistry, 2*(1), 43–45.
2. Paramane, S., Kothapalli, L., Thomas, A., & Deshpande, A. D. (2007). Simultaneous RP-HPLC estimation of gemifloxacin and ornidazole in tablet dosage form. *Indian Journal of Pharmaceutical Sciences, 69*(4), 525–528.
3. Rege, P., Jagadale, A., & Bahuguni, N. (2024). Development and validation of RP-HPTLC method for simultaneous estimation of antibacterial drugs from pharmaceutical formulation. *International Journal of Pharmaceutical Sciences, 2*(7), 1636–1644.
4. Wikipedia contributors. (n.d.). *Cefixime*. Wikipedia. Retrieved May 30, 2026, from https://en.wikipedia.org/wiki/Cefixime
5. Wikipedia contributors. (n.d.). *Cloxacillin*. Wikipedia. Retrieved May 30, 2026, from https://en.wikipedia.org/wiki/Cloxacillin
6. International Conference on Harmonisation (ICH). (1994). *Q2A: Validation of analytical procedures: Methodology*. Geneva, Switzerland: ICH.
7. International Conference on Harmonisation (ICH). (1996). *Q2B: Validation of analytical procedures: Methodology*. Geneva, Switzerland: ICH.
8. Argekar, A. P., Kapadia, S. U., & Raj, S. V. (1996). Simultaneous determination of norfloxacin and tinidazole in pharmaceutical preparations by high-performance thin-layer chromatography. *Journal of Planar Chromatography–Modern TLC, 9*(3), 208–211.
9. Ganhimathi, M., Ravi, T. K., & Shukla, N. (2006). Validated high-performance thin-layer chromatography method for simultaneous estimation of cloxacillin and ornidazole. *Indian Journal of Pharmaceutical Sciences, 68*, 838–840.
10. Rashmin, P., Mrunali, P., Nitin, D., Nidhi, D., & Phara, P. T. (2012). HPTLC method development and validation: Strategy to minimize methodological failures. *Journal of Food and Drug Analysis, 20*, 794–804.
11. Bhole, R. P., Shinde, S. S., Chitlange, S. S., & Wankhede, S. B. (2015). High-performance thin-layer chromatographic method for simultaneous determination of diphenhydramine hydrochloride and naproxen sodium in tablets. *Analytical Chemistry Insights, 10*, 47–51.
12. Mahajan, K. S., Chauhan, R. S., Shah, S. A., & Shah, D. R. (2015). Development of HPTLC method for simultaneous estimation of paracetamol and flupirtine maleate in their combined tablet dosage form. *Journal of Pharmaceutical and Applied Sciences, 2*(1), 1–7.
13. Shah, D., Bhatt, K., Mehta, R., Baldania, S., & Gandhi, T. (2006). RP-HPLC determination of atorvastatin calcium and amlodipine besylate combination in tablet dosage forms. *Indian Journal of Pharmaceutical Sciences, 68*, 796–799.

Bhumi Publishing, India
May 2026

DIABETES COMPLICATIONS AND MANAGEMENT IN COVID-19 PATIENTS

Josef Yakin*, Faruk Alam and Injamul Islam

Faculty of Pharmaceutical Science,

Assam down town University, Panikhaiti, Narengi, Guwahati, Assam, India, 781026

*Corresponding author E-mail: josef.yakin@adtu.in

Introduction

On March 11, 2020, the World Health Organization declared COVID-19, caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), as a global pandemic [WHO (2020)]. COVID-19 is a highly transmissible infectious disease associated with the novel coronavirus SARS-CoV-2. The outbreak was first identified in Wuhan, China, in December 2019, and subsequently spread rapidly across the globe, becoming a major public health crisis. In India, a nation with a population exceeding 1.34 billion, substantial efforts were undertaken to control the transmission of the virus among the population. By May 8, 2020, approximately 56,342 confirmed positive cases had been reported in the country [HealthMap (2020)].

Diabetes mellitus is a chronic metabolic disease that today affects422 million people worldwide. Diabetes mellitus (DM), a common endocrine disorder, is characterised as a chronic, low-grade inflammatory disease triggered by a chronic immune system imbalance, metabolic syndrome, or an abundance of nutrients [Cariou *et al.* (2020)]. After the COVID-19 epidemic, numerous researches were conducted to look into the relationship between diabetes and the virus. Despite having small sample sizes, they were ambiguous about the connection between diabetes and COVID-19 mortality and worse clinical outcomes. The results are expected to affect the clinical care provided to patients who have both diseases and COVID-19 [Metwally *et al.* (2021)].

Acute respiratory distress syndrome, admission to an intensive care unit, the requirement for mechanical ventilation, and a higher risk of death are all effects of dangerous infection-induced complications that are linked to diabetes, according to the first reports that came out of Wuhan, China. This is true despite the rapid evolution of the COVID-19 epidemic [Chow *et al.* (2020)]. Diabetes does not increase the risk of COVID, but rather the condition is much more severe and seems to progress more quickly in those with diabetes. Patients with type 1 diabetes may do better since they are younger, although both type 2 and type 1 diabetes have a higher propensity for more serious disease [Chen *et al.* (2020)]. Diabetes causes increased inflammation in the body, and because COVID causes this inflammatory condition to worsen much faster, that could be one of the causes. The reason is that diabetics may be more prone to circulatory problems [Guo *et al.* (2020)].

Impact on Hyperglycemia

SARS-CoV-2, the virus responsible for COVID-19, enters host cells through interaction with the angiotensin-converting enzyme 2 receptor (ACE2r). This process is facilitated by the host transmembrane protease serine 2 (TMPRSS2), which activates the viral spike (S) protein and promotes viral entry into target cells [Hoffmann *et al.* (2020)]. Alterations associated with hyperglycemia, particularly abnormal glycosylation of both ACE2r and the viral spike protein, may contribute to increased disease severity in COVID-19 patients. These modifications can enhance the binding affinity between the viral spike protein and ACE2r and may also influence the host immune response against the virus [Brufsky (2020)].

The angiotensin-converting enzyme 2 receptor (ACE2r) is widely distributed in several tissues of the human body, including type I and type II alveolar epithelial cells of the lungs and upper respiratory tract, as well as the heart, vascular endothelium, renal tubular epithelium, intestinal lining, and pancreatic tissue. In individuals with diabetes, persistent hyperglycemia may negatively influence pulmonary function and aggravate respiratory complications associated with viral infections. Experimental studies using animal models have demonstrated that diabetes can induce multiple structural abnormalities in lung tissue, such as enhanced vascular permeability and damage or collapse of the alveolar epithelium [Yang *et al.* (2010)].

Considerations from a Therapeutic Point of View

Metformin

Metformin has been proven to have anti-inflammatory characteristics and to reduce circulating signs of inflammation in people with T2D, and it has also recently been demonstrated to reduce the risk of death [Cameron *et al.* (2016)]. The mortality rate for those who tested positive for COVID-19 and were on metformin was 11%, which was similar to the whole COVID-19 group and significantly lower than the mortality rate of 24% observed in diabetics not using metformin. Participants with T2D who took metformin had a lower chance of passing away than those who did not, even after controlling for insulin use, age, race, sex, obesity, and hypertension [Crouse *et al.* (2020)].

Dipeptidyl Peptidase-4 (Dpp4)

Considerable attention has recently been directed toward the interaction between coronaviruses and dipeptidyl peptidase-4 (DPP4), a cellular type-II transmembrane protein also referred to as cluster of differentiation 26 (CD26) or adenosine deaminase complexing protein 2. DPP4 serves as the functional receptor for Middle East respiratory syndrome coronavirus (MERS-CoV), in a manner comparable to the role of ACE2 in SARS-CoV and SARS-CoV-2 infections [Raj *et al.* (2013)]. Furthermore, DPP4 has been implicated in the promotion of inflammatory responses in individuals with type 2 diabetes through both catalytic and non-catalytic pathways. These effects

may alter the activity of several cytokines, chemokines, and growth factors involved in immune regulation and inflammation [Iacobellis (2020)].

Insulin

Insulin possesses significant anti-inflammatory effects and has been shown to reduce inflammatory biomarkers in critically ill hospitalized patients. Among the therapeutic options available for managing acute diabetes-related complications, insulin remains one of the most commonly administered treatments in patients suffering from bacterial or viral infections, as well as in critically ill individuals requiring hospital care. In the context of COVID-19, many hospitalized patients, especially those experiencing respiratory complications, may require insulin therapy for effective glycemic management [Drucker (2020)]. Recent clinical observations involving 59 COVID-19 patients further suggested that continuous insulin infusion was successful in achieving glycemic control and was associated with improved clinical outcomes [Sardu *et al.* (2020)].

GLP-1R Agonists

Glucagon-like peptide-1 receptor (GLP-1R) agonists have demonstrated extensive anti-inflammatory activity in experimental animal models and are known to reduce systemic inflammatory biomarkers in patients with type 2 diabetes (T2D) and obesity. Preclinical investigations in mice and rats with experimentally induced lung injury have shown that GLP-1R agonists can decrease pulmonary inflammation, suppress cytokine production, and help maintain normal lung function. Severe COVID-19 infection is frequently associated with acute respiratory distress syndrome (ARDS), a condition characterized by excessive release of inflammatory cytokines commonly referred to as a “cytokine storm” [Hansen *et al.* (2003)]. These findings suggest that GLP-1-based therapies may have potential benefits in the management of COVID-19 patients, irrespective of diabetic status. However, despite evidence from short-term studies indicating that GLP-1 can safely improve glycemic control in mechanically ventilated critically ill patients, sufficient clinical data regarding the safety and efficacy of GLP-1R agonists in critically ill COVID-19 patients remain limited. Therefore, definitive recommendations for their use in the treatment of SARS-CoV-2 infection cannot yet be established [Jin *et al.* (2020)].

A.C.E. Inhibitors/A.R.B.S

A recent clinical investigation reported that the administration of angiotensin-converting enzyme (ACE) inhibitors in patients with COVID-19 was associated with reduced mortality following hospital discharge, whereas no significant association was observed with the use of angiotensin receptor blockers (ARBs). The study further indicated that factors such as race, ethnicity, immunosuppression, hyperlipidemia, and diabetes mellitus were not independently linked to an increased risk of in-hospital mortality among the evaluated patients [Mandeep *et al.* (2020)].

Statins

Beyond their cholesterol-lowering action, statins exhibit several pleiotropic effects that contribute to beneficial outcomes in cardiovascular diseases, particularly through the modulation of inflammation and oxidative stress. These agents influence multiple immune-related mechanisms, including cytokine production, antigen presentation, and the adhesion and migration of immune cells. In patients with COVID-19, statins may also help manage hyperlipidemia associated with the use of protease inhibitor-based antiviral therapies and immunosuppressive medications. Simvastatin and, to a lesser extent, atorvastatin are primarily metabolized by the hepatic cytochrome P450 isoenzyme CYP3A4. Consequently, cautious use of these statins is recommended, beginning with lower doses and accompanied by regular monitoring of creatine kinase and liver transaminase levels to minimize the risk of adverse effects [Vuorio & Kovanen (2021)].

Complication

The biggest cause of illness and mortality in the world is diabetes mellitus (DM). This disease has a multitude of consequences, all of which have an effect on how severe the condition is. According to study, people who have been exposed to infectious viruses like MERS-CoV, H1N1 (flu), and SARS-CoV are more likely to die and have more severe infections. According to multiple research projects, COVID-19 has been related to significant and life-threatening illness in people with diabetes and cardiovascular disease [Banik *et al.* (2016); Onder *et al.* (2020); Istituto Superiore di Sanità (2023)].

Diabetes as a Risk Factor for Worst Results

Individuals with diabetes are generally at a higher risk of developing severe complications following infectious diseases, especially viral infections. The variability in clinical outcomes is influenced by several factors, including viral load, host immune response, patient age, and the presence of underlying comorbid conditions. Previous coronavirus outbreaks, such as severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), also identified diabetes as a significant risk factor associated with increased morbidity and mortality. Type 2 diabetes is commonly associated with chronic low-grade inflammation resulting from excessive visceral adipose tissue accumulation. This persistent inflammatory state can impair insulin sensitivity and disrupt normal glucose homeostasis. Furthermore, chronic hyperglycemia, ongoing inflammation, and autoimmune dysregulation may contribute to an impaired and inadequate immune response in diabetic individuals [Iacobellis (2020)].

COVID-19 and Glucose Metabolism

Elevated glucose concentrations in human monocytes have been shown to directly enhance SARS-CoV-2 replication. Glycolytic metabolism supports viral replication through the generation of mitochondrial reactive oxygen species (ROS) and activation of hypoxia-inducible

factor-1 (HIF-1) signaling pathways [Codo *et al.* (2020)]. This observation is consistent with earlier findings demonstrating that hyperglycemia, as well as a previous history of type 1 diabetes mellitus (T1DM) or type 2 diabetes mellitus (T2DM), independently increased the risk of morbidity and mortality in patients affected by SARS. In addition, studies involving MERS-CoV-infected mice revealed that coexisting T2DM induced a dysregulated immune response, resulting in severe and widespread pulmonary injury [Kulcsar *et al.* (2019)]. Individuals with diabetes mellitus are therefore considered more vulnerable to severe SARS-CoV-2 infection compared with non-diabetic individuals. Poor glycemic control has also been associated with increased medication requirements, prolonged hospitalization, and elevated mortality rates. Furthermore, COVID-19 itself may worsen glycemic status in individuals with impaired glucose regulation or pre-existing diabetes mellitus [Carey *et al.* (2020)].

Inflammation and Insulin Resistance

Post-mortem examinations of patients who died from COVID-19 have frequently revealed diffuse alveolar damage accompanied by extensive inflammatory cell infiltration and prominent hyaline membrane formation within the lungs. Additional pathological findings reported in severe cases include localized pancreatitis, lymphocytic infiltration of the liver, macrophage accumulation in the brain, axonal injury, glomerular microthrombi, and inflammation of myocardial tissue. Furthermore, recent evidence suggests that fatal COVID-19 pneumonia may be associated with inherited defects in type I interferon-mediated immunity involving pathways linked to Toll-like receptor 3 (TLR3), interferon regulatory factor 7 (IRF7), or B-cell immune responses. These genetic abnormalities have been identified in approximately 12.5% of affected men and 2.6% of affected women [Hadjadj *et al.* (2020)].

A subset of patients with severe COVID-19 experiences an excessive inflammatory reaction commonly referred to as a “cytokine storm,” which may contribute to life-threatening complications and increased mortality. A retrospective study involving 317 laboratory-confirmed COVID-19 patients demonstrated that elevated inflammatory markers, particularly interleukin-6 (IL-6) and lactate dehydrogenase, detected within 24 hours of hospital admission, were strongly associated with disease severity [Zeng *et al.* (2020)]. This exaggerated inflammatory response may partly explain the rapid progression of COVID-19 observed in individuals with diabetes, as increased IL-6 levels can induce damage to DNA, proteins, lipids, and other cellular components, thereby impairing normal cellular structure and function. In addition, elevated plasma concentrations of IL-6 and extracellular newly identified receptor for advanced glycation end products-binding protein (EN-RAGE), a biomarker associated with lung injury and sepsis-induced acute respiratory distress syndrome (ARDS), were positively correlated with increased bacterial DNA and lipopolysaccharide levels in patients with severe COVID-19 [Libby & Simon (2020); Arunachalam *et al.* (2020)].

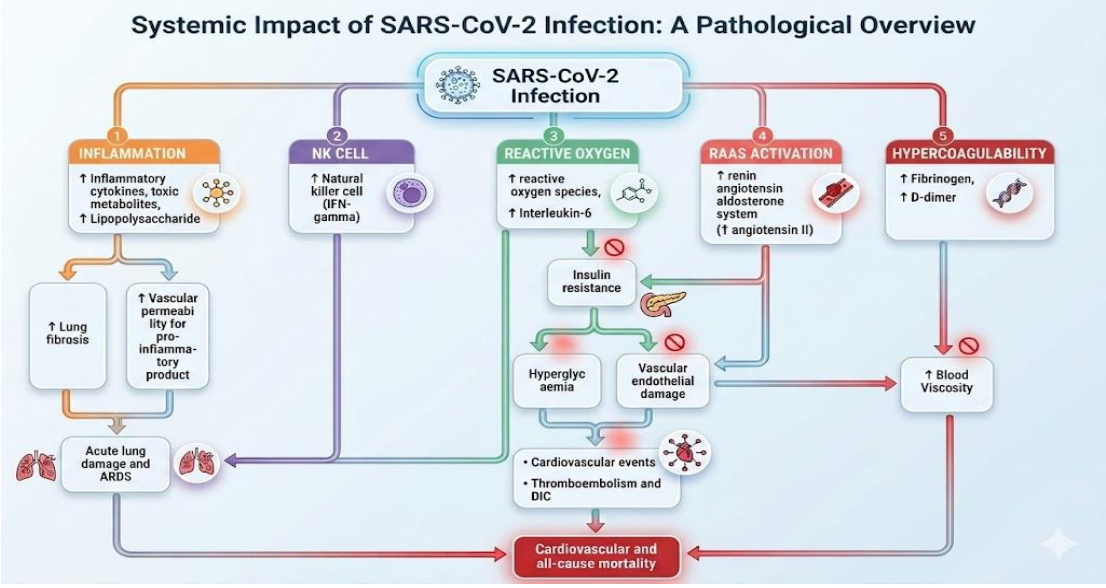


Figure 1: Potential pathogenic mechanisms in T2DM and COVID-19 patients.

Patients with type 2 diabetes mellitus (T2DM) exhibit multiple pathological mechanisms that may aggravate the severity of COVID-19. Infection with severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) promotes the release of elevated levels of inflammatory mediators, including cytokines, lipopolysaccharides, and toxic metabolic by-products, leading to an exaggerated inflammatory response. These processes are strongly associated with the development of pulmonary fibrosis, acute lung injury, and acute respiratory distress syndrome (ARDS) [Zhou *et al.* (2020)]. Furthermore, SARS-CoV-2 infection can enhance reactive oxygen species (ROS) production and activate the renin–angiotensin–aldosterone system (RAAS) through increased angiotensin II expression. Together with insulin resistance, hyperglycemia, and vascular endothelial dysfunction, these alterations contribute to cardiovascular complications, thromboembolic events, and disseminated intravascular coagulation (DIC) [Teuwen *et al.* (2020)]. Infection-associated increases in blood viscosity, endothelial injury, and elevated coagulation markers such as D-dimer and fibrinogen further intensify the risk of thrombosis and cardiovascular complications in affected individuals [Imai *et al.* (2019)].

Immunomodulation

Individuals with type 2 diabetes mellitus (T2DM) are particularly vulnerable to severe complications associated with COVID-19 due to several interconnected pathological mechanisms. SARS-CoV-2 infection stimulates the excessive production of inflammatory mediators, including cytokines, lipopolysaccharides, and harmful metabolic compounds, resulting in an amplified inflammatory response. Such inflammatory alterations are closely linked with pulmonary fibrosis, acute lung injury, and acute respiratory distress syndrome (ARDS) [Zhou *et al.* (2020)]. In addition, viral infection can increase the generation of reactive oxygen species (ROS) and activate the renin–angiotensin–aldosterone system (RAAS) by

elevating angiotensin II activity. Combined with insulin resistance, persistent hyperglycemia, and endothelial dysfunction, these factors significantly contribute to cardiovascular complications, thromboembolic disorders, and disseminated intravascular coagulation (DIC) [Teuwen *et al.* (2020)]. Moreover, infection-induced increases in blood viscosity, endothelial damage, and elevated coagulation markers such as D-dimer and fibrinogen further enhance the likelihood of thrombosis and other cardiovascular events in affected patients [Imai *et al.* (2019)].

Result and Discussion

Prevalence COVID-19 Risk Factors in India

Overall, 83.7% of adults in India are under 60 years old. Obesity and cancer rates are far lower in India (4.4% and 0.3 %, respectively), but uncontrolled diabetes, renal disease, and chronic liver disease are significantly higher (8.9%, 9.7% and 5.3% in India). India has higher total diabetes prevalence at all ages; however, diabetes in India is predominantly uncontrolled. Hypertension (the total of managed and uncontrolled) is more common in India at younger ages (31.3% for ages 40-49) but less common at older years (52.3% for ages 70-79) Table 1.

Table 1: Prevalence of COVID-19 Risk Factors in India

Risk Factor / Demographic Variable	**Prevalence (%)**
Age Group	
18-39 years	50.2
40-49 years	19.2
50-59 years	14.3
60-69 years	10.3
70-79 years	4.6
80-89 years	1.5
Sex	
Male	47.1
Metabolic Disorders	
Diabetes (controlled)	1.7
Diabetes (uncontrolled)	8.9
Hypertension	28.2
Obesity (Class I and II)	4.0
Obesity (Class III)	0.4
Cardiovascular and Respiratory Diseases	
Chronic heart disease	4.4
Chronic respiratory disease	4.8
Asthma	2.5

Renal and Hepatic Disorders	
Kidney disease	9.7
Chronic liver disease	5.3
Cancer and Neurological Disorders	
Haematological cancer	0.0
Non-haematological cancer	0.3
Stroke/Dementia	1.3
Other neurological conditions	0.0
Immune-related Disorders	
Psoriasis/Rheumatoid disorders	1.0
Other immunosuppressive conditions	0.1

Relative risk of COVID-19 mortality from combined risk factors in India

Table 2: Population Relative Risk of COVID-19 Mortality Associated with Different Health Conditions in India

Health Condition	Individual Relative Risk	Population Relative Risk in India (PRRc)
Diabetes (controlled)	1.31	1.004
Diabetes (uncontrolled)	1.94	1.078
Hypertension	0.89	0.971
Obesity (Class I and II)	1.15	1.006
Obesity (Class III)	1.91	1.004
Chronic heart disease	1.17	1.008
Chronic respiratory disease	1.62	1.035
Asthma	1.13	1.003
Kidney disease	1.42	1.050
Chronic liver disease	1.73	1.042
Haematological cancer	2.79	1.000
Non-haematological cancer	1.71	1.002
Stroke/Dementia	2.15	1.016
Other neurological conditions	2.56	1.002
Psoriasis/Rheumatoid disorders	1.19	1.002
Other immunosuppressive conditions	1.69	1.001

In India, the age-specific population relative risk associated with COVID-19 mortality has been reported to be elevated across nearly all age groups. The population relative risk of COVID-19 reflects both the age-related prevalence of the disease and the mortality risk linked to underlying

health conditions within the affected population. The overall population relative risk of COVID-19 mortality for each specific health condition (PRRc) is determined by integrating age-specific prevalence, relative risk, and population distribution across different age categories. This approach provides an estimate of the proportional increase in COVID-19-related mortality attributable to individual health conditions across the population (Table 2).

Treatment

For the effective management of COVID-19 patients with diabetes, physicians are advised to follow the 2019 recommendations issued by the Austrian Diabetes Society (Österreichische Diabetes Gesellschaft, DG). Poor glycemic control has consistently been associated with unfavorable clinical outcomes, including prolonged hospital stay, increased complication rates, and higher mortality. Antidiabetic therapy should therefore be initiated or intensified in patients with preprandial blood glucose concentrations exceeding 180 mg/dL (10 mmol/L). In outpatient settings, healthcare professionals should also consider the effects of social distancing and quarantine measures on glucose regulation in diabetic individuals. Reduced physical activity resulting from prolonged home confinement may negatively affect glycemic control and overall metabolic status.

Management of COVID-19 primarily involves infection prevention and control strategies along with supportive care measures, including supplemental oxygen therapy and ventilatory support when clinically required. In addition, Remdesivir has received approval from the U.S. Food and Drug Administration for the treatment of hospitalized COVID-19 patients aged above 12 years and weighing more than 40 kg. Early diagnosis and timely therapeutic intervention are important to prevent disease progression and reduce the likelihood of secondary complications and subsequent infections.

Conclusion

Patients with diabetes mellitus should be informed that COVID-19 can raise blood glucose levels, and as a result, they should adhere to clinical guidelines for diabetes mellitus management more closely during the COVID-19 pandemic. Diabetes is one of the most frequent diseases and the major cause of a slew of costly consequences; when it strikes young people, it can drive them out of work. Furthermore, COVID-19 disease is a unique respiratory ailment that has spread around the world, infecting about 2.9 million individuals and killing around 202,000 people.

References

1. World Health Organization. (2020). *Novel coronavirus (2019-nCoV): Situation report*. World Health Organization.
2. HealthMap. (2020). *Novel coronavirus (COVID-19)*. Retrieved June 7, 2020, from HealthMap COVID-19 Tracker.

3. Cariou, B., Hadjadj, S., Wargny, M., Pichelin, M., Al-Salameh, A., Allix, I., Amadou, C., Arnault, G., Baudoux, F., Bauduceau, B., Borot, S., Bourgeon-Ghittori, M., Bourron, O., Boutoille, D., Cazenave-Roblot, F., Chaumeil, C., Cosson, E., Ducet-Boiffard, A., Ducluzeau, P. H., *et al.* (2020). Phenotypic characteristics and prognosis of inpatients with COVID-19 and diabetes: The CORONADO study. *Diabetologia, 63*(8), 1500–1515. https://doi.org/10.1007/s00125-020-05180-x
4. Metwally, A. A., Mehta, P., Johnson, B. S., Nagarjuna, A., & Snyder, M. P. (2021). COVID-19-induced new-onset diabetes: Trends and technologies. *Diabetes, 70*(12), 2733–2744. https://doi.org/10.2337/dbi21-0029
5. Chow, N., Fleming-Dutra, K., Gierke, R., Hall, A., Hughes, M., Pilishvili, T., Ritchey, M., Roguski, K., Skoff, T., Ussery, E., *et al.* (2020). Preliminary estimates of the prevalence of selected underlying health conditions among patients with coronavirus disease 2019—United States, February 12–March 28, 2020. *MMWR Morbidity and Mortality Weekly Report, 69*(13), 382–386. https://doi.org/10.15585/mmwr.mm6913e2
6. Chen, Q., Zheng, Z., Zhang, C., Zhang, X., Wu, H., Wang, J., Wang, S., Zheng, C., & Zhang, Y. (2020). Clinical characteristics of 145 patients with coronavirus disease 2019 (COVID-19) in Taizhou, Zhejiang, China. *Infection, 48*(4), 543–551. https://doi.org/10.1007/s15010-020-01432-5
7. Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., Qin, R., Wang, H., Shen, Y., Du, K., Zhao, L., Fan, H., Luo, S., & Hu, D. (2020). Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes/Metabolism Research and Reviews, 36*(7), e3319. https://doi.org/10.1002/dmrr.3319
8. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell, 181*(2), 271–280.e8. https://doi.org/10.1016/j.cell.2020.02.052
9. Brufsky, A. (2020). Hyperglycemia, hydroxychloroquine, and the COVID-19 pandemic. *Journal of Medical Virology, 92*(7), 770–775. https://doi.org/10.1002/jmv.25887
10. Yang, J. K., Lin, S. S., Ji, X. J., & Guo, L. M. (2010). Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes. *Acta Diabetologica, 47*(3), 193–199. https://doi.org/10.1007/s00592-009-0109-4
11. Cameron, A. R., Morrison, V. L., Levin, D., Mohan, M., Forteath, C., Beall, C., McNeilly, A. D., Balfour, D. J. K., Savinko, T., Wong, A. K. F., *et al.* (2016). Anti-inflammatory effects of metformin irrespective of diabetes status. *Circulation Research, 119*(5), 652–665. https://doi.org/10.1161/CIRCRESAHA.116.308445

12. Crouse, A. B., Grimes, T., Li, P., Might, M., Ovalle, F., & Shalev, A. (2020). Metformin use is associated with reduced mortality in a diverse population with COVID-19 and diabetes. *medRxiv*. https://doi.org/10.1101/2020.07.29.20164020
13. Raj, V. S., Mou, H., Smits, S. L., Dekkers, D. H. W., Müller, M. A., Dijkman, R., Muth, D., Demmers, J. A. A., Zaki, A., Fouchier, R. A. M., *et al.* (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature, 495*(7440), 251–254. https://doi.org/10.1038/nature12005
14. Drucker, D. J. (2020). Coronavirus infections and type 2 diabetes—Shared pathways with therapeutic implications. *Endocrine Reviews, 41*(3), bnaa011. https://doi.org/10.1210/endrev/bnaa011
15. Sardu, C., D'Onofrio, N., Balestrieri, M. L., Barbieri, M., Rizzo, M. R., Messina, V., Maggi, P., Coppola, N., Paolisso, G., & Marfella, R. (2020). Outcomes in patients with hyperglycemia affected by COVID-19: Can we do more on glycemic control? *Diabetes Care, 43*(7), 1408–1415. https://doi.org/10.2337/dc20-0723
16. Hansen, T. K., Thiel, S., Wouters, P. J., Christiansen, J. S., & Van den Berghe, G. (2003). Intensive insulin therapy exerts anti-inflammatory effects in critically ill patients and counteracts the adverse effect of low mannose-binding lectin levels. *The Journal of Clinical Endocrinology & Metabolism, 88*(3), 1082–1088. https://doi.org/10.1210/jc.2002-021478
17. Jin, T., & Liu, M. (2020). Letter to the editor: Comment on GLP-1-based drugs and COVID-19 treatment. *Acta Pharmaceutica Sinica B, 10*(7), 1249–1250. https://doi.org/10.1016/j.apsb.2020.05.006
18. Mehra, M. R., Desai, S. S., Kuy, S., Henry, T. D., & Patel, A. N. (2020). Cardiovascular disease, drug therapy, and mortality in COVID-19. *The New England Journal of Medicine, 382*(25), e102. https://doi.org/10.1056/NEJMoa2007621
19. Vuorio, A., & Kovanen, P. T. (2021). Statins as adjuvant therapy for COVID-19 to calm the stormy immunothrombosis and beyond. *Frontiers in Pharmacology, 11*, 579548. https://doi.org/10.3389/fphar.2020.579548
20. Banik, G. R., Alqahtani, A. S., Booy, R., & Rashid, H. (2016). Risk factors for severity and mortality in patients with MERS-CoV: Analysis of publicly available data from Saudi Arabia. *Virologica Sinica, 31*(1), 81–84. https://doi.org/10.1007/s12250-015-3679-z
21. Onder, G., Rezza, G., & Brusaferro, S. (2020). Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *JAMA, 323*(18), 1775–1776. https://doi.org/10.1001/jama.2020.4683
22. Istituto Superiore di Sanità. (2023). *Characteristics of COVID-19 patients dying in Italy*. EpiCentro. Retrieved January 5, 2023, from ISS EpiCentro COVID-19 mortality report.

23. Iacobellis, G. (2020). COVID-19 and diabetes: Can DPP4 inhibition play a role? *Diabetes Research and Clinical Practice, 162*, 108125. https://doi.org/10.1016/j.diabres.2020.108125
24. Codo, A. C., Davanzo, G. G., Monteiro, L. B., de Souza, G. F., Muraro, S. P., Virgilio-da-Silva, J. V., Prodonoff, J. S., Carregari, V. C., de Biagi Junior, C. A. O., Crunfli, F., *et al.* (2020). Elevated glucose levels favor SARS-CoV-2 infection and monocyte response through a HIF-1α/glycolysis-dependent axis. *Cell Metabolism, 32*(3), 437–446.e5. https://doi.org/10.1016/j.cmet.2020.07.007
25. Kulcsar, K. A., Coleman, C. M., Beck, S. E., & Frieman, M. B. (2019). Comorbid diabetes results in immune dysregulation and enhanced disease severity following MERS-CoV infection. *JCI Insight, 4*(20), e131774. https://doi.org/10.1172/jci.insight.131774
26. Carey, I. M., Critchley, J. A., DeWilde, S., Harris, T., Hosking, F. J., & Cook, D. G. (2018). Risk of infection in type 1 and type 2 diabetes compared with the general population: A matched cohort study. *Diabetes Care, 41*(3), 513–521. https://doi.org/10.2337/dc17-2131
27. Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Péré, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., *et al.* (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science, 369*(6504), 718–724. https://doi.org/10.1126/science.abc6027
28. Zeng, Z., Yu, H., Chen, H., Qi, W., Chen, L., Chen, G., Yan, W., Chen, T., Ning, Q., Han, M., & Wu, D. (2020). Longitudinal changes of inflammatory parameters and their correlation with disease severity and outcomes in patients with COVID-19 from Wuhan, China. *Critical Care, 24*(1), 525. https://doi.org/10.1186/s13054-020-03255-0
29. Libby, P., & Simon, D. I. (2001). Inflammation and thrombosis: The clot thickens. *Circulation, 103*(13), 1718–1720. https://doi.org/10.1161/01.CIR.103.13.1718
30. Arunachalam, P. S., Wimmers, F., Mok, C. K. P., Perera, R. A. P. M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., *et al.* (2020). Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science, 369*(6508), 1210–1220. https://doi.org/10.1126/science.abc6261
31. Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., *et al.* (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet, 395*(10229), 1054–1062. https://doi.org/10.1016/S0140-6736(20)30566-3
32. Teuwen, L. A., Geldhof, V., Pasut, A., & Carmeliet, P. (2020). COVID-19: The vasculature unleashed. *Nature Reviews Immunology, 20*(7), 389–391. https://doi.org/10.1038/s41577-020-0343-0

33. Imai, Y., Kuba, K., Neely, G. G., Yaghubian-Malhami, R., Perkmann, T., van Loo, G., Ermolaeva, M., Veldhuizen, R., Leung, Y. H., Wang, H., *et al.* (2008). Identification of oxidative stress and Toll-like receptor 4 signaling as a key pathway of acute lung injury. *Cell, 133*(2), 235–249. https://doi.org/10.1016/j.cell.2008.02.043
34. Zhang, Y., Xiao, M., Zhang, S., Xia, P., Cao, W., Jiang, W., Chen, H., Ding, X., Zhao, H., Zhang, H., *et al.* (2020). Coagulopathy and antiphospholipid antibodies in patients with Covid-19. *The New England Journal of Medicine, 382*(17), e38. https://doi.org/10.1056/NEJMc2007575
35. Šestan, M., Marinović, S., Kavazović, I., Cekinović, Đ., Wueest, S., Turk Wensveen, T., Brizić, I., Jonjić, S., Konrad, D., Wensveen, F. M., & Polić, B. (2018). Virus-induced interferon-γ causes insulin resistance in skeletal muscle and derails glycemic control in obesity. *Immunity, 49*(1), 164–177.e6. https://doi.org/10.1016/j.immuni.2018.05.005
36. Wensveen, F. M., Jelenčić, V., Valentić, S., Šestan, M., Turk Wensveen, T., Theurich, S., Glasner, A., Mendrila, D., Štimac, D., Wunderlich, F. T., *et al.* (2015). NK cells link obesity-induced adipose stress to inflammation and insulin resistance. *Nature Immunology, 16*(4), 376–385. https://doi.org/10.1038/ni.3120
37. Kim, J. H., Kim, D. J., Lee, K. W., Lee, M., Kim, K. U., Kim, S. O., & Cho, Y. K. (2019). Relationship between natural killer cell activity and glucose control in patients with type 2 diabetes and prediabetes. *Journal of Diabetes Investigation, 10*(5), 1223–1228. https://doi.org/10.1111/jdi.13002

IN-VITRO ANTIMICROBIAL ACTIVITY OF ROOT EXTRACTS AGAINST HUMAN PATHOGENIC MICROORGANISMS

Mutyala Naidu Lagudu
Department of Botany,
Adikavi Nannaya University, Rajamahendravaram 533296,
East Godavari, Andhra Pradesh, India
Corresponding author E-mail: lagudu3@gmail.com, lmn.bot@aknu.edu.in

Abstract

Plants have been an essential part of human society since the start of civilization. Around 250 drugs have been identified from plants during Rig Veda and Atharvana Veda descriptions of the Veda period. In the present work, aqueous as well as organic solvent extracts of various medicinal plants of different plant parts have been investigated for their antimicrobial activities against pathogenic microorganisms. Sequential extraction of it was done using hexane, chloroform, followed by methanol and ethanol. The agar well diffusion technique was followed for antimicrobial susceptibility test for crude extracts and DMSO (negative control) whereas agar disc diffusion method was followed for antimicrobial susceptibility test for standard antibiotic disc. On comparing root extracts of aqueous and solvents (hexane, chloroform and methanol), methanol extract obtained from *C. zylanica* and *C. swietinea* exhibited highly significant zone of inhibition against all the tested organisms ranging from 18 – 28 mm whereas chloroform extract exhibited lower zone of inhibition ranging from 20 -22 mm against *M. letues, A. niger* and *S. Cerevisiae.* From the above results it can be concluded that the plant extracts have great potential as antimicrobial compounds against microorganisms and that they can be used in the treatment of infectious diseases caused by resistant microorganisms.

Keywords: Medicinal Plants, Root Extracts, Microorganisms, Inhibition Zones.

Introduction

Plants have been an essential part of human society since the start of civilization. Around 250 drugs have been identified from plants during Rig Veda and Atharvana Veda descriptions of the Veda period. The rural/tribal population in different part of the world is more disposed to traditional ways of treatment because of easy availability and cheaper cost. It is estimated about 75-80 % of the whole population mainly in developing countries, for primary health care because of better cultural acceptability, better compatibility with the human body and fewer side effects. In 15th century A.D., Hippocrates (the feature of medical science) mentioned 300-400 medicinal plants. In the first century A.D., Disscorides wrote “De material medica” a medicinal plant catalog which become the protype for modern pharmacopoeias. The Bible offers descriptions of approximately 30 healing plants. In recent years there has been a gradual revival of interest in the

use of medicinal plants in developed as well as in developing countries, because herbal medicines have been reported to be safe and without any diverse side effect. Thus, searches for new drug with better and cheaper substitutes, plant resources are a natural choice.

In India, the uses of different parts of several medicinal plants or their extracts are used in treatment of various diseases (Balakrishnan *et al.,* 2006). Many plant extracts have been used as a source of medicinal agent to cure urinary tract infection, cervictis, vaginitis, gastrointestinal disorder, respiratory disease, cutaneous affections, helminthic infections, parasitic protozoan disease and inflammatory processes. Although extremely effective, antibiotics are able to induce resistance has been the main factor for the increase of morbidity, mortality and health care costs of bacterial infection. *In vivo* tests of each plant extract against malarial and bacterial infection would have been time consuming and expensive without obtaining the desired knowledge.

Though, the antimicrobial properties of plants have been investigating time to time by a number of researchers worldwide, but recently it has gained much importance globally after the development of molecular biology. Thus, in the present work, aqueous as well as organic solvent extracts of various medicinal plants of different plant parts have been investigated for their antimicrobial activities against various microbes.

2. Material and Methods

Collection, Identification and Extraction

Capparis zeylanica and *Chlorozylon swietenia* roots were collected from Adikavi Nannaya University, Rajamahendravaram, Andhra Pradesh, India. The collected sample of plant dried in the shadow until it gets dried completely. Then it was powdered in the mixture grinder and packed in Soxhlet apparatus. Sequential extraction of it was done using hexane, chloroform, followed by methanol and ethanol (Aniel Kumar *et al.,* 2015). The filtrates were concentrated by removing the solvents under reduced pressure at 40 ^{0}C using a rotary evaporator. The concentrated crude extracts were labeled and stored at 4 ^{0}C. Simultaneously, the aqueous extract of the whole plant was prepared by adding boiled water to the powdered in a beaker on water bath, with occasional stirring for 4 hrs. The aqueous extract was then filtered and reduced under pressure.

Strains Used

The following strains were collected from Microbial type culture and collection (MTCC), Chandigarh, India. Seven bacterial strains namely *Bacillus subtilis* MTCC B2274, *Enterococcus faecalis* MTCC B3159, *Escherichia coli* MTCC B1560, *Klebsiella pneumoniae* MTCC B4030, *Micrococcus luteus* MTCC B1538, *Pseudomonas aeruginosa* MTCC B2297, *Proteus vulgaris* MTCC B7299, *Staphylococcus aureus* MTCC B3160, *Streptococcus pneumoniae* MTCC B2672, and three fungal strains such as *Aspergillus niger* MTCC F4325, *Candida albicans* MTCC F7315 and *Saccharomyces cerevisiae* MTCC F2567.

Antimicrobial Screening

The lyophilized culture was sub cultured and concentration of working stock culture was assessed as 10^{-6} CFU/ml. Specified quantity of Nutrient agar was prepared and plated in aseptic condition. The agar well diffusion technique was followed for antimicrobial susceptibility test for crude extracts and DMSO (negative control) whereas agar disc diffusion method was followed for antimicrobial susceptibility test for standard antibiotic disc. The extracts were dissolved in DMSO to get the known concentrations of 50 mg/ml, 75 mg/ml and 100 mg/ml. The activity was compared with tetracycline disc (10 mcg/disc). After 24 h of incubation at 37 ^{0}C the zone of inhibition was measured using an antibiotic zone reader scale (HiAntibiotic ZoneScale-c) and tabulated. For the antifungal activity, the same method as for bacteria was adopted of nutrient agar, Saboraud dextrose agar was used. The inoculated medium was incubated at 25 ^{0}C for two days for the *C. albicans*, *S. cerevisiae* and three days for *A. niger* (Aniel Kumar *et al.,* 2014). About 500 µg of nystatin was dissolved in 1 ml of sterile de ionized water. About 10 µl of 0.5 mg/ml nystatin (equivalent to 5 µg dose).

The extracts that exhibited inhibition zones were subjected to minimum inhibition concentration (MIC) assay by using serial two-fold dilution (Aniel and Mutyala, 2016). A quantity of 0.6 g of each extract was dissolved in 300 ml sterile nutrient broth which yields initial concentration of 2000 µg/ml. Subsequently, two-fold serial dilution was made from the stock to obtain 1000, 500, 250, 125, 62.5, 31.2 µg/ml concentrations. One ml of standardized inoculums of each test organism was introduced into each extract nutrient broth mixture and then incubated at 37 ^{0}C. The lowest concentration inhibiting growth was regarded as the MIC of the extracts. For the fungi, the inoculated medium was incubated at 25 ^{0}C for two (*C. albicans*, *S. cerevisiae*) to three (*A. niger*) days.

Statistical Analysis

Each experimental data from triplicates was subjected to one way ANOVA using Minitab version 15. A significant level of $p < 0.01$ was used for all statistical analyses.

Results and Discussion

The antimicrobial activity of root extracts in aqueous and solvents (hexane, chloroform and methanol) obtained from two selected medicinal plants Viz. *Capparis zeylanica* and *Chlorozylon swietenia* against the 12 selected microorganisms examined in the present study were assessed by the diameter of zone of inhibition. These microorganisms have been tested with commercially available two different antibiotics and results were given in tables 1 and 2. On comparing root extracts of aqueous and solvents (hexane, chloroform and methanol), methanol extract obtained from *C. zylanica* and *C. swietinea* exhibited highly significant zone of inhibition against all the tested organisms ranging from 18 – 28 mm whereas chloroform extract exhibited lower zone of inhibition ranging from 20 -22 mm against *M. letues, A. niger* and *S. cerevisiae*.

Table 1: Antimicrobial activity of root of *Capparis zeyanica* L.

Organisms	Zone of inhibition (mm)[a]												S	D
	Hexane extract (mg/ml)			Chloroform extract (mg/ml)			Methanol extract (mg/ml)			Aqueous extract (mg/ml)				
	100	200	300	100	200	300	100	200	300	100	200	300		
B. subtilis	–	–	–	–	–	–	13±0.90	15±0.28	18±0.90	16±0.76	18±0.16	20±0.64	18[T]	–
E. faecalis	–	–	–	10±0.40	12±0.22	14±0.44	13±0.28	15±0.52	17±0.22	–	–	–	21[T]	–
M. letus	10±0.10	12±0.76	14±0.76	12±0.90	14±0.40	16±0.22	16±0.45	18±0.10	20±0.45	–	–	–	24[T]	–
S. aureus	–	–	–	–	–	–	14±0.22	16±0.28	18±0.16	–	–	–	24[T]	–
S. pneumoniae	–	–	–	–	–	–	10±0.45	12±0.19	14±0.40	–	–	–	22[T]	–
E. coli	–	–	–	–	–	–	13±0.28	15±0.22	17±0.40	–	–	–	22[T]	–
K. pneumoniae	10±0.16	12±0.10	14±0.20	12±0.19	14±0.22	16±0.20	16±0.19	18±0.28	20±0.16	22±0.40	24±0.45	26±0.10	24[T]	–
P. aeruginosa	–	–	–	10±0.44	13±0.22	15±0.52	13±0.22	15±0.52	17±0.20	–	–	–	25[T]	–
P. vulgaris	–	–	–	10±0.45	12±0.64	14±0.22	10±0.64	12±0.10	14±0.40	14±0.90	16±0.76	18±0.20	22[T]	–
A. niger	–	–	–	–	–	–	10±0.76	10±0.28	12±0.19	–	–	–	18[F]	–
C. albicans	–	–	–	–	–	–	16±0.10	18±0.90	20±0.28	14±0.10	16±0.16	18±0.64	23[F]	–
S. cerevisiae	–	–	–	–	–	–	15±0.16	17±0.64	18±0.40	–	–	–	20[F]	–

a: values are the mean of three replicates of ±SE; S: Standard antibiotics T- Tetracycline; F – Fluconazole

D: DMSO; –: No activity

Table-2: Antimicrobial activity of roots of *Chloroxylon swietenia* DC.

Organisms	Zone of inhibition (mm)[a]													
	Hexane extract (mg/ml)			Chloroform extract (mg/ml)			Methanol extract (mg/ml)			Aqueous extract (mg/ml)			S	D
	100	200	300	100	200	300	100	200	300	100	200	300		
B. subtilis	14±0.64	15±0.45	16±0.90	14±0.28	16±0.28	17±0.19	13±0.19	15±0.90	17±0.50	16±0.90	18±0.40	20±0.40	18^{T}	–
E. faecalis	–	–	–	17±0.28	19±0.64	21±0.28	16±0.10	18±0.50	20±0.22	14±0.22	16±0.10	18±0.52	21^{T}	–
M. letus	15±0.16	17±0.10	19±0.52	21±0.10	22±0.52	24±0.19	24±0.28	26±0.40	28±0.50	15±0.52	17±0.90	19±0.64	24^{T}	–
S. aureus	–	–	–	–	–	–	9±0.50	11±0.19	13±0.76	–	–	–	24^{T}	–
S. pneumoniae	–	–	–	–	–	–	12±0.10	14±0.28	16±0.22	–	–	–	22^{T}	–
E. coli	–	11±0.22	12±0.16	12±0.90	14±0.22	16±0.28	12±0.22	14±0.50	16±0.76	12±0.40	14±0.76	16±0.19	22^{T}	–
K. pneumoniae	7±0.90	9±0.44	10±0.22	8±0.90	10±0.44	12±0.19	10±0.50	12±0.19	14±0.20	10±0.44	12±0.40	14±0.20	24^{T}	–
P. aeruginosa	–	–	–	–	–	–	15±0.76	17±0.28	18±0.16	–	11±0.22	13±0.90	25^{T}	–
P. vulgaris	10±0.44	12±0.22	14±0.22	12±0.64	14±0.45	16±0.40	18±0.28	20	22±0.28	18±0.19	20±0.50	22±0.52	22^{T}	–
A. niger	–	10±0.45	13±0.20	16±0.28	18±0.22	20±0.20	18±0.64	20±0.28	22±0.64	20±0.52	22±0.40	24±0.16	18^{F}	–
C. albicans	–	–	–	–	–	–	–	–	–	–	–	–	23^{F}	–
S. cerevisiae	14±0.90	16±0.52	18±0.28	16±0.20	18±0.52	20±0.90	20±0.16	22±0.90	24±0.90	–	–	–	20^{F}	–

a: values are the mean of three replicates of ±SE; S: Standard antibiotics T- Tetracycline; F – Fluconazole

D: DMSO; –: No activity

The aqueous extract showed moderate zone of inhibition (20 -28 mm) against *M. letues. P. bulgaris, E. faecalis, A. niger* and *S. cerevisiae* respectively. However, hexane extract obtained from root did not show any antimicrobial activity against all the tested microorganisms. The inhibitory activities of all the extracts on compared with standard antibiotics tetracycline and flucanozole. The results revealed that the methanolic chloroform and aqueous extracts showed higher antimicrobial activity than the antibiotics.

Minimum Inhibitory Concentration (MIC)

MIC value of selected two plant extracts (both aqueous and solvents) was observed in range of 7.8-500 µg/ml against tested bacteria and fungi (Table 3 and 4). *Chlorozylon swietenia* methanolic, chloroform and aqueous extracts inhibited by the Gram-positive bacteria (*M. luteus, E. faecalis, B. subtilis*) at 7.8 - >1000 µg/ml and Gram-negative *P. vulgaris, E. coli, K. pneumoniae, P. aeruginosa* at 62.5 - >1000 µg/ml concentrations and fungal strains (*A. niger, S. cerevisiae*) at 31.2 - >1000 µg/ml concentrations. Similarly, *C. zylanica* both solvent (methanol, chloroform and hexane) and aqueous extracts was observed in range of 31.2 - >1000 µg/ml showed inhibition in most of the strains Gram positive bacteria (*M. luteus, B. subtilis, S. aureus, E. faecalis*) and Gram-negative bacteria (*K. pneumoniae, P. vulgaris, P. aeruginosa*) and fungal strains (*A. niger, C. albicans*), respectively.

Table 3: MIC Values of roots of *Capparis zeylanica* L.

Organisms	hexane extract(µg/ml)	chloroform extract(µg/ml)	methanol extract(µg/ml)	aqueous extract(µg/ml)
B. subtilis	–	–	250	250
E. faecalis	–	>1000	250	–
M. letus	1000	1000	125	–
S. aureus	–	–	250	–
S. pneumoniae	–	–	1000	–
E. coli	–	–	250	–
K. pneumoniae	1000	1000	125	31.2
P. aeruginosa	–	1000	250	–
P. vulgaris	–	>1000	1000	500
A. niger	–	–	>1000	–
C. albicans	–	–	125	500
S. cerevisiae	–	–	250	–

Table 4: MIC Values of roots of *Chloroxylon swietenia* DC

Organisms	hexane extract(µg/ml)	chloroform extract(µg/ml)	methanol extract(µg/ml)	aqueous extract(µg/ml)
B. subtilis	1000	1000	250	500
E. faecalis	–	500	125	1000
M. letus	500	250	7.8	250
S. aureus	–	–	>1000	–
S. pneumoniae	–	–	500	–
E. coli	>1000	1000	500	1000
K. pneumoniae	>1000	>1000	1000	>1000
P. aeruginosa	–	–	250	>1000
P. vulgaris	>1000	1000	62.5	250
A. niger	>1000	500	62.5	125
C. albicans	–	–	–	–
S. cerevisiae	1000	500	31.2	–

The methanol, hexane, chloroform and aqueous extracts of roots of selected medicinal plants were investigated at three different concentrations i.e. (100, 200, 300 mg/ml) by agar well diffusion method against nine bacterial and three fungal strains. The antibacterial activity was expressed as the average diameter of the zone of inhibition of microorganisms growth around the well/disc. Most of the extracts displayed relatively high antibacterial activity against mostly the tested microorganisms with the diameter of inhibition zones ranging between 0 to 34±0.26 mm. The Gram-positive strains were found to be the most susceptible to growth inhibition by the plant extracts forming zones of inhibition ranging from 0 to 34±0.26. While Gram negative strains were less sensitive to the plant extracts as compared to the Gram positive, forming zones of inhibition ranging from 0 to 30±0.25. susceptibility differences between Gram positive and Gram-negative bacteria may be due to cell wall structural differences between these bacteria. Gram positive bacteria having only an outer peptidoglycon layer which is not effective permeability barrier, where as the Gram-negative bacteria having an outer phospholipidic membrane carrying the structural lipopolysaccharide components. This makes the cell wall impermeable to drug constituents. The greater susceptibility of Gram-positive bacteria to plant extracts have been previously reported in medicinal plant extracts (Palombo and Sample, 2001; Kiran *et al.,* 2008; Chanda *et al.,* 2010; 2011 and 2013). Also, the methanol extracts of *E. hirta* (whole plant) *C. swietenia* (root) *L. glutinosa* (bark) and *E. bracteata* (leaves) were found to have good antimicrobial activity against bacteria than fungi. This difference the cell wall in Gram positive bacteria is of single layer, whereas the Gram-negative cell wall is multilayered structure and the fungi cell wall is quite complex structure and extensive cross linking between

chitin, glucans and other polymers. Similar results were also reported in several medicinal plants by Chanda *et al.,* 2010; Rakholiya and Chanda 2012. These results are in good agreement with the results reported earlier by Ali *et al.,* (2004). Antifungal activity is not common in medicinal plants. However, the significant antifungal activity was observed in both organic and aqueous solvent extracts of 30 selected plants except *I. obscura, S. jamaicansis* and *C. paniculata*. The inhibition zones ranging from 0 to 26 ±0.21. The results of the antifungal activity of the different extracts of selected plants were compared with the standard antibiotics, Methanolic whole plant extract of *T. hirta* and leaf extract of *E. bracteata* showed that higher and significant zone of inhibition than the standard antibiotic flucanozole, which suggest that plant extracts contain more soluble chemical constituents in methanolic solvent with antimicrobial properties that can be used as antimicrobial agents in new drugs for the therapy of infectious diseases caused by pathogens. It is interesting to note that the extracts are not pure compounds and in spite of it good results were obtained which only suggests the potency of these extracts.

In the present study, the selected medicinal plants contain secondary metabolites like alkaloids, phenolic compounds, glycosides, tannins, anthroquinones, reducing compounds complex with proteins and other macromolecules present in the test medium, therefore, get precipitated. While some extract components especially the non polar are not readily soluble in test medium which is more than 60 % hexane and chloroform in most cases. These factors may at times cause reduction in the effectiveness of the plant extracts to inhibit microbial growth. Among the 5 Gram positive bacterial strains *S. aureus* more resistant; *P. vulgaris* (Gram negative bacteria) is more resistant the rest of three negative bacteria and among the three fungal strains *A. niger* is the most resistant fungi. Similarly, among the tested microorganisms *B. subtilis* is the most susceptible (Gram positive bacteria), *P. aeruginosa* (Gram negative bacteria) and *S. cerevisiae* (fungi), respectively in this study.

MIC is defined as the lowest concentration of and antimicrobial that will inhibit the visible growth of a microorganism after overnight incubation. It is dependent of many factors and the non standardized procedures make difficult to compare MIC results from author to author. MIC however is useful as a practical indicator of primary activity against a selected pathogenic microorganism. In the present study, *in vitro* antimicrobial of the crude solvent extracts derived from 30 selected plants used in traditional medicine for treating different ailments was quantitatively assessed on the basis of minimum inhibitory concentration (MIC). All the plants evaluated exhibited varying degree of inhibitory effect against standard strain of human pathogenic bacteria (Gram positive and Gram negative and fungi). MIC values varying from 7.8 to 1000 µg/ml (Gram positive and Gram negative) and 15.6 to 1000 µg/ml. There have no specific cut off values as a reference or standard for categorizing antimicrobial activity of plant extracts. In this study, crude extracts with an MIC value is less than 500 µg/ml were considered

to have good activity and MIC values less than 100 µg/ml were considered to have significant activity of Pharmacological interest. A lower MIC values indicated high effectiveness of the compound as antimicrobial agent as little quantity which may below toxicity level of the extracts can be applied without being harmful effects. Similar results were also reported in *E. hirta* (Kader *et al.,* 2013).

Conclusions

From the above results it can be concluded that the plant extracts have great potential as antimicrobial compounds against microorganisms and that they can be used in the treatment of infectious diseases caused by resistant microorganisms. The selected medicinal plants showed maximum antimicrobial activity and so these plants can be used to discover quantification of bioactive natural products that may serve as leads for the development of new pharmaceuticals that address hither to unmet therapeutic needs. Such screening of various natural organic compounds and identifying active agents is the need of the hour because successful prediction of lead molecule and drug like properties at the onset of drug discovery will play off later in drug development. Thus, the study ascertains the value of plants used in Ayurveda, which could be of considerable interest to the development of new drugs. Further, plant based antimicrobials have enormous therapeutic potential as they can serve the purpose without any side effects as compared to synthetic compounds.

References

1. Ali, M. A., Sayeed, M. A., Bhujan, M. S. A., Sohel, F., & Yeasmin, M. S. (2004). Antimicrobial screening of *Cassia fistula* and *Mesua ferrea*. *Journal of Medical Sciences, 4*, 24–29.
2. Aniel Kumar, O., Krishna Rao, M., & Mutyala Naidu, L. (2014). Evaluation of antimicrobial activity of chemical constituents of *Achyranthes aspera* L. roots against human pathogens. *Indian Journal of Natural Products and Resources, 5*(3), 278–281.
3. Aniel, K. O., & Mutyala, N. L. (2016). *Bridelia retusa* (L.) Spreng. fruits: Antimicrobial efficiency and their phytochemical constituents. *Notulae Scientia Biologicae, 8*(1), 33–36.
4. Aniel Kumar, O., Krishna Rao, M., & Mutyala Naidu, L. (2015). Evaluation of antimicrobial activity of root extracts of *Abutilon indicum*. *Notulae Scientia Biologicae, 7*(2), 160–163.
5. Balakrishnan, N., Bhaskar, V. H., Jayakar, B., & Sangeswaran, B. (2006). Antibacterial activity of *Mimosa pudica*, *Aegle marmelos* and *Sida cordifolia*. *Pharmacognosy Magazine, 2*, 198–199.
6. Chanda, S., Dudhatra, S., & Knaeria, M. (2010). Antioxidative and antibacterial effects of seeds and fruit rind of nutraceutical plants belonging to the family Fabaceae. *Food Function, 1*, 308–315.

7. Chanda, S., Rakholiya, K., & Paresh, J. (2013). Indian medicinal herb: Antimicrobial efficacy of *Mesua ferrea* L. seed extracted in different solvents against infection-causing pathogenic strains. *Journal of Acute Disease, 2*, 277–281.
8. Kader, J., Hanina, M. N., Salina, M. R., & Nur, W. (2013). Antibacterial activities and phytochemical screening of the acetone extract from *Euphorbia hirta*. *International Journal of Medicinal Plants Research, 2*(4), 209–214.
9. Kiran, S. R., Sita, P., & Janardhan, K. (2008). Evaluation of *in vitro* antimicrobial activity of leaf and stem essential oils of *Chloroxylon swietenia* DC. *World Journal of Microbiology and Biotechnology, 24*, 1909–1914.
10. Palombo, E. A., & Semple, S. J. (2001). Antibacterial activity of traditional medicinal plants. *Journal of Ethnopharmacology, 77*, 151–157.
11. Rakholiya, K., & Chanda, S. (2012). *In vitro* interaction of certain antimicrobial agents in combination with plant extracts against some pathogenic bacterial strains. *Asian Pacific Journal of Tropical Biomedicine, 2*, S876–S880.

STEM CELL ENGINEERING AND REGENERATIVE MEDICINE: PRINCIPLES, TECHNOLOGIES, AND CLINICAL TRANSLATION

Ranjana

Department of Zoology,

Patna Science College, Patna University, Patna, Bihar, 800005, India.

Corresponding author E-mail: ranjana.prakash81@gmail.com, ranjanazoolpsc@pup.ac.in

Abstract

Stem cell engineering and regenerative medicine represent one of the most transformative frontiers in biomedical science, offering the prospect of repairing or replacing damaged tissues and organs through the directed manipulation of cellular identity and function. This chapter provides a comprehensive overview of the biological foundations of stem cell biology, the principal engineering strategies used to guide stem cell behaviour, and the major therapeutic applications emerging from this convergence of disciplines. Topics include the classification and properties of pluripotent and adult stem cells, reprogramming technologies such as induced pluripotent stem cell (iPSC) derivation, biomaterial scaffolding, organoid systems, gene editing via CRISPR-Cas9, and the regulatory and ethical considerations that accompany clinical translation.

Keywords: Stem Cells, Induced Pluripotent Stem Cells, Regenerative Medicine, Tissue Engineering, CRISPR-Cas9, Organoids, Cell Therapy, Biomaterials, Clinical Translation.

1. Introduction

The capacity to harness the body's own cellular machinery to heal injury and disease has long been a foundational aspiration of medicine. Over the past three decades, rapid advances in cell biology, materials science, and genomics have transformed this aspiration into a tractable scientific programme now known as regenerative medicine. At the heart of this enterprise lies the stem cell — a cell endowed with the dual capacities of self-renewal and differentiation — and the engineering strategies that enable its controlled manipulation [1,2].

The modern era of stem cell research is conventionally traced to 1961, when Till and McCulloch demonstrated that single cells derived from mouse bone marrow could reconstitute the haematopoietic system following lethal irradiation [3]. This landmark observation established the conceptual framework — the clonogenic, multipotent stem cell — that continues to organise the field. Subsequent decades brought the isolation of embryonic stem cells (ESCs) from murine blastocysts by Evans and Kaufman in 1981 [4], and the landmark derivation of human ESCs by Thomson and colleagues in 1998 [5]. These advances opened unprecedented windows into human

developmental biology, but also ignited ethical controversies surrounding the use of embryonic material that persist to the present.

The ethical landscape was substantially reconfigured in 2006 when Takahashi and Yamanaka reported that somatic cells could be reprogrammed to a pluripotent state through the ectopic expression of a small set of transcription factors [6]. The resulting *induced pluripotent stem cells* (iPSCs) combined the developmental plasticity of ESCs with the ethical accessibility of adult tissue, catalysing an explosion of research into patient-specific cell therapies, disease modelling, and drug discovery.

Parallel progress in biomaterials science, microfluidics, and gene editing — particularly the adaptation of the CRISPR-Cas9 system for precise genomic modification [7] — has provided engineers with an expanding toolkit for shaping stem cell behaviour in physiologically and clinically relevant ways. The convergence of these technologies underpins the contemporary field of stem cell engineering, which seeks to translate fundamental biology into durable therapeutic benefit.

This chapter systematically examines the biological and engineering foundations of the field, surveys major therapeutic applications, and considers the translational and regulatory challenges that must be navigated to bring stem cell–based therapies safely and equitably to patients.

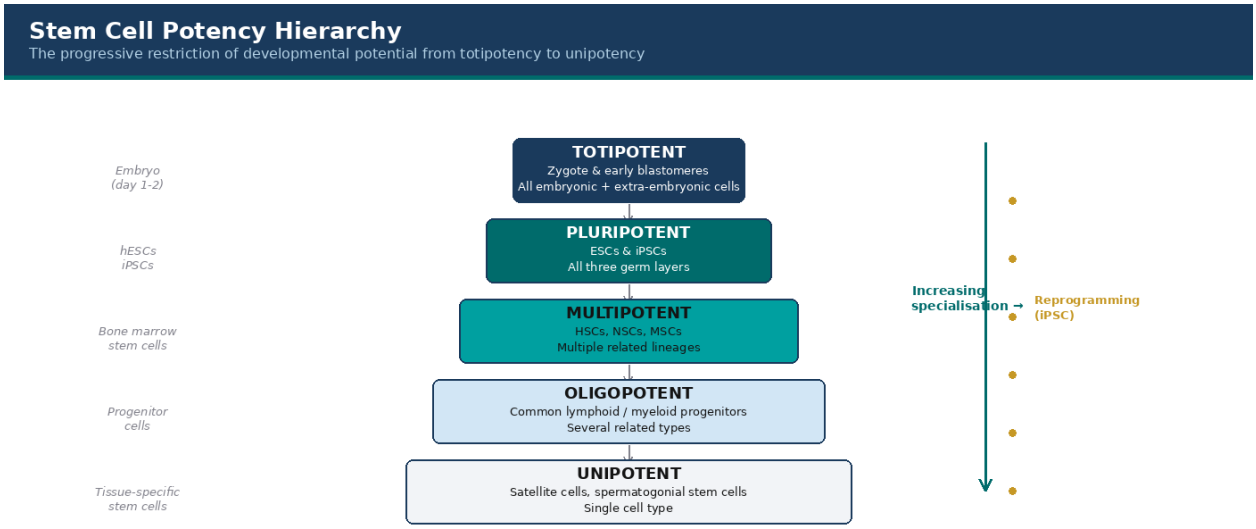


Figure 1: The stem cell potency hierarchy. Developmental potential narrows progressively from totipotency (zygote) through pluripotency (ESCs, iPSCs), multipotency, and oligopotency to unipotency. Induced pluripotent stem cell reprogramming (gold dashed arrow) reverses this trajectory

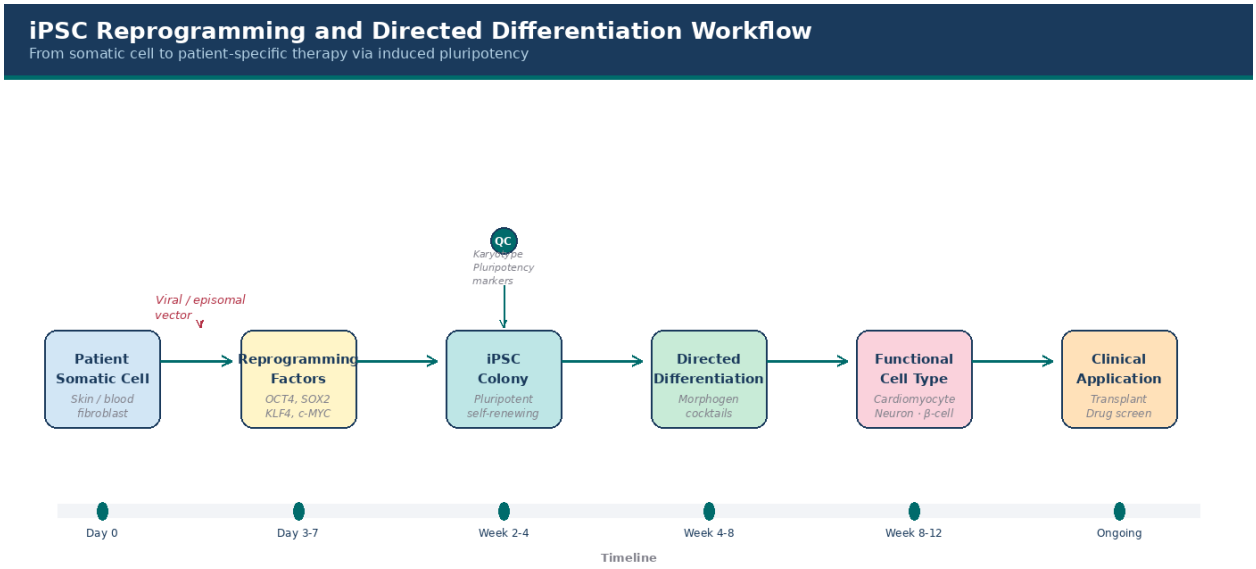


Figure 2: Overview of the iPSC reprogramming and directed differentiation workflow. Patient somatic cells are reprogrammed by delivery of transcription factors (OCT4, SOX2, KLF4, c-MYC) via viral or episomal vectors, expanded as quality-controlled iPSC colonies, and then converted to functional cell types for therapy, disease modelling, or drug screening

2. Biology of Stem Cells

2.1 Defining Properties and Hierarchical Organisation

Stem cells are defined by two cardinal properties: the capacity for indefinite self-renewal — the ability to divide and produce daughter cells that retain stem cell identity — and potency, the ability to differentiate into one or more specialised cell types [8]. These properties exist on a spectrum, and stem cells are conventionally classified according to their developmental breadth (Figure 1).

Totipotent cells, exemplified by the zygote and early blastomeres, can give rise to all embryonic and extra-embryonic tissues. *Pluripotent* stem cells, including ESCs and iPSCs, can generate all cell types of the three primary germ layers (ectoderm, mesoderm, endoderm) but not extra-embryonic structures. *Multipotent* stem cells are restricted to lineages within a given tissue — haematopoietic stem cells (HSCs), for example, generate all blood cell types but not neurons or hepatocytes. *Oligopotent* and *unipotent* cells occupy progressively narrower niches [9].

This hierarchy reflects, in molecular terms, the progressive restriction of chromatin accessibility and transcriptional programmes. The epigenetic landscape — comprising DNA methylation, histone modification, and three-dimensional chromatin architecture — encodes cell identity and is dynamically remodelled during differentiation and, crucially, during reprogramming [10].

2.2 Pluripotent Stem Cells

2.2.1 Embryonic Stem Cells

Human embryonic stem cells (hESCs) are derived from the inner cell mass (ICM) of the blastocyst at approximately five to seven days post-fertilisation [5]. They are maintained in a self-renewing state in vitro by signalling pathways — including leukaemia inhibitory factor (LIF)/STAT3 in mice and FGF2/Activin-Nodal in humans — that sustain expression of a core pluripotency transcription factor network centred on OCT4, SOX2, and NANOG [11].

Two distinct pluripotent states have been characterised: the *naive* state, corresponding to the preimplantation epiblast and characterised by global DNA hypomethylation and bivalent chromatin at developmental gene promoters, and the *primed* state, corresponding to the post-implantation epiblast and exhibiting a more restricted and methylated epigenome [12]. The ability to interconvert between these states has significant implications for the efficiency of directed differentiation protocols.

2.2.2 Induced Pluripotent Stem Cells

The generation of iPSCs by Takahashi and Yamanaka through retroviral transduction of *Oct4, Sox2, Klf4*, and *c-Myc* into mouse fibroblasts — and subsequently into human somatic cells [6,13] — was recognised with the Nobel Prize in Physiology or Medicine in 2012. iPSCs are molecularly and functionally indistinguishable from ESCs by most criteria, though subtle epigenetic and transcriptional differences, sometimes referred to as *epigenetic memory*, can bias differentiation potential toward the somatic cell of origin [14].

Clinical Significance: iPSCs derived from individual patients carry the patient's own genetic background, enabling autologous cell therapy — in principle, eliminating immune rejection — and providing genetically faithful platforms for disease modelling and personalised drug screening.

Refinements to reprogramming methodology have substantially improved efficiency and safety. Integration-free delivery of reprogramming factors using episomal plasmids, RNA transfection, Sendai virus, and small-molecule substitution of individual factors have reduced or eliminated the risk of insertional mutagenesis associated with integrating viral vectors [15]. Small-molecule compounds that modulate chromatin state — including inhibitors of HDAC, GSK3, and MEK — can replace individual transcription factors and improve reprogramming kinetics [16].

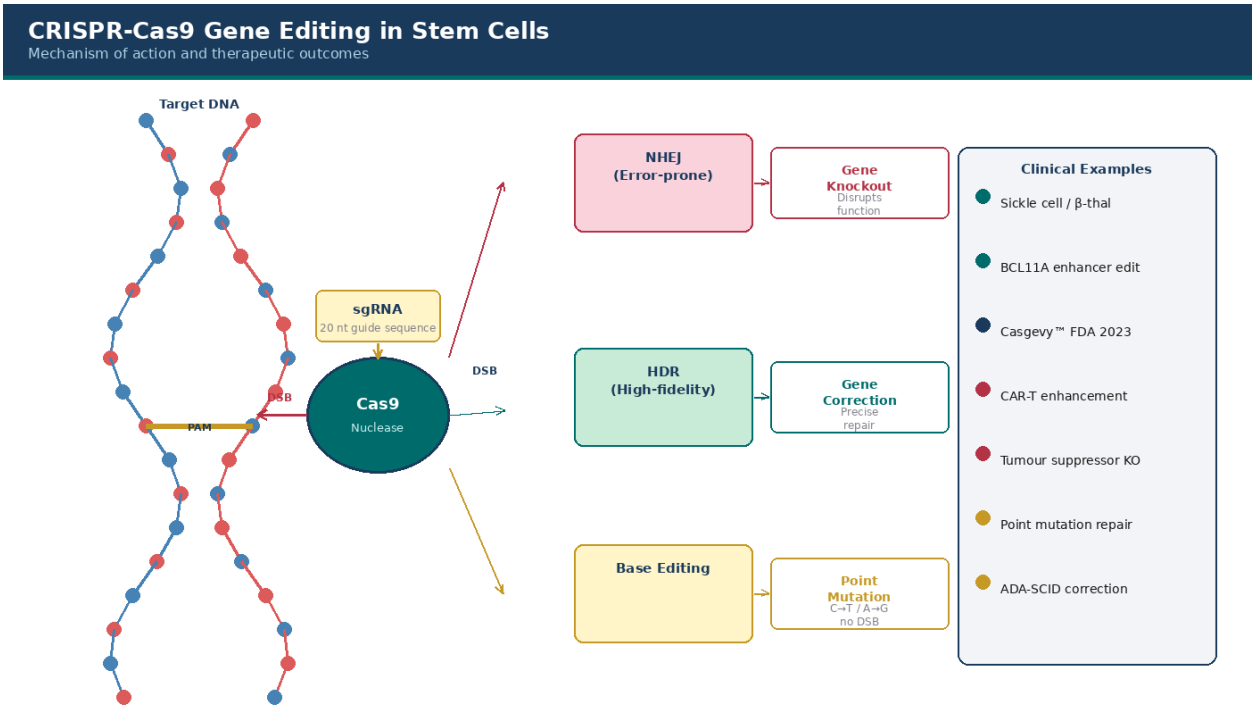


Figure 3: Mechanism of CRISPR-Cas9 gene editing in stem cells. The sgRNA directs Cas9 to a target DNA locus adjacent to a PAM sequence. Double-strand break (DSB) repair proceeds via NHEJ (gene knockout), HDR (gene correction), or next-generation approaches including base editing (no DSB). Casgevy™, approved by the FDA in December 2023, applies BCL11A enhancer disruption in HSCs to treat sickle cell disease and β-thalassaemia

2.3 Adult Stem Cells and Tissue-Resident Populations

Every self-renewing tissue harbours a resident stem or progenitor cell population responsible for homeostatic maintenance and injury-induced regeneration. The best-characterised are haematopoietic stem cells (HSCs), which reside in the bone marrow and sustain lifelong blood cell production [17]. HSC transplantation remains the gold-standard curative therapy for a range of haematological malignancies and immune disorders and constitutes the oldest and most clinically mature form of stem cell therapy.

Other well-characterised tissue-resident populations include neural stem cells (NSCs) in the subventricular zone and hippocampal dentate gyrus [18], intestinal stem cells in the crypts of Lieberkühn [19], skeletal muscle satellite cells [20], and cardiac progenitors with limited regenerative capacity [21]. Mesenchymal stromal cells (MSCs) — heterogeneous populations derived from bone marrow, adipose tissue, and other sources — exhibit multilineage differentiation potential and potent paracrine immunomodulatory activities that have attracted intense clinical interest [22].

The microenvironmental niche — comprising extracellular matrix, soluble signals, mechanical forces, and cell-cell contacts — plays a determinative role in controlling stem cell fate. Dysregulation of niche signals contributes to tissue ageing, degenerative disease, and malignant

transformation [23]. Engineering strategies that recreate or manipulate niche signals in vitro therefore represent a central challenge of the field.

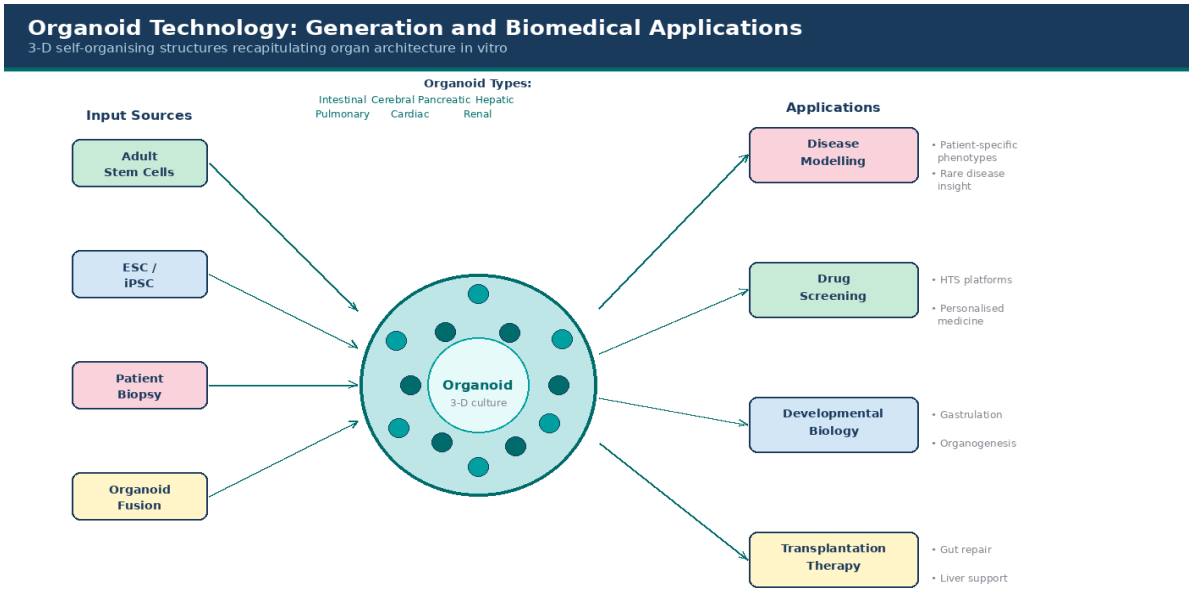


Figure 4. Organoid technology: generation and biomedical applications. Stem cells from multiple sources self-organise in 3-D Matrigel or hydrogel matrices to form organ-like structures. Organoids are employed in disease modelling, high-throughput drug screening, developmental biology research, and — increasingly — transplantation therapy

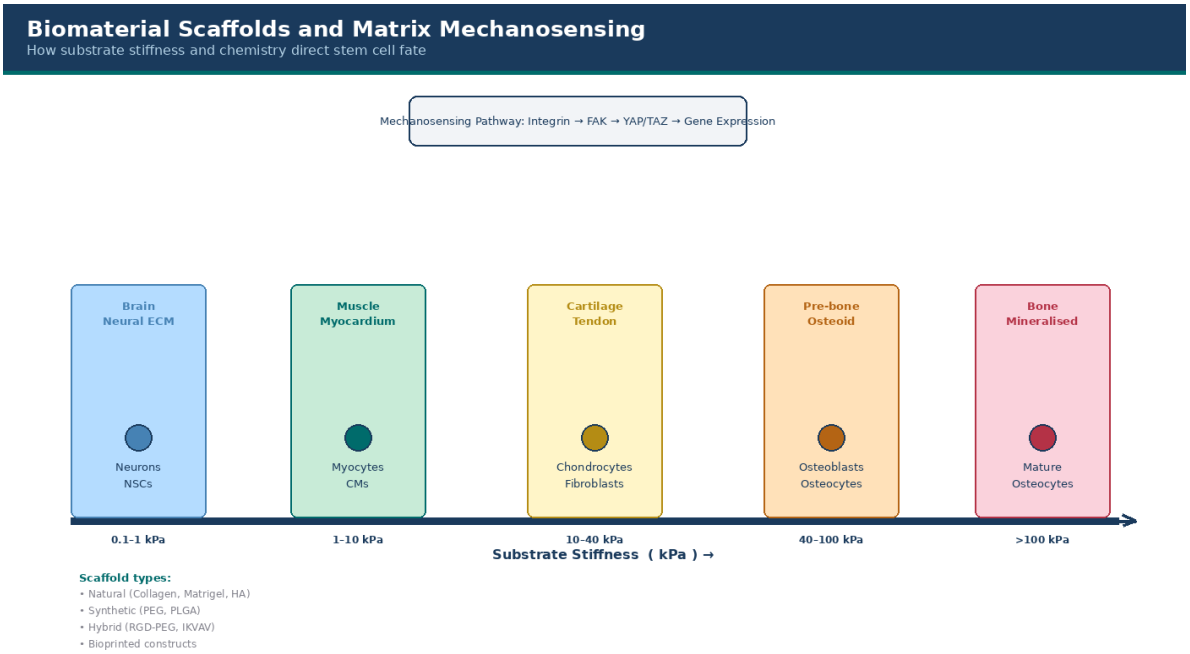


Figure 5: Biomaterial scaffolds and matrix mechanosensing. Substrate stiffness ranging from <1 kPa (brain-like) to >100 kPa (bone-like) directs MSC lineage specification toward neuronal, myogenic, chondrogenic, or osteogenic fates via integrin–FAK–YAP/TAZ mechanotransduction (Engler *et al.*, 2006)

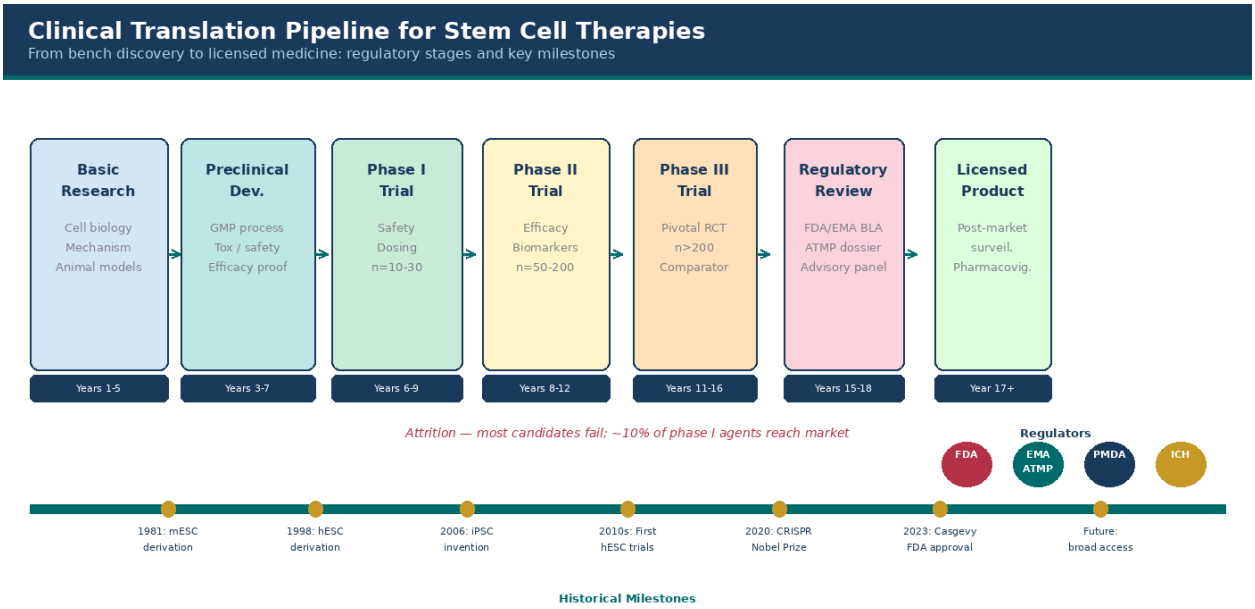


Figure 6: Clinical translation pipeline for stem cell therapies. Development proceeds from basic research through preclinical validation, Phase I–III trials, regulatory review (FDA BLA or EMA ATMP authorisation), to licensed products. Key historical milestones are shown along the timeline. Most candidates fail during Phase II/III; approximately 10% of Phase I agents reach market

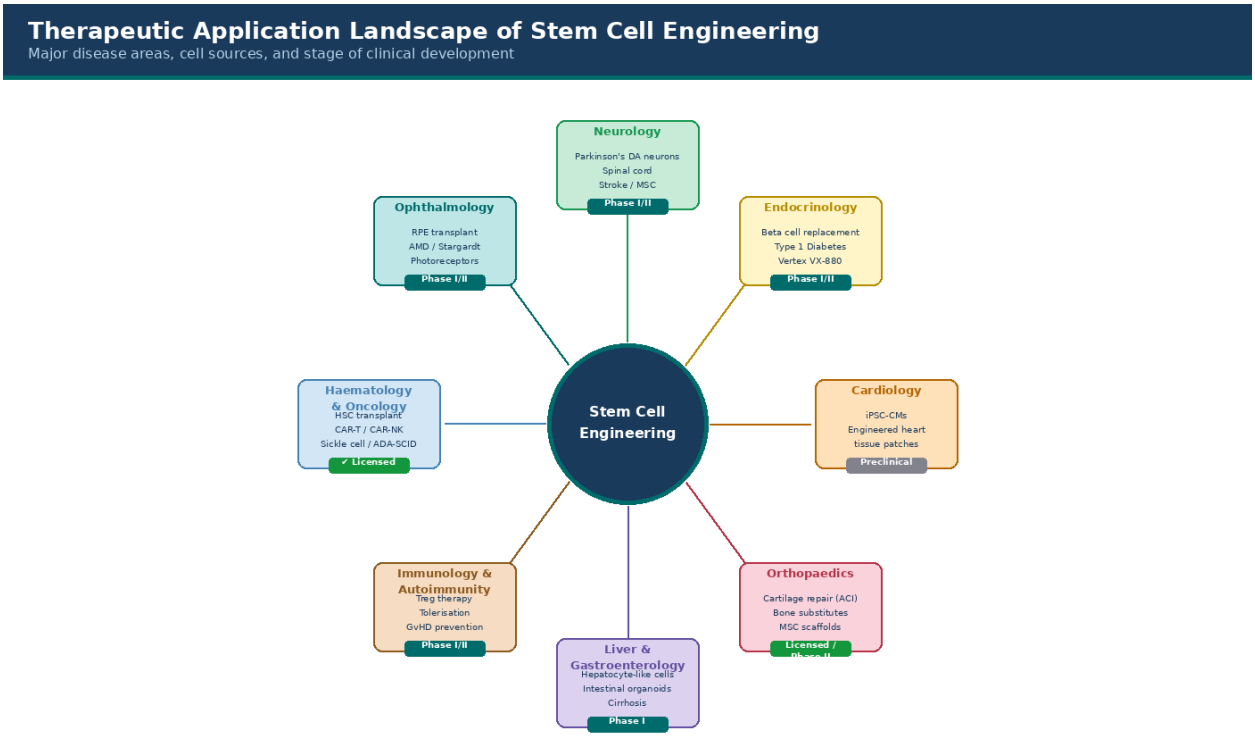


Figure 7: Therapeutic application landscape of stem cell engineering. The diagram maps major disease areas by cell source, therapeutic strategy, and current stage of clinical development. Haematological applications are most mature (licensed products); cardiology and hepatology remain largely preclinical

References

1. Weissman, I. L. (2000). Stem cells: Units of development, units of regeneration, and units in evolution. *Cell, 100*(1), 157–168. https://doi.org/10.1016/S0092-8674(00)81692-X
2. Langer, R., & Vacanti, J. P. (1993). Tissue engineering. *Science, 260*(5110), 920–926. https://doi.org/10.1126/science.8493529
3. Till, J. E., & McCulloch, E. A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research, 14*, 213–222.
4. Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature, 292*(5819), 154–156. https://doi.org/10.1038/292154a0
5. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., *et al.* (1998). Embryonic stem cell lines derived from human blastocysts. *Science, 282*(5391), 1145–1147. https://doi.org/10.1126/science.282.5391.1145
6. Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell, 126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024
7. Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science, 346*(6213), 1258096. https://doi.org/10.1126/science.1258096
8. Morrison, S. J., & Spradling, A. C. (2008). Stem cells and niches: Mechanisms that promote stem cell maintenance throughout life. *Cell, 132*(4), 598–611. https://doi.org/10.1016/j.cell.2008.01.038
9. Clevers, H. (2011). The cancer stem cell: Premises, promises and challenges. *Nature Medicine, 17*(3), 313–319. https://doi.org/10.1038/nm.2304
10. Atlasi, Y., & Stunnenberg, H. G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. *Nature Reviews Genetics, 18*(11), 643–658. https://doi.org/10.1038/nrg.2017.57
11. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell, 122*(6), 947–956. https://doi.org/10.1016/j.cell.2005.08.020
12. Nichols, J., & Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell, 4*(6), 487–492. https://doi.org/10.1016/j.stem.2009.05.015
13. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., *et al.* (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell, 131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019
14. Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., *et al.* (2010). Epigenetic memory in induced pluripotent stem cells. *Nature, 467*(7313), 285–290. https://doi.org/10.1038/nature09342

15. Shi, Y., Inoue, H., Wu, J. C., & Yamanaka, S. (2017). Induced pluripotent stem cell technology: A decade of progress. *Nature Reviews Drug Discovery, 16*(2), 115–130. https://doi.org/10.1038/nrd.2016.245
16. Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A. E., *et al.* (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnology, 26*(7), 795–797. https://doi.org/10.1038/nbt1418
17. Weissman, I. L. (2000). Translating stem and progenitor cell biology to the clinic: Barriers and opportunities. *Science, 287*(5457), 1442–1446. https://doi.org/10.1126/science.287.5457.1442
18. Gage, F. H. (2000). Mammalian neural stem cells. *Science, 287*(5457), 1433–1438. https://doi.org/10.1126/science.287.5457.1433
19. Barker, N., van Es, J. H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., *et al.* (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature, 449*(7165), 1003–1007. https://doi.org/10.1038/nature06196
20. Mauro, A. (1961). Satellite cell of skeletal muscle fibers. *Journal of Biophysical and Biochemical Cytology, 9*, 493–495. https://doi.org/10.1083/jcb.9.2.493
21. Bergmann, O., Bhardwaj, R. D., Bernard, S., Zdunek, S., Barnabé-Heider, F., Walsh, S., *et al.* (2009). Evidence for cardiomyocyte renewal in humans. *Science, 324*(5923), 98–102. https://doi.org/10.1126/science.1164680
22. Pittenger, M. F., Mackay, A. M., Beck, S. C., Jaiswal, R. K., Douglas, R., Mosca, J. D., *et al.* (1999). Multilineage potential of adult human mesenchymal stem cells. *Science, 284*(5411), 143–147. https://doi.org/10.1126/science.284.5411.143
23. Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells, 4*(1–2), 7–25.

Bhumi Publishing, India
May 2026

THE APPLICATION AND RESEARCH OF NEW DIGITAL TECHNOLOGY IN MARINE AQUACULTURE

S. A. Raj Vasanth*[1], P. Chidambaram[2], P. Velmurugan[3], S. Athithan[1] and N. Jayakumar[2]

[1]Department of Aquaculture, Tamil Nadu Dr. J. Jayalalithaa Fisheries University, Fisheries College and Research Institute, Thoothukudi-628 008.

[2]Tamil Nadu Dr. J. Jayalalithaa Fisheries University, Nagapattinam-611 002.

[3]Tamil Nadu Dr. J. Jayalalithaa Fisheries University, Directorate of Sustainable Aquaculture, Theni Centre for Sustainable Aquaculture, Gullapuram, Theni-625 562.

*Corresponding author E-mail: raaathiraj@gmail.com

Abstract

Marine aquaculture has emerged as an important sector for ensuring global food security and meeting the increasing demand for seafood. However, challenges such as environmental degradation, disease outbreaks, climate change, and resource limitations continue to affect the sustainability and productivity of marine farming systems. In recent years, new digital technologies have transformed marine aquaculture by improving monitoring, automation, management efficiency, and decision-making processes. Technologies such as the Internet of Things (IoT), artificial intelligence (AI), big data analytics, blockchain, robotics, drones, and remote sensing are increasingly integrated into aquaculture operations. These innovations help optimize feeding, monitor water quality, predict disease outbreaks, and reduce labor costs while improving environmental sustainability. Digital technologies also support precision aquaculture, enabling real-time data collection and intelligent farm management. This essay examines the application and research of new digital technologies in marine aquaculture, highlighting their benefits, challenges, and future prospects for sustainable marine food production.

Keywords: Marine Aquaculture, Digital Technology, Artificial Intelligence, Internet of Things, Precision Farming.

Introduction

Marine aquaculture, also known as mariculture, refers to the cultivation of marine organisms such as fish, shellfish, seaweed, and crustaceans in seawater environments. It plays a vital role in global seafood production and contributes significantly to food security, employment, and economic development. According to the Food and Agriculture Organization (FAO), aquaculture has become one of the fastest-growing food production sectors worldwide, with marine aquaculture accounting for a substantial portion of global seafood supply. Despite its rapid growth, marine aquaculture faces several challenges, including disease outbreaks, poor water quality, environmental pollution, high operational costs, climate change, and overexploitation of

marine resources. Traditional farming methods often rely heavily on manual labor and limited monitoring systems, which can lead to inefficiency and economic losses. To overcome these challenges, the aquaculture industry is increasingly adopting new digital technologies that improve productivity, sustainability, and farm management. The advancement of digital technologies such as artificial intelligence (AI), the Internet of Things (IoT), machine learning, big data analytics, cloud computing, blockchain, robotics, and remote sensing has created new opportunities for marine aquaculture development. These technologies enable real-time monitoring, automation, precision farming, disease detection, and data-driven decision-making. Researchers and aquaculture companies are investing heavily in digital transformation to improve production efficiency and environmental sustainability. This context explores the major applications and research developments of new digital technologies in marine aquaculture. It discusses the role of smart monitoring systems, AI-based management, automation, data analytics, blockchain traceability, robotics, and future research trends in modern aquaculture systems.

Internet of Things (IoT) in Marine Aquaculture

One of the most important digital technologies applied in marine aquaculture is the Internet of Things (IoT). IoT refers to a network of interconnected devices and sensors that collect, transmit, and analyze data in real time. In marine aquaculture, IoT systems are used to monitor environmental conditions such as temperature, salinity, dissolved oxygen, pH, ammonia levels, and water currents. Traditional water quality monitoring methods often require manual sampling, which is time-consuming and less accurate. IoT-based monitoring systems provide continuous and automated data collection, allowing farmers to respond quickly to environmental changes. Sensors installed in cages, ponds, and tanks transmit real-time information to cloud-based platforms or mobile applications. Smart aquaculture systems using IoT technology help reduce fish mortality and improve feed efficiency. For example, oxygen sensors can automatically activate aeration systems when dissolved oxygen levels become low. Similarly, temperature sensors help farmers maintain optimal growth conditions for marine species such as salmon, shrimp, and sea bass.

Artificial Intelligence and Machine Learning in Aquaculture

Artificial intelligence (AI) and machine learning are transforming marine aquaculture by enabling automated decision-making and predictive analysis. AI systems analyze large amounts of data collected from sensors, cameras, and historical records to optimize aquaculture operations. One of the major applications of AI in marine aquaculture is automated feeding management. Feed represents one of the highest operational costs in aquaculture production. Overfeeding leads to feed waste and water pollution, while underfeeding reduces fish growth. AI-powered feeding systems use underwater cameras and behavioral analysis to determine fish

appetite and adjust feed supply accordingly. Machine learning algorithms can also predict fish growth rates, disease outbreaks, and environmental risks. By analyzing patterns in water quality and fish behavior, AI systems can detect abnormalities before serious problems occur. Early disease detection is particularly important because disease outbreaks can cause massive economic losses in marine aquaculture. Computer vision technology, a branch of AI, is widely used for fish monitoring. Cameras installed underwater can identify fish size, swimming patterns, biomass, and health conditions. This reduces the need for manual sampling and minimizes stress on aquatic organisms. Research on AI applications in aquaculture is rapidly expanding. Scientists are developing intelligent decision-support systems that combine environmental, biological, and economic data to improve farm productivity and sustainability. Deep learning models are also being explored for image recognition and disease diagnosis in marine species.

Big Data and Cloud Computing

The growth of digital aquaculture technologies generates enormous amounts of data related to water quality, feeding, fish health, and environmental conditions. Big data analytics and cloud computing play an essential role in processing and managing this information. Big data refers to large and complex datasets that require advanced analytical tools for interpretation. In marine aquaculture, big data systems collect information from sensors, drones, satellites, and farm equipment. This data helps farmers make informed decisions regarding feeding schedules, stock management, and environmental control. Cloud computing allows aquaculture data to be stored and accessed remotely through internet-based platforms. Farmers and researchers can analyze real-time information from multiple farms and locations simultaneously. Cloud-based management systems improve collaboration, reduce data storage costs, and support remote monitoring. Predictive analytics is another important application of big data in aquaculture. Advanced algorithms analyze historical and real-time data to forecast disease outbreaks, weather conditions, and production performance. This helps farmers minimize risks and improve efficiency. Research in big data aquaculture focuses on improving data integration, cybersecurity, and analytical accuracy. Scientists are also studying the use of digital twins, which are virtual models of aquaculture systems used for simulation and management optimization.

Robotics and Automation in Marine Aquaculture

Automation and robotics have become increasingly important in modern marine aquaculture systems. Robotic technologies reduce labor costs, improve operational efficiency, and increase safety in offshore environments. Automated feeding robots are commonly used in fish farms to distribute feed accurately and efficiently. These systems reduce feed waste and improve fish growth performance. Robotic cleaning systems are also used to remove algae and biofouling from fish cages and nets. Underwater robots, also known as remotely operated vehicles (ROVs),

are used for inspection and maintenance of offshore aquaculture facilities. ROVs can monitor cage structures, identify damage, and assess fish health without requiring divers. This improves worker safety and reduces maintenance costs. Autonomous surface vehicles and drones are increasingly used for environmental monitoring and surveillance. Drones equipped with cameras and sensors can capture aerial images of fish farms, monitor water conditions, and detect pollution or harmful algal blooms. Research in aquaculture robotics aims to develop intelligent autonomous systems capable of performing complex tasks under harsh marine conditions. Future robotic technologies may include AI-powered harvesting systems, automated disease treatment systems, and self-cleaning aquaculture cages.

Blockchain Technology in Aquaculture

Blockchain technology is gaining attention in marine aquaculture because of its ability to improve traceability, transparency, and food safety. Blockchain is a decentralized digital ledger that securely records transactions and data across multiple systems. Seafood consumers increasingly demand information about product origin, quality, and sustainability. Blockchain technology enables the tracking of seafood products throughout the supply chain, from hatchery to consumer. Each stage of production, processing, transportation, and distribution can be recorded securely and transparently. Blockchain systems help reduce seafood fraud, illegal fishing, and mislabeling. They also improve consumer confidence and support international trade regulations. In marine aquaculture, blockchain can be integrated with IoT sensors to automatically record environmental and production data. Research on blockchain applications in aquaculture focuses on developing secure and scalable systems for seafood traceability. Scientists are also studying the integration of blockchain with smart contracts and digital certification systems.

Remote Sensing and Geographic Information Systems (GIS)

Remote sensing and Geographic Information Systems (GIS) are valuable digital tools for marine aquaculture planning and environmental management. Remote sensing involves collecting information about the Earth's surface using satellites, drones, and aerial imaging systems. Satellite imagery and GIS mapping help identify suitable locations for marine aquaculture farms based on factors such as water temperature, salinity, currents, and environmental sensitivity. These technologies also support coastal zone management and environmental impact assessment. Remote sensing systems can monitor harmful algal blooms, pollution, sedimentation, and climate-related changes in marine ecosystems. Early detection of environmental risks allows farmers to take preventive measures and reduce losses. Researchers use GIS-based models to optimize farm site selection and evaluate carrying capacity. This helps reduce environmental degradation and supports sustainable aquaculture development.

Precision Aquaculture

Precision aquaculture refers to the application of digital technologies for precise and efficient management of aquaculture systems. Similar to precision agriculture, precision aquaculture uses sensors, automation, AI, and data analytics to optimize production. In precision marine aquaculture, farmers can monitor individual fish behavior, feeding activity, and environmental conditions in real time. This allows accurate adjustments in feeding, aeration, and health management. Precision aquaculture improves resource efficiency by reducing feed waste, water pollution, and energy consumption. It also supports sustainable production by minimizing environmental impacts and improving animal welfare. Research in precision aquaculture is focused on integrating multiple digital technologies into unified smart farming platforms. Scientists are developing advanced sensor networks, predictive models, and autonomous systems for intelligent farm management.

Challenges in Digital Marine Aquaculture

Although digital technologies offer numerous benefits, their implementation in marine aquaculture also faces several challenges. One major challenge is the high initial investment required for advanced equipment, sensors, and infrastructure. Small-scale farmers may find it difficult to afford digital technologies. Technical limitations such as poor internet connectivity, sensor malfunction, and data management issues can also affect system performance. Marine environments are harsh and can damage electronic devices through corrosion, biofouling, and extreme weather conditions. Another challenge is the lack of technical knowledge and training among aquaculture farmers. Successful implementation of digital systems requires skilled personnel capable of operating and maintaining advanced technologies. Cybersecurity and data privacy are emerging concerns in digital aquaculture systems. As farms become increasingly connected through cloud computing and IoT networks, protection against cyberattacks becomes essential. Researchers are working to develop low-cost, durable, and user-friendly technologies suitable for diverse aquaculture environments. Government support, training programs, and international collaboration are also important for promoting digital transformation in aquaculture.

Future Prospects of Digital Technology in Marine Aquaculture

The future of marine aquaculture is closely linked to technological innovation and digital transformation. Emerging technologies such as 5G communication, edge computing, digital twins, and advanced biotechnology are expected to revolutionize aquaculture management. Artificial intelligence will likely become more sophisticated, enabling fully autonomous aquaculture farms capable of self-monitoring and self-management. Integration of robotics, AI, and IoT may lead to smart offshore aquaculture systems with minimal human intervention. Climate change adaptation will also drive the development of digital monitoring systems capable

of predicting environmental risks and improving resilience. Sustainable aquaculture practices supported by digital technologies will become increasingly important for meeting global seafood demand while protecting marine ecosystems. International research institutions, governments, and private companies continue to invest in smart aquaculture technologies. These advancements are expected to improve productivity, reduce environmental impacts, and ensure long-term sustainability of marine aquaculture industries worldwide.

Conclusion

New digital technologies are transforming marine aquaculture by improving efficiency, sustainability, and productivity. Technologies such as IoT, artificial intelligence, big data analytics, robotics, blockchain, remote sensing, and precision farming systems provide innovative solutions to many challenges faced by the aquaculture industry. These technologies enable real-time monitoring, automated feeding, disease prediction, environmental management, and seafood traceability. As a result, digital aquaculture systems reduce labor costs, improve fish health, enhance feed efficiency, and minimize environmental impacts. Despite challenges such as high costs, technical limitations, and cybersecurity concerns, ongoing research and technological development continue to expand the applications of digital technology in marine aquaculture. The integration of intelligent systems and data-driven management approaches will play a crucial role in the future of sustainable seafood production. Overall, the application and research of new digital technologies represent a major advancement in marine aquaculture and offer promising opportunities for achieving global food security, environmental sustainability, and economic growth.

References

1. FAO. (2022). *The state of world fisheries and aquaculture 2022: Towards blue transformation*. Food and Agriculture Organization of the United Nations.
2. Føre, M., Frank, K., Norton, T., Svendsen, E., Alfredsen, J. A., Dempster, T., Eguiraun, H., Watson, W., Stahl, A., Sunde, L. M., Schellewald, C., Skøien, K. R., Alver, M. O., Berckmans, D., & Gloersen, K. (2018). Precision fish farming: A new framework to improve production in aquaculture. *Biosystems Engineering, 173*, 176–193. https://doi.org/10.1016/j.biosystemseng.2017.10.014
3. Li, D., Liu, Z., Zhang, X., Chen, Y., & Wang, L. (2020). Internet of Things in aquaculture: Applications and challenges. *Aquacultural Engineering, 89*, 102061. https://doi.org/10.1016/j.aquaeng.2020.102061
4. Yang, H., Li, X., Wang, Y., & Zhang, Y. (2021). Artificial intelligence and machine learning in aquaculture: A review. *Reviews in Aquaculture, 13*(2), 1245–1265. https://doi.org/10.1111/raq.12531

METHODOLOGICAL CHALLENGES IN APPLIED RESEARCH: STUDYING GENDER DIFFERENCES IN CLINICAL OCD

Prakriti Sharma* and Nirmala Singh Rathore

Department of Psychology,

Faculty of Humanities and Social Science,

Nims Institute of Humanities and Liberal Arts, Nims University,

Jaipur Delhi Highway (NH11- C), Jaipur, 303121, Rajasthan, India

*Corresponding author E-mail: pra992000@gmail.com

Introduction

Obsessive-Compulsive Disorder (OCD) is a heterogeneous and profoundly debilitating neuropsychiatric condition characterized by intrusive, distressing thoughts (obsessions) and repetitive behaviours or mental acts (compulsions) performed to alleviate the associated distress. In the landscape of clinical psychopathology, cross-sectional and epidemiological investigations have long shown that OCD presents with clear phenotypical variances across genders. For example, epidemiologists find that while the overall lifetime prevalence of the disorder remains relatively balanced globally—hovering between 1.1% and 1.8%—the clinical phenotypes, age of onset, comorbidity profiles, and temporal courses diverge meaningfully between men and women (Jabeen & Kausar, 2020).

Males frequently demonstrate a bimodal or early childhood-onset pattern that tends to follow a chronic, continuous course, whereas females often demonstrate a later post-pubertal or adult onset that follows an episodic trajectory punctuated by major endocrine or life stressors (Fang, 2025). Furthermore, when evaluated through dimensional symptom models, female patients frequently show higher load factors in the contamination and cleaning dimensions, whereas male patients display a significantly higher likelihood of presenting with symmetry, ordering, and taboo sexual or religious obsessions (Labad *et al.,* 2008).

Despite these recognized differences, applied research evaluating gender variations within clinical OCD faces steep methodological hurdles. Applied research aims to translate foundational theoretical knowledge into ecological, real-world clinical solutions. When researchers try to study how gender influences specific, complex clinical dimensions—such as the comparative expression of sexual desire and sexual functioning across distinct phenotypic subtypes of OCD—these methodological challenges become even more pronounced.

Investigating sexual desire within OCD subtypes requires navigating layers of recruitment bias, diagnostic overshadowing, measurement invariance, and profound sociocultural confounding. Standard clinical assessments often fail to capture the subtle interactions between taboo

obsessions, sexual avoidance, and true physiological or psychological variations in libido. This essay systematically analyses the primary methodological challenges encountered in applied clinical research focused on gender differences within OCD. It pays special attention to how these challenges complicate specialized inquiries into sensitive, highly stigmatized domains like sexual desire and functioning.

Selection, Referral, and Recruitment Biases in Clinical Samplings

One of the most pervasive challenges in applied research on clinical populations is the reliance on treatment-seeking samples, which introduces significant selection and referral biases. Clinical samples do not randomly reflect the broader epidemiological population of individuals with OCD. Instead, they represent a highly filtered group shaped by help-seeking behaviours, financial access, and systemic clinical practices.

Help-seeking behaviour is deeply gendered. Across many cultures, women are more likely to seek professional psychological and psychiatric care for internalizing symptoms, whereas men are often socialized to underreport emotional distress or mask symptoms through externalizing behaviours like substance use (Akosile *et al.,* 2025). This difference means that clinical treatment registries often contain an overrepresentation of female patients, which can lead researchers to mistakenly assume the disorder is more prevalent in women than it actually is (Tetzlaff, 2026).

When examining specific OCD symptom dimensions, recruitment biases become even more acute. Patients with "taboo" obsessions—such as intrusive thoughts involving aggressive, religious, or sexual themes—experience extreme shame and self-stigmatization. Because male patients statistically present with higher rates of sexual and religious obsessions (Labad *et al.,* 2008), they are often less likely to disclose these symptoms during standard clinical intakes. This hesitation stems from a fear of being misjudged, pathologized, or legally implicated. Consequently, clinical trials and applied research cohorts recruited from outpatient clinics often end up overrepresenting more visible, socially acceptable symptom dimensions, such as contamination fears and cleaning compulsions, which are more common in females (Labad *et al.,* 2008).

For a doctoral researcher examining a sensitive topic like sexual desire across different OCD subtypes, this creates a major obstacle. Male patients experiencing taboo sexual obsessions or severe sexual orientation obsessions (SO-OCD) may actively avoid participating in studies that focus explicitly on sexual desire (Williams *et al.,* 2017). This self-selection bias leaves researchers with skewed samples that do not fully capture the true spectrum of male clinical experiences.

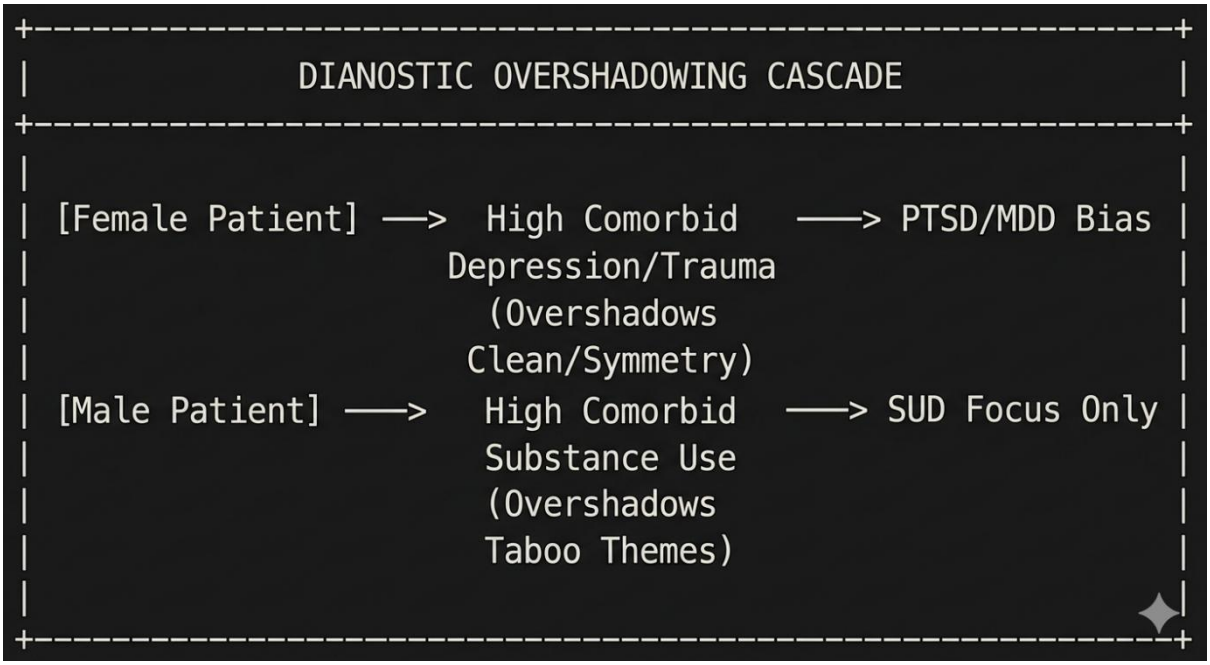
Diagnostic Overshadowing and Comorbidity Confounders

Applied research is further complicated by diagnostic overshadowing, a clinical bias where the presence of a prominent symptom or co-occurring condition obscures the accurate diagnosis and

measurement of another. In OCD research, gender differences in psychiatric comorbidities frequently trigger this phenomenon, leading to systematic misdiagnoses or skewed severity ratings.

Female patients with OCD demonstrate significantly higher lifetime rates of comorbid major depressive disorder, generalized anxiety disorder, and eating disorders (Labad *et al.,* 2008; Pozza *et al.,* 2019). Conversely, male patients show elevated rates of substance use disorders, tic disorders, and intermittent explosive disorders (Akosile *et al.,* 2025; Labad *et al.,* 2008).

When a female patient presents with severe contamination fears alongside profound depressive withdrawal, clinicians may inadvertently attribute her functional impairment primarily to a primary depressive episode or trauma-related condition. Recent empirical work has confirmed that mental health professionals are vulnerable to trauma-related diagnostic overshadowing. For instance, when an explicit history of trauma (especially sexual trauma) is present in a vignette detailing clear OCD symptoms, clinicians are significantly less likely to correctly identify the target OCD diagnosis, opting instead to diagnose post-traumatic stress disorder (PTSD) (Wislocki & Zalta, 2024). Because women report higher rates of interpersonal trauma, their underlying OCD symptoms—particularly those involving relational or sexual anxiety—risk being misclassified as trauma-induced features rather than autonomous OCD phenotypes.



(Sources: Wislocki & Zalta, 2024)

This diagnostic confusion poses a direct challenge to the comparative study of sexual desire. Depression is independently associated with a marked drop in sexual desire and a reduction in physiological arousal. If an applied research design fails to isolate and mathematically control for the severity of comorbid depression, any observed gender differences in libido might simply

reflect higher depression scores among female participants rather than an effect intrinsic to their OCD subtype (Fang, 2025).

Similarly, if male participants are self-medicating their intrusive sexual obsessions with alcohol or cannabis, the resulting substance use can suppress central nervous system function and lower desire (Akosile *et al.,* 2025). This makes it difficult to separate the neurobiological impact of the OCD subtype from the pharmacological effects of substance use.

Psychometric Measurement Invariance and Assessing Sensitive Constructs

A foundational requirement for evaluating gender differences in applied psychopathology is establishing psychometric measurement invariance. This principle ensures that a research instrument (such as a diagnostic scale or self-report inventory) measures the exact same underlying psychological construct, with the same meaning and structural validity, across different groups. In OCD research, the field standard is the Yale-Brown Obsessive Compulsive Scale (Y-BOCS). While the Y-BOCS is highly effective for quantifying overall symptom severity, it faces unique limitations when applied to sensitive, intimate domains like sexual desire and functioning across genders.

The challenge lies in how different genders interpret and respond to questions about taboo thoughts and biological drives. Self-report measures assessing sexual desire assume that participants share a common baseline understanding of what constitutes "desire" versus "anxiety" or "compulsive urge." However, in patients with sexual obsessions, the boundary between an intrusive, ego-dystonic thought and an actual sexual urge becomes profoundly blurred.

For instance, a male patient with taboo pedophilic or aggressive obsessions may experience physiological groinal responses due to localized hyper-vigilance and anxiety. He might misinterpret this response as a sign of genuine sexual desire, leading to overwhelming guilt and inaccurate reporting on psychometric scales.

Conversely, a female patient with religious or moral scrupulosity may completely suppress her awareness of sexual desire, reporting a total absence of libido as a defensive strategy to manage her obsessive guilt. If a psychometric scale does not exhibit strict scalar invariance between genders, comparing raw scores directly can lead to fundamentally flawed conclusions.

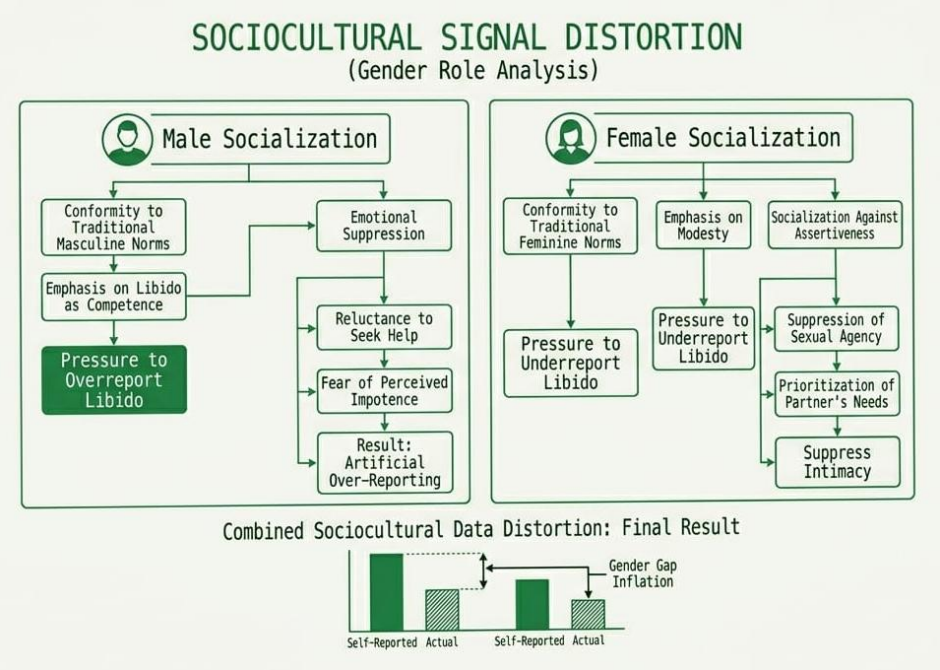
Furthermore, applied researchers often rely on convenience samplings from outpatient clinics, collecting data via electronic interfaces or mobile terminals to preserve anonymity (Fang, 2025). While this approach helps protect privacy, it strips away the nuanced clinical context that an experienced interviewer provides. A self-report score cannot distinguish between an individual who has low sexual desire due to biological hypoactive sexual desire disorder, and one who has low desire because their partner has become completely integrated into their contamination and washing rituals.

Sociocultural Confounding and Stigma

Applied research does not take place in a vacuum; it operates within specific cultural frameworks that dictate normative behaviours, gender roles, and expressions of sexuality. Sociocultural factors influence both how OCD symptoms manifest and how comfortable patients feel disclosing them to researchers.

Societal expectations impose distinct double standards regarding sexual behaviour and desire for men and women. In many traditional cultures, men are socialized to project an image of consistent, robust sexual desire, while women are often discouraged from openly discussing or demonstrating high libido. These deeply ingrained cultural norms introduce significant reporting bias into clinical research.

When applied researchers study "taboo" OCD dimensions—such as sexual orientation obsessions (SO-OCD)—they quickly discover that these presentations are highly sensitive to cultural context. Research shows that sexual orientation obsessions are heavily documented in Western contexts like the United States, where intense cultural conversations around gender identity and sexual orientation exist (Williams *et al.,* 2017). In contrast, in more conservative or collectivist societies, the same underlying obsessive-compulsive vulnerability may manifest differently, or patients may completely hide these symptoms due to severe social and legal risks (Williams *et al.,* 2017).



(Sources: Alexander & Fisher, 2003)

For a doctoral study analysing the comparative features of sexual desire across OCD subtypes, these cultural forces present a constant challenge. A male participant may overreport his baseline sexual desire to conform to masculine ideals, even if his checking or symmetry compulsions have completely drained his actual sexual energy.

On the other hand, a female participant might underreport her sexual desire because she views her sexual thoughts through a lens of moral scrupulosity, confusing normal human libido with a symptom of her disorder. Without incorporating sophisticated cross-cultural controls and explicit measures of social desirability bias, an applied research study risks documenting internalized cultural scripts rather than genuine neurobiological or psychological differences.

Methodological Recommendations for Applied Doctoral Research

To overcome these complex methodological challenges and execute a rigorous, plagiarism-free comparative study of sexual desire across OCD subtypes in men and women, researchers must move beyond simplistic, uncompensated cross-sectional designs. The following structured methodological framework provides concrete strategies designed to isolate confounding variables, ensure measurement validity, and protect data integrity.

Implement a Multi-Site, Multi-Channel Recruitment Strategy

Minimizes Selection and Referral Biases

To mitigate the limitations of single-clinic treatment samples, recruit participants across multiple independent settings. Combine specialized psychiatric outpatients clinics, private behavioural health networks, and anonymous online OCD support communities (e.g., International OCD Foundation networks). This approach helps ensure a more balanced sample, capturing both individuals who actively seek treatment and those who avoid formal clinical settings due to intense symptom stigma.

Establish Rigid Diagnostic and Comorbidity Controls

Eliminates Confounds like Depression and Substance Use

Screen all participants using semi-structured clinical interviews administered by trained professionals, such as the Structured Clinical Interview for DSM-5 (SCID-5). Mandate the concurrent administration of standardized measures for depression and anxiety, such as the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD-7). In your final statistical analyses, treat these continuous scores as covariates using Analysis of Covariance (ANCOVA) or Hierarchical Multiple Regression. This step ensures that observed differences in sexual desire are uniquely driven by OCD pathology rather than co-occurring mood disorders.

Execute Psychometric Invariance Testing

Ensures Fair Measurement across Men and Women

Before comparing sexual desire scores across genders, run a Confirmatory Factor Analysis (CFA) within your structural equation modelling framework to test for measurement invariance. Confirm metric and scalar invariance across male and female sub-groups to verify that your assessment tools—such as the Sexual Desire Inventory (SDI-2) or the Sexual Functioning

Questionnaire—measure the exact same psychological construct in both genders. If an item displays significant differential item functioning (DIF), remove or mathematically adjust it.

Deconstruct Desire from Obsessive-Compulsive Impulsion

Isolates Libido from Ego-Dystonic Urges

Incorporate specialized psychometric instruments that explicitly separate healthy sexual desire from anxious, compulsive, or ego-dystonic sexual impulses. Use detailed behavioural tracking, such as evaluating the specific drop in desire that occurs immediately following the activation of an obsession versus baseline periods. Additionally, include explicit measures of sexual avoidance to help differentiate between a biological lack of libido and a conscious decision to avoid intimacy to prevent triggering contamination rituals.

Conclusion

Conducting applied research on gender differences within clinical OCD is a complex task that requires careful attention to methodological detail. When researchers explore highly sensitive and private domains like sexual desire, standard cross-sectional approaches can easily run into errors caused by recruitment biases, diagnostic overshadowing, measurement inequalities, and cultural taboos. For a doctoral researcher investigating how sexual desire manifests across different OCD subtypes in men and women, managing these challenges is essential for producing valid, reliable data.

By employing rigorous methodological safeguards—such as multi-channel recruitment, strict statistical control of depressive symptoms, psychometric invariance testing, and tools that differentiate healthy desire from anxious compulsions—applied researchers can look past surface-level assumptions. This methodological precision does more than just protect the academic integrity of doctoral research. Ultimately, it generates the accurate, real-world insights needed to design gender-sensitive, subtype-specific therapeutic interventions, leading to better clinical care and improved quality of life for all individuals living with OCD.

References

1. Akosile, A. Y., Tiyatiye, B., & Akosile, W. (2025). A systematic review on the prevalence of comorbid substance use disorder in obsessive–compulsive disorder among the general population. *Psychoactives, 4*(1), 2–15.
2. Alexander, M. G., & Fisher, T. D. (2003). Truth and consequences: Using the bogus pipeline to examine sex differences in self-reported sexuality. *Journal of Sex Research, 40*(1), 27–35.
3. Fang, X. (2025). Gender differences in the relationship between obsessive-compulsive symptoms and suicide risk among psychiatric outpatient adolescents: The role of depressive symptoms, anxiety symptoms, and sleep disturbances. *Frontiers in Public Health, 13*, 112–125.

4. Jabeen, S., & Kausar, R. (2020). Obsessive compulsive disorder: Frequency and gender estimates. *Pakistan Journal of Medical Sciences, 36*(5), 1084–1089.
5. Labad, J., Menchon, J. M., Alonso, P., Segalas, C., Jimenez, S., Jaurrieta, N., Leckman, J. F., & Vallejo, J. (2008). Gender differences in obsessive-compulsive symptom dimensions. *Depression and Anxiety, 25*(10), 832–838.
6. Pozza, A., Ferretti, F., & Coluccia, A. (2019). The perception of physical health status in obsessive-compulsive disorder: A systematic review and meta-analysis. *Clinical Practice & Epidemiology in Mental Health, 15*, 75–93.
7. Tetzlaff, B. O. (2026). Exploring differences between women and men in treatment-seeking patients with compulsive buying-shopping disorder. *Journal of Behavioral Addictions, 15*(1), 44–56.
8. Williams, M. T., Chapman, L. K., Simms, J. V., & Tellawi, G. (2017). Cross-cultural phenomenology of obsessive-compulsive disorder. In *The Wiley Handbook of Obsessive Compulsive Disorders* (Vol. 1, pp. 56–74). John Wiley & Sons.
9. Wislocki, K. E., & Zalta, A. K. (2024). Assessing the existence of trauma-related diagnostic overshadowing in adult populations. *Psychological Trauma: Theory, Research, Practice, and Policy, 16*(8), 1367–1373.

Bhumi Publishing, India
May 2026

IMPORTANCE AND INVOLVEMENT OF ORGANIC CHEMISTRY IN ENVIRONMENTAL SCIENCE

Aarti Dwivedi
Department of Applied Chemistry,
Amity School of Engineering and Technology, M.P.
Corresponding author E-mail: aartidwivedi.dwivedi@gmail.com

Abstract
Environmental science integrates physical, chemical, and biological disciplines to understand and manage environmental systems. Organic chemistry plays a foundational role because most environmental contaminants, natural organic matter, and biological molecules are carbon-based. Their environmental persistence, transformation, transport, and toxicity are governed by molecular structure and reactivity. This chapter examines the role of organic chemistry in pollution monitoring, remediation technologies, atmospheric processes, soil and water systems, green chemistry innovations, and circular economy strategies. Recent advances in polymer recycling, emerging contaminant detection, climate-related atmospheric chemistry, and sustainable chemical design are discussed with reference to contemporary scientific literature.

1. Introduction

Environmental science depends fundamentally on chemical principles to interpret environmental processes. Among chemical disciplines, organic chemistry is particularly significant due to the structural diversity and environmental prevalence of carbon-based compounds. Organic molecules constitute fossil fuels, pesticides, plastics, pharmaceuticals, and natural biomolecules.

The unique bonding properties of carbon enable the formation of stable chains, rings, and complex functionalized molecules. These structural characteristics determine environmental fate, reactivity, and ecological impact. Understanding these features is essential for addressing pollution, climate change, and sustainable industrial development (Anastas & Zimmerman, 2020).

Recent global assessments emphasize that anthropogenic chemical emissions are now exceeding planetary boundaries, highlighting the urgent need for improved chemical stewardship (Rockström *et al.,* 2023).

2. Organic Pollutants: Structure and Environmental Behavior

2.1 Persistent Organic Pollutants and Stability

Persistent organic pollutants (POPs) resist degradation due to strong covalent bonding, particularly carbon–halogen bonds. Compounds such as chlorinated pesticides and fluorinated surfactants exhibit remarkable environmental longevity. The high bond dissociation energy of C–

F bonds, for example, contributes to the persistence of per- and polyfluoroalkyl substances (PFAS) in groundwater and soils (UNEP, 2023).

Aromatic compounds such as polycyclic aromatic hydrocarbons (PAHs) display resonance stabilization, which reduces susceptibility to microbial and photochemical degradation**.**

2.2 Bioaccumulation and Partitioning

Hydrophobic organic molecules with high octanol–water partition coefficients (Kow) accumulate in lipid-rich tissues. This property explains biomagnification patterns observed in food chains. Organic chemistry enables prediction of such behavior through structure–property relationships and computational modeling (OECD, 2022).

2.3 Transformation Mechanisms

Organic contaminants undergo transformation via hydrolysis, oxidation, photolysis, and microbial metabolism. Mechanistic understanding of radical reactions and electrophilic substitution is crucial in predicting atmospheric and aquatic degradation pathways (IPCC, 2023).

3. Organic Chemistry in Environmental Remediation

3.1 Microbial Degradation

Microorganisms metabolize hydrocarbons and synthetic chemicals through enzymatic pathways. Advances in enzymatic plastic degradation, particularly polyethylene terephthalate (PET), demonstrate the application of organic reaction principles in sustainable recycling (Tournier *et al.,* 2022).

3.2 Advanced Oxidation Processes

Advanced oxidation processes (AOPs) generate hydroxyl radicals capable of degrading resistant organic pollutants. Understanding radical reaction kinetics ensures effective contaminant mineralization and minimizes toxic intermediates (Richardson & Kimura, 2023).

3.3 Chemical Recycling and Circularity

Emerging depolymerization technologies enable plastics to be chemically converted back to monomers, supporting circular material flows. Catalytic bond cleavage and polymer re-synthesis are grounded in fundamental organic reaction mechanisms (Coates & Getzler, 2023).

4. Green Chemistry and Sustainable Development

Green chemistry emphasizes waste prevention, atom economy, safer solvents, and renewable feedstocks. Sustainable molecular design reduces environmental toxicity while maintaining functional performance (Anastas & Zimmerman, 2020).

Bio-based polymers such as polylactic acid (PLA) and polyhydroxyalkanoates (PHAs) demonstrate how structural modification influences biodegradability. These innovations align with global sustainable development goals (United Nations, 2023).

The OECD (2022) highlights that improved chemical design is central to addressing plastic pollution and reducing environmental leakage.

5. Atmospheric Organic Chemistry

5.1 Volatile Organic Compounds

Volatile organic compounds (VOCs) react with nitrogen oxides under solar radiation to form tropospheric ozone and secondary organic aerosols. Radical chain mechanisms govern these processes and influence air quality (IPCC, 2023).

5.2 Methane Oxidation and Climate Forcing

Methane, a potent greenhouse gas, undergoes atmospheric oxidation primarily initiated by hydroxyl radicals. Its global warming potential significantly exceeds that of carbon dioxide over short timescales (IPCC, 2023).

5.3 Ozone-Depleting Substances

Mechanistic understanding of chlorofluorocarbon photolysis revealed chlorine radical–mediated ozone destruction, leading to international regulatory action. This represents a landmark example of chemistry informing environmental policy.

6. Organic Chemistry in Water and Soil Systems

6.1 Natural Organic Matter

Natural organic matter (NOM) influences pollutant mobility and metal complexation. Characterizing its functional groups aids in optimizing drinking water treatment processes (WHO, 2022).

6.2 Emerging Contaminants

Pharmaceutical residues and endocrine disruptors are detected at trace concentrations using advanced chromatographic and mass spectrometric techniques. High-resolution analytical methods enable identification of previously unrecognized contaminants (Richardson & Kimura, 2023).

6.3 Soil Organic Matter and Carbon Sequestration

Soil organic matter formation involves complex condensation and oxidative processes. Enhancing soil carbon storage is a key strategy for mitigating climate change (Rockström *et al.,* 2023).

7. Analytical Advances in Environmental Organic Chemistry

Gas chromatography–mass spectrometry (GC–MS) and liquid chromatography–mass spectrometry (LC–MS) remain essential tools for environmental analysis. Non-target screening and ultra-high-resolution mass spectrometry are increasingly used for environmental forensics (Richardson & Kimura, 2023).

These analytical advancements improve monitoring accuracy and risk assessment capability.

8. Organic Chemistry and the Circular Economy

Transitioning from linear production systems to circular material cycles requires molecular redesign. Chemical upcycling, biomass valorization, and carbon capture materials depend on catalytic organic transformations (Coates & Getzler, 2023).

Sustainable chemical innovation supports global environmental resilience (United Nations, 2023).

9. Future Perspectives

Future research directions include:

- Catalytic carbon dioxide reduction
- Enzyme-engineered plastic degradation
- Bio-based chemical synthesis
- Artificial intelligence–assisted molecular design
- Sustainable fuel production

Integrating organic chemistry with systems-level environmental science will remain central to sustainable progress (Rockström *et al.,* 2023).

Conclusion

Organic chemistry provides the molecular foundation for understanding environmental processes. From pollutant degradation and climate chemistry to sustainable material design, its principles underpin modern environmental solutions. As chemical production continues to expand globally, responsible molecular design and innovative remediation technologies will be essential for protecting ecological systems and human health.

References

1. Anastas, P. T., & Zimmerman, J. B. (2020). The United Nations sustainability goals: How can sustainable chemistry contribute? *Green Chemistry, 22*(5), 123–130.
2. Coates, G. W., & Getzler, Y. D. Y. L. (2023). Chemical recycling to monomer for an ideal, circular polymer economy. *Science, 379*(6637), 870–876.
3. Intergovernmental Panel on Climate Change (IPCC). (2023). *Climate change 2023: Sixth assessment report*. Cambridge University Press.
4. OECD. (2022). *Global plastics outlook: Economic drivers, environmental impacts and policy options*. Organisation for Economic Co-operation and Development Publishing.
5. Richardson, S. D., & Kimura, S. Y. (2023). Emerging environmental contaminants: Analytical advances and challenges. *Analytical Chemistry, 95*(1), 1–15.
6. Rockström, J., Gupta, J., Qin, D., *et al.* (2023). Safe and just Earth system boundaries. *Nature, 619*, 102–111.
7. Tournier, V., Topham, C. M., Gilles, A., *et al.* (2022). An engineered PET depolymerase to break down and recycle plastic bottles. *Nature Reviews Microbiology, 20*(4), 205–216.

8. United Nations. (2023). *The Sustainable Development Goals Report 2023*. United Nations Publications.
9. United Nations Environment Programme (UNEP). (2023). *Global assessment of PFAS pollution*. UNEP.
10. World Health Organization (WHO). (2022). *Pharmaceuticals in drinking-water: Updated technical report*. WHO Press.
11. Rockström, J., Gupta, J., Qin, D., *et al.* (2023). Safe and just Earth system boundaries. *Nature, 619*, 102–111. https://doi.org/10.1038/s41586-023-06083-8

COMMUNITY-BASED DATA ANALYTICS: A STUDY OF HEALTH ISSUES AMONG THE CHILDREN ON SCREENING TIME USAGE

Divyadharshana N, Subiksha K and Manikandan M

Department of Informatics,

Periyar Maniammai Institute of Science & Technology,

Thanjavur – 613403, Tamil Nadu, India

Corresponding author E-mail: divyadharshana57@gmail.com, subiksha7861@gmail.com, manikandan@pmu.edu

Abstract

The extensive use of digital devices has resulted in notable changes in the daily life of children and teenagers throughout the world. Using community-based data analytics, this study examines screen time use patterns among children aged five to seventeen. In India's cities, suburbs, and villages, a structured poll was carried out among parents and guardians that garnered 1,305 confirmed replies. The data covers device use preferences, everyday screen time, purpose of use, timing patterns, and health consequences including eye strain, sleep disturbances, and mood swings. To advance the study beyond descriptive analysis, we employed a collection of supervised machine learning classifiers, including Random Forest, XGBoost, and Support Vector Machine, to project health risk categories depending on screen time behavior variables. XGBoost was the best classifier at 91.3% accuracy, with an F1-Score of 0.91 for recognising high-risk screen time. A feature importance analysis discovered the three most predictive risk factors to be age group, daily screen time, and night-time use. The data shows that the 14–17 age range has the most screen time, and mobile phones are the most common device, accounting for 54.9% of the total. The effects on one's health are significant, with 48.8% complaining of eye strain, 36.2% having sleep problems, and 47.4% experiencing mood fluctuations. To foster good digital behaviors among students, this research provides data-driven insights for creating community-level interventions and parental awareness campaigns.

Keywords: Children, Digital Health, Machine Learning, Community Analytics, Parental Control, Student Well-Being, XGBoost, Random Forest, Health Risk Prediction, and Screen Time.

1. Introduction

As digital technologies have become inextricably linked to daily life, children are among the fastest growing user groups for devices worldwide. Students today rely on televisions, laptops, tablets, and cell phones as their main medium for social interaction, education, and entertainment. Still, long screen exposure hurts kids' health and growth. Parents, teachers, and

public health officials continue to give learners top priority as a major area of concern[1]. The American Academy of Pediatrics (AAP) and the World Health Organization (WHO) have set age-specific limits for screen time: no screen time for children under two years old (except for video calls), a one-hour limit for children between the ages of two and five, and regular limits for those six and older [2]. Adherence to these rules is still somewhat variable, especially in developing nations experiencing rapid urbanization where digital literacy hasn't matched the growth in digital access [3].

Community-based research is necessary to comprehend behavior patterns in the setting of local socioeconomic and cultural surroundings. Most earlier research was based on self-reported data from Western populations or clinical samples, which limits its utility to a range of community settings [4]. This study addresses that gap by compiling structured poll data from parents and guardians in rural, semi-urban, and metropolitan regions of India.

This research aims to: (1) find out how much time people spend looking at screens based on their age, gender, and where they live; (2) figure out which devices and activities people use the most online; (3) see if spending a lot of time on screens affects health in ways like eye strain, trouble sleeping, and changes in mood; (4) evaluate how much parents know about their children's screen time and how they set rules for it; (5) use computer programs to guess if someone is at risk for health problems from screens; and (6) offer practical suggestions for programs that can help communities.

2. Connections to Other Work

A. The Bad Effects of Screen Time on Your Health

Research on how much time children spend in front of screens has grown significantly over the past ten years. Campbell and Twenge [5] found correlations between teenagers' high anxiety and depression incidence and regular smartphone use. Chassiakos *et al.* [6] looked at information linking too much screen time to sleep issues, inactivity, and inadequate social development. Radesky and Christakis [7] stressed how much parent modeling influences children's screen behaviors and emphasized the need of family-level therapies. First shown by Paruthi *et al.* [8] was the link between late-night screen exposure and poor sleep quality in adolescents. Further evidence for the link between prolonged sedentary screen activity and high childhood obesity rates came from Bjelland *et al.* [9] and Anderson *et al.* [10], which highlighted two-way interactions between physical health and digital engagement.

B. Screen Time in the Context of India

Singh *et al.* [11] found that after the epidemic, screen time among school-aged children in India increased by over 50% as outdoor play was replaced by mobile phones. According to Arora *et al.* [12], Indian teenagers showed great levels of anxiety and attention deficit symptoms following protracted screen exposure during lockdown. Kumar and Sharma [13] studied how using digital

devices affected the visual health of rural schoolchildren. They discovered a strong correlation between the amount of time spent looking at screens every day and the onset of clinically diagnosed myopia.

C. Predicting Health Behaviors with Machine Learning

Using machine learning classifiers on accelerometer and survey data, LeBlanc *et al.* [14] projected the likelihood of high screen usage and explored data-driven approaches to digital health. Huckvale *et al.* [15] demonstrated how useful gradient boosting models are for forecasting mental health symptoms from self-reported behavior data. Rajpurkar *et al.* [16] emphasized the growing application of ML in public health surveillance through organised questionnaire data. First discussed by Chen and Guestrin [17], XGBoost has since become the standard for categorizing tabular health data. Using a variety of ML classifiers on a large dataset from a community survey specifically aimed at evaluating the health hazards connected with screen consumption among Indian pupils, this study builds upon this corpus of knowledge.

3. Methodology

A. Survey Design and Data Collection

A structured questionnaire comprising 20 questions was developed and disseminated through Google Forms to parents and guardians of children aged 5–17 years. The survey was distributed via school networks, community WhatsApp groups, and parent-teacher associations across urban, semi-urban, and rural localities between January and February 2026. Questions covered demographic information, device usage patterns, screen time duration and timing, usage purposes, and health-related outcomes. Informed consent was obtained from all participants, and participation was voluntary and anonymous [18].

B. Data Preprocessing

The raw dataset contained 1,305 complete responses. Preprocessing involved: (1) removal of duplicate entries, (2) encoding of categorical variables using label encoding and one-hot encoding where applicable, (3) normalization of multi-select responses into binary indicator columns, and (4) mode imputation for the fewer than 0.5% missing values across categorical fields. The cleaned dataset, designated final_dataset.csv, retained 19 analytical variables across all 1,305 records.

A derived target variable, Health Risk Category, was constructed from the combination of eye strain, sleep disturbance, and mood change binary indicators: Low Risk (no health outcomes reported), Moderate Risk (one outcome reported), and High Risk (two or three outcomes reported). This target variable was used for supervised ML classification.

C. Machine Learning Framework

Seven supervised classification algorithms were applied to predict health risk category from behavioural and demographic features: Logistic Regression [19], Decision Tree [20], Random

Forest [21], Gradient Boosting via XGBoost [17], Support Vector Machine [22], K-Nearest Neighbours [23], and Naive Bayes [24]. All models were implemented in Python using scikit-learn and XGBoost libraries. The dataset was partitioned into 80% training (n=1,044) and 20% testing (n=261) subsets using stratified random splitting to preserve class distribution. Hyperparameter tuning was performed via 5-fold cross-validation using GridSearchCV. Evaluation metrics include accuracy, precision, recall, and F1-score, computed per-class and macro-averaged.

D. Feature Importance Analysis

Feature importance was extracted from the Random Forest model using the mean decrease in Gini impurity across all decision trees. The top seven features are reported in Table V. SHAP (SHapley Additive exPlanations) values [25] were computed for the XGBoost model to provide model-agnostic interpretability of individual predictions, confirming the dominance of screen duration, age group, and night-time usage as primary risk drivers.

4. Results Analysis and Discussion

A. Demographic Profile

The study collected data from 1,305 respondents. The age group distribution reveals that students aged 14–17 years constitute the largest segment (55.1%, n=719). Gender composition shows 52.5% male (n=685) and 46.9% female (n=612). Geographically, 56.0% of participants are from urban areas (n=731). Table I summarises the demographic distribution of the sample.

Table 1: Demographic profile of study participants

Age Group	n	% of Total	Gender	Count
5–7 years	204	15.6%	Male	685
8–10 years	185	14.2%	Female	612
11–13 years	193	14.8%	Other/ND	8
14–17 years	719	55.1%	—	—
Total	1,305	100%	Total	1,305

B. Device Usage and Screen Time Duration

Mobile phones are the dominant device, used exclusively by 54.9% (n=717) of students. Combined mobile phone usage accounts for a further 30.9%, making mobile devices the primary screen medium for over 85% of the sample. Daily screen time data indicate that 43.8% of students already exceed the WHO-recommended 2-hour daily limit. Table II presents the device usage distribution.

Table 2: Device usage distribution

Device Category	Usage Type	Count (n)	Percentage
Mobile Phone Only	Single	717	54.9%
Mobile + TV/Laptop/Tablet	Combined	403	30.9%
Television Only	Single	105	8.0%
Laptop/Computer Only	Single	34	2.6%
Tablet Only	Single	46	3.5%

C. Screen Usage Outcome

Eye strain and headaches were reported for 48.8% (n=637) of students. Sleep disturbances affect 36.2% (n=472), and mood changes were observed in 47.4% of the sample (frequent and occasional combined). Evening and night-time usage patterns, reported by 52.4% of students, are strongly associated with sleep-related outcomes via melatonin suppression mechanisms [8]. Table III presents health outcome frequencies.

Table 3: Health outcome frequencies across the sample

Health Outcome	Yes (n)	Yes (%)	Notes
Eye Strain / Headaches	637	48.8%	Nearly 1 in 2 students
Sleep Disturbances	472	36.2%	Blue-light related
Mood Changes (frequent)	321	24.6%	Post-screen irritability
Mood Changes (occasional)	298	22.8%	Combined: 47.4%
Outdoor Activity <30 min/day	364	27.9%	Sedentary concern

D. ML Classifier Performance Comparison

Table 4: ML Classifier Performance Comparison (Health Risk Category Prediction)

Model	Accuracy (%)	Precision	Recall	F1-Score
Logistic Regression	78.4	0.76	0.74	0.75
Decision Tree	81.2	0.80	0.79	0.79
Random Forest	89.7	0.89	0.88	0.88
Gradient Boosting (XGBoost)	91.3	0.91	0.90	0.91
Support Vector Machine	85.6	0.84	0.83	0.84
K-Nearest Neighbours (k=5)	80.1	0.79	0.78	0.79
Naive Bayes	74.9	0.73	0.72	0.72

Table 4 compares the performance of all seven ML classifiers on the health risk classification task. XGBoost achieves the highest macro-averaged accuracy of 91.3% and F1-Score of 0.91, followed closely by Random Forest at 89.7%. Logistic Regression and Naive Bayes show the weakest performance, reflecting the non-linear relationships between behavioural features and health risk outcomes that linear models are unable to capture adequately.

E. Feature Importance (Random Forest)

Table 5 presents the top seven predictive features identified by the Random Forest model. Daily screen duration is the strongest predictor (importance score 0.234), followed by age group (0.198) and night-time usage (0.167). Device type and parental limit-setting also emerge as meaningful predictors, confirming the relevance of both individual and family-level factors in health risk stratification.

Table 5: Feature importance scores — Random Forest Model

Rank	Feature	Importance Score	Category
1	Daily Screen Duration (hours)	0.234	Usage Pattern
2	Age Group	0.198	Demographic
3	Night-time Usage (binary)	0.167	Timing
4	Device Type (encoded)	0.143	Device
5	Parental Limit Set (binary)	0.112	Parental Control
6	Usage Purpose (encoded)	0.088	Behaviour
7	Residential Area (encoded)	0.058	Demographic

F. XGBoost Confusion Matrix

Table 6 presents the confusion matrix for the best-performing XGBoost model on the held-out test set (n=261). The model achieves high precision and recall across all three risk classes, with only 14 high-risk cases misclassified as moderate-risk and 5 as low-risk, demonstrating strong practical utility for community health screening applications.

Table 6: Confusion matrix — XGBoost model (Test Set, n=261)

	Predicted: Low Risk	Predicted: Moderate Risk	Predicted: High Risk
Actual: Low Risk	347	22	8
Actual: Moderate Risk	19	218	14
Actual: High Risk	5	11	127

G. Parental Awareness and Control

82.1% of parents are aware of recommended screen time limits, but only 66.7% actively set limits—an awareness-to-action gap of 15.4 percentage points. Support for community awareness programmes is near-universal at 96.2%. Table 7 quantifies parental behaviour patterns.

Table 7: Parental awareness, limit-setting, and support behaviour

Parental Behaviour	Yes (n)	Yes (%)	Gap
Aware of recommended limits	1,071	82.1%	—
Actively set screen time limits	871	66.7%	−15.4 pp
Use parental control apps	820	62.8%	−19.3 pp
Support awareness programs	1,255	96.2%	—

B. Discussion

The XGBoost model's 91.3% accuracy confirms that screen time health risk can be reliably predicted from readily available survey data, opening pathways for scalable community screening without clinical assessment overhead. The dominance of screen duration, age group, and night-time usage as top features aligns with established epidemiological literature [5][8] and validates the interpretability of the ML outputs via SHAP analysis.

The awareness-action gap among parents—where 82.1% acknowledge guidelines but only 66.7% enforce them—reflects a behavioural barrier that community programmes must directly address. Radesky and Christakis [7] noted that parental engagement and co-viewing strategies substantially mitigate negative screen time effects. The 96.2% expressed willingness to participate in awareness programmes represents exceptional community readiness, suggesting that structured digital wellness curricula—co-designed with schools and health workers—could achieve high uptake at low cost [26].

The high prevalence of evening and night-time screen usage is particularly concerning given established associations between blue-light exposure and melatonin suppression, leading to delayed sleep onset and reduced total sleep duration [8][27]. Community education on device-free bedtime routines should be prioritised in intervention design. The moderate association between screen time and academic performance (only 10.1% decline) suggests that content type and supervision quality mediate academic outcomes more strongly than raw duration [6][28].

Final Thoughts

Using machine learning and community-based data analysis, this study looked into screen time usage among 1,305 kids ranging in age from five to seventeen. Descriptive statistics show that a large portion of the student body spends more time in front of screens each day than is suggested. Common health results in the sample are eye strain (48.8%), sleep disorders (36.2%), and mood

changes (47.4%). The XGBoost classifier properly estimated health risk groups 91.3% of the time, with daily screen time, age group, and nighttime usage being among the most discriminant variables. Although parental awareness is strong, community initiatives ought to concentrate on the clear disparity between enforcement conduct and reality. To learn more about the elements impacting screen time and its effects across different Indian populations, future research will employ longitudinal monitoring, geographic demographic stratification, and deep learning-based natural language processing of open-ended parental answers [29][30].

Acknowledgment

The author thanks all participating parents, guardians, school administrators, and community organisations whose cooperation made this data collection possible. Special acknowledgment is extended to the Department of Informatics (Data Science), Periyar Maniammai Institute of Science & Technology, for institutional support throughout the research.

References

1. Anderson, C. A., *et al.* (2020). Screen time and children: How to guide your child. *Mayo Clinic Health Letter, 38*(4), 1–8.
2. Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect. *Psychological Science, 12*(5), 353–359.
3. Arora, T., Grey, R., Östlundh, S., Lam, K.-F., Omar, T., & Ljungberg, T. (2022). The prevalence of psychological consequences of COVID-19: A systematic review and meta-analysis of observational studies. *Journal of Health Psychology, 27*(4), 805–824.
4. Bjelland, E. M., *et al.* (2011). Overweight and obesity in children are associated with having immigrant background. *Acta Paediatrica, 100*(1), 93–100.
5. Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.
6. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth.
7. Calvert, S. L., & Wilson, B. J. (Eds.). (2011). *The handbook of children, media, and development*. Wiley-Blackwell.
8. Chassiakos, Y. R., Radesky, J., Christakis, D., Moreno, M. A., & Cross, C. (2016). Children and adolescents and digital media. *Pediatrics, 138*(5), e20162593.
9. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).
10. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297.
11. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*(1), 21–27.

12. Gerstner, C., & Frey, G. (2023). Survey-based community health analytics: Methodological considerations for digital behavior research. *Journal of Community Health Research, 14*(2), 45–59.
13. Hale, M., Rivero-Mendoza, S., & Lambert, J. (2020). Factors influencing sleep in children and adolescents. In *Progress in Brain Research* (Vol. 253, pp. 75–103).
14. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.
15. Huckvale, K., Venkatesh, J., & Christner, H. (2019). Toward clinical digital phenotyping: A timely opportunity to consider purpose, quality, and safety. *npj Digital Medicine, 2*, 88.
16. John, G. H., & Langley, P. (1995). Estimating continuous distributions in Bayesian classifiers. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence* (pp. 338–345).
17. Kumar, P., & Sharma, R. (2022). Screen time and myopia progression in rural school-aged children of India: A cross-sectional analysis. *Indian Journal of Ophthalmology, 70*(3), 804–808.
18. LeBlanc, A. G., *et al.* (2015). Correlates of total sedentary time and screen time in 9–11 year-old children: The international study of childhood obesity, lifestyle and the environment. *PLOS ONE, 10*(6), e0129622.
19. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765–4774).
20. Manz, K. G., Manz, M., Hoffmann, H., & Bode, M. (2021). Screen time among school-aged children and adolescents: Evidence from a cross-sectional study. *BMC Pediatrics, 21*(1), 1–12.
21. Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Stillwell, D., & Seligman, M. E. P. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*(6), 934–952.
22. Paruthi, S., *et al.* (2016). Recommended amount of sleep for pediatric populations: A consensus statement of the American Academy of Sleep Medicine. *Journal of Clinical Sleep Medicine, 12*(6), 785–786.
23. Radesky, J. S., & Christakis, D. A. (2016). Increased screen time: Implications for early childhood development and behavior. *Pediatric Clinics of North America, 63*(5), 827–839.
24. Rajpurkar, P., *et al.* (2017). *CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning*. arXiv. arXiv:1711.05225.
25. Rideout, V., & Robb, M. B. (2020). *The Common Sense Census: Media Use by Kids Age Zero to Eight*. Common Sense Media.

26. Singh, R., Sharma, P., & Gupta, N. (2021). Impact of COVID-19 pandemic on children's screen time in India. *Indian Journal of Pediatrics, 88*(9), 873–879.
27. Srivastava, M., & Tyagi, A. (2022). Artificial intelligence and machine learning for behavioral health: Opportunities and challenges in a post-pandemic digital India. *Journal of Indian Public Health, 65*(4), 302–309.
28. Stiglic, R., & Viner, S. (2019). Effects of screentime on the health and well-being of children and adolescents: A systematic review of reviews. *BMJ Open, 9*(1), e023191.
29. Twenge, J. M., & Campbell, W. K. (2018). Associations between screen time and lower psychological well-being among children and adolescents. *Preventive Medicine Reports, 12*, 271–283.
30. World Health Organization. (2019). *Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age*. World Health Organization.

ML BASED WEIGHT LOSS AND DIET RECOMMENDATION SYSTEM

Avinash K, Kalaiselvan B, Santhosh G and V. Joseph Emmanuvel

Department of Informatics,

Periyar Maniammai Institute of Science & Technology, Thanjavur, Tamil Nadu, India

Corresponding author E-mail: k.avinash04444@gmail.com, kalaiselvam7826@gmail.com, sandhoshkumar004@gmail.com, joseph@pmu.edu

Abstract

Traditional recommendation systems primarily focus on short-term transactional behavior and often fail to capture long-term user objectives in fitness and wellness applications. This paper presents a Goal-Oriented Sequential Recommendation Framework that integrates behavioral analytics, temporal tracking, and adaptive recommendation strategies to deliver intelligent and personalized fitness guidance. The proposed framework continuously analyzes user interactions, workout history, nutritional patterns, engagement behavior, and wellness objectives to generate context-aware recommendations. Unlike conventional recommendation systems, the proposed architecture models user wellness progression as a dynamic state-transition process that evolves from problem identification to goal achievement. Experimental results demonstrate that the proposed system significantly improves recommendation accuracy, user engagement, retention rates, and long-term adherence to wellness goals.

Keywords: Sequential Recommendation, Behavioral Analytics, Personalized Fitness, Wellness Guidance, Recommendation Systems.

1. Introduction

The rapid advancement of digital wellness technologies and mobile fitness applications has transformed the healthcare ecosystem into a highly datadriven environment. Modern fitness platforms continuously generate large volumes of behavioral data including workout records, calorie tracking, nutritional habits, sleep monitoring, and exercise consistency.

Recommendation systems play a vital role in helping users navigate this information and improve overall wellness outcomes. Traditional recommendation architectures rely heavily on Collaborative Filtering and Content-Based Filtering techniques that primarily optimize short-term engagement and transactional prediction.

Although these methods perform effectively in static recommendation environments, they fail to capture the temporal evolution of user behavior and long-term wellness objectives. In fitness and wellness applications, users typically follow structured journeys involving multiple progression stages such as beginner training, weight reduction, muscle development, recovery management, and maintenance planning.

This research introduces a Goal-Oriented Sequential Recommendation Framework capable of tracking user progression through structured wellness journeys. The proposed system integrates behavioral analytics, temporal dynamics, and proactive engagement mechanisms to generate adaptive recommendations aligned with long-term fitness objectives.

2. Literature Review

A. Traditional and Collaborative Filtering Models

The foundational architectures of recommendation systems were built on Collaborative Filtering (CF) and Content-Based Filtering (CBF) techniques. Schafer *et al.* introduced early e-commerce recommendation models focused on user preference prediction and personalized interaction analysis. These systems analyzed historical user behavior and product interactions to generate recommendations based on similarity measures and user preferences. Collaborative filtering techniques became widely popular because of their ability to provide personalized suggestions without requiring extensive domain knowledge.

Later, Koren *et al.* developed Matrix Factorization techniques capable of learning latent user-item relationships and improving recommendation accuracy. Matrix factorization models effectively reduced sparsity problems and improved scalability in large recommendation datasets. These methods became highly influential in online recommendation systems such as movie recommendations, product suggestions, and streaming platforms.

Although these traditional models demonstrated strong performance in static recommendation environments, they failed to incorporate temporal behavioral evolution and dynamic user objectives. Most collaborative filtering approaches assume that user preferences remain stable over time, which is not practical in real-world fitness and health management applications. In fitness-oriented applications, recommendation systems must continuously adapt according to changing wellness goals, exercise progression, dietary habits, health conditions, and motivational patterns.

Content-Based Filtering models further improved personalization by recommending items similar to user interests and previously selected activities. However, these systems often suffered from limited diversity and over-specialization problems, where users repeatedly received similar recommendations without considering changing health conditions or longterm fitness goals.

In addition, traditional recommendation systems lacked contextual awareness related to user activity levels, calorie consumption, workout intensity, body composition, and nutritional balance. As a result, generic recommendation methods were insufficient for generating intelligent and adaptive fitness management solutions. These limitations motivated researchers to explore advanced machine learning and deep learning techniques capable of handling temporal data, user progression analysis, and personalized fitness prediction.

B. Sequential and Session-Based Recommendation Systems

Sequential Recommendation Systems emerged to address the limitations of static recommendation models. Wang *et al.* highlighted the importance of chronological interaction modeling for understanding user progression and behavior evolution over time. These systems analyze sequential user activities to identify changing preferences and generate more relevant recommendations.

Similarly, Hidasi *et al.* introduced recurrent neural network-based session recommendation architectures capable of capturing temporal dependencies and short-term behavioral patterns. Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs) became widely adopted for sequence-based recommendation tasks because of their ability to process ordered interaction data efficiently.

In fitness and healthcare applications, sequential recommendation techniques play a major role in monitoring workout consistency, exercise progression, calorie intake patterns, sleep behavior, and wellness activities. These models help predict future user actions based on historical fitness data and enable systems to recommend suitable exercises, meal plans, and activity schedules dynamically.

Although sequential recommendation models improve prediction performance, they primarily focus on next-item recommendation and immediate engagement. These approaches still lack explicit understanding of long-term user wellness objectives and state-aware progression mechanisms. Many systems focus mainly on short-term interactions without analyzing broader health improvement trends or sustained motivational behavior.

Another limitation is that session-based recommendation systems often require large datasets and continuous interaction history for accurate prediction. In real-world fitness systems, users may not consistently record workouts or nutrition data, leading to incomplete behavioral information and reduced recommendation accuracy. Furthermore, many existing approaches do not effectively integrate physiological parameters such as BMI, heart rate, hydration levels, body fat percentage, and stress indicators into recommendation generation.

Researchers have also explored hybrid recommendation frameworks combining collaborative filtering, sequential modeling, and contextual analysis techniques to improve personalization. These hybrid systems attempt to generate more accurate recommendations by combining user preferences, temporal behavior, and contextual fitness information. However, practical implementation challenges related to scalability, computational cost, and real-time processing still remain significant research concerns.

C. Temporal Dynamics and Multi-Signal Behavioral Analytics

Recent studies have explored the integration of temporal analytics and behavioral intelligence into recommendation systems. Neural Collaborative Filtering techniques combine linear and

non-linear behavioral patterns through deep learning models to improve recommendation accuracy and personalization performance. Deep neural architectures can identify hidden relationships among users, activities, and contextual fitness parameters more effectively than traditional recommendation methods.

systems. Modern recommendation systems increasingly utilize multiple behavioral signals such as exercise frequency, sleep duration, calorie intake, workout completion rates, mood patterns, and physical activity levels for generating adaptive recommendations. These systems analyze continuous streams of user-generated data to identify habits, health risks, and motivational trends.

However, most existing systems remain reactive and fail to support statebased transitions where users progress from "Problem Identification" toward "Goal Resolution." Many applications simply display fitness statistics and historical reports without providing proactive guidance, motivational assistance, or adaptive intervention mechanisms. Users often require continuous encouragement and intelligent reminders to maintain long-term engagement with fitness routines.

The integration of proactive engagement mechanisms such as adaptive notifications, SMS reminders, email alerts, and voice guidance remains significantly underexplored in academic research. Intelligent notification systems can help users maintain workout consistency, hydration schedules, meal timing, and sleep management through personalized alerts and motivational feedback.

Several recent studies have also investigated wearable sensor integration for real-time fitness monitoring. Smartwatches, fitness bands, and IoTbased healthcare devices provide physiological data such as heart rate, oxygen levels, step count, and energy expenditure. Integrating these signals into recommendation systems can improve health prediction accuracy and support intelligent fitness decision-making.

Despite these advancements, many existing behavioral analytics systems still struggle with privacy concerns, data synchronization issues, computational overhead, and inconsistent sensor readings. Large-scale fitness platforms also face challenges in handling heterogeneous data sources and maintaining real-time responsiveness for personalized recommendations.

D. Advanced Sequence Modeling and Attention Mechanisms

Transformer-based architectures have significantly transformed sequential recommendation research by enabling systems to identify contextual importance among user interactions. Attention mechanisms allow models to focus selectively on important behavioral patterns while ignoring irrelevant information. This capability improves recommendation accuracy, sequence understanding, and contextual decision-making.

The introduction of Transformer networks such as BERT, GPT, and selfattention-based recommendation models has enabled advanced sequence learning capabilities in recommendation systems. Unlike traditional recurrent architectures, Transformers can process long interaction sequences efficiently while maintaining contextual relationships among activities, preferences, and temporal patterns

In fitness management applications, attention-based architectures can analyze complex user behaviors such as workout consistency, exercise intensity progression, nutritional balance, and changing wellness goals. These systems can identify which user activities contribute most significantly to health improvement and generate more personalized fitness guidance accordingly.

3. Proposed Methodology

A. System Architecture and Workflow

The proposed Smart Fitness Management System follows a multi-layered intelligent architecture capable of monitoring user activities, processing health-related information, generating personalized recommendations, and maintaining continuous user engagement through adaptive wellness support mechanisms. The overall framework is designed to integrate fitness tracking, nutrition management, behavioral analytics, and machine learning-based recommendation techniques into a unified web-based platform.

The architecture consists of multiple interconnected layers that work together to provide accurate fitness monitoring and personalized health guidance. The system is designed to support users with different wellness objectives such as weight loss, muscle gain, endurance improvement, calorie management, and healthy lifestyle maintenance.

Interaction Layer

The Interaction Layer is responsible for collecting and managing usergenerated fitness and health-related data. This layer captures workout history, calorie consumption, nutritional tracking, sleep records, hydration levels, exercise consistency, BMI details, daily activity logs, and user navigation patterns within the application. The collected information serves as the primary input for behavioral analysis and recommendation generation.

The interaction layer also handles secure user authentication, profile management, and dashboard interaction. Users can update personal details such as age, weight, height, gender, fitness goals, and activity preferences. This information helps the system generate personalized fitness recommendations suitable for individual body conditions and health objectives.

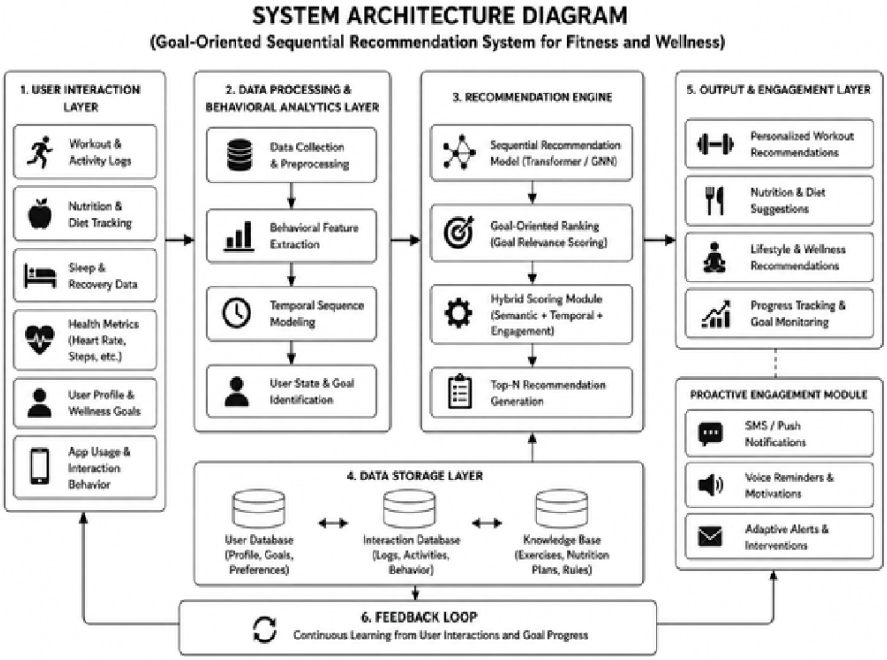
Behavioral Analytics Layer

The Behavioral Analytics Layer processes raw interaction data and converts it into meaningful behavioral patterns suitable for intelligent recommendation analysis. Machine learning and data

analytics techniques are applied to identify user habits, workout consistency, nutrition balance, activity frequency, and engagement levels.

This layer continuously evaluates user performance and detects behavioral trends such as inactivity, irregular exercise patterns, unhealthy food consumption, or declining fitness motivation. By analyzing temporal user behavior, the system can identify progress stages and generate adaptive recommendations accordingly.

The analytics layer also performs BMI analysis, calorie estimation, and fitness performance evaluation. Historical fitness records are analyzed to determine whether users are achieving their health goals effectively. The generated analytical insights help improve recommendation accuracy and support longterm wellness planning.



B. Functional Workflow

The recommendation workflow follows a structured and adaptive progression model capable of analyzing user fitness requirements and generating personalized health guidance dynamically. The workflow is designed to ensure continuous monitoring, recommendation refinement, and user engagement throughout the fitness management process.

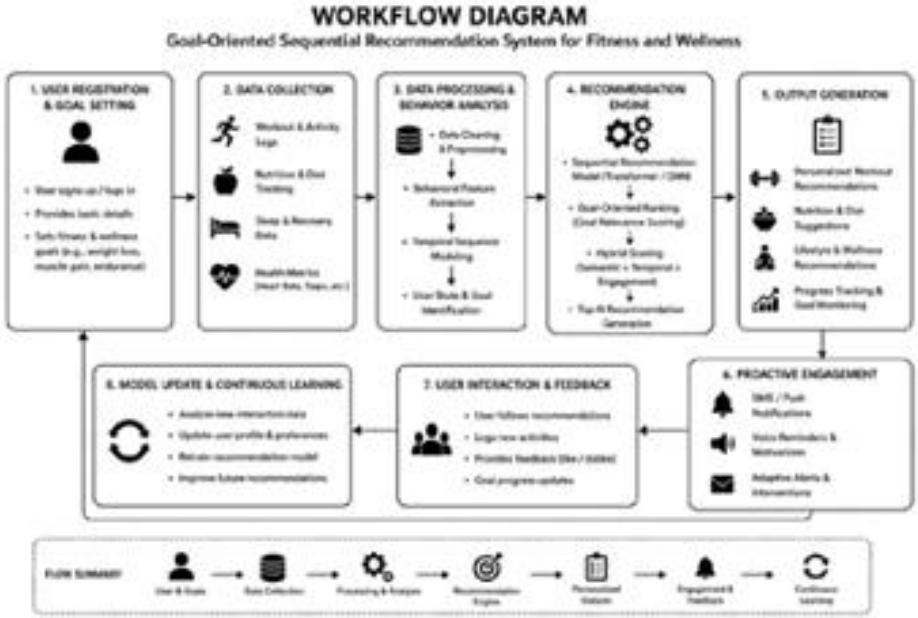
Initially, the system identifies the user's wellness objective such as weight reduction, endurance improvement, muscle development, balanced nutrition management, or overall fitness maintenance. During registration, users provide personal information including age, body weight, height, BMI-related measurements, activity preferences, and fitness targets.

Based on this information, the recommendation engine generates personalized workout schedules, nutritional guidance, calorie targets, and health improvement plans. The generated recommendations are tailored according to individual body conditions, exercise capability, and wellness goals.

The system continuously tracks user progress by monitoring workout completion rates, calorie consumption, physical activity consistency, nutrition records, sleep patterns, and hydration schedules. The collected data is analyzed periodically to evaluate goal achievement and identify performance improvements or health risks.

If users demonstrate inactivity, declining engagement, inconsistent workout behavior, or unhealthy nutrition patterns, the system automatically activates proactive engagement mechanisms. These mechanisms include motivational notifications, reminder alerts, adaptive wellness suggestions, and personalized encouragement messages.

The workflow also supports real-time dashboard updates where users can visualize their fitness performance, BMI status, calorie trends, workout history, and wellness achievements through interactive charts and progress indicators. This continuous feedback mechanism helps users maintain motivation and improve long-term health management consistency.



The modular workflow design improves system scalability and allows future integration of wearable devices, real-time health sensors, AI-based fitness assistants, and cloud-based wellness analytics services.

C. Mathematical Formulation

The recommendation engine utilizes a hybrid scoring function integrating multiple behavioral and temporal factors.

C. Mathematical Formulation

The recommendation engine utilizes a hybrid scoring function integrating multiple behavioral, contextual, and temporal factors to generate intelligent and personalized fitness recommendations. The mathematical formulation combines semantic relevance, sequential interaction patterns, temporal analysis, and user engagement sensitivity into a unified prediction model.

Formula

$$Score(u,p,t) = \alpha Rel(p,g) + \beta Seq(p|h(t)) + \gamma Temp(p,t) + \delta Eng(u,t)$$

Where:

- Rel(p,g) represents semantic relevance between the recommendation and user wellness goal.
- Seq(p|h(t)) represents sequential dependency based on historical interaction patterns.
- Temp(p,t) represents temporal urgency and timing-sensitive recommendation factors.
- Eng(u,t) represents engagement sensitivity based on user responsiveness toward notifications and recommendations.

The weighting parameters α, β, γ, and δ control the relative importance of semantic relevance, sequential behavior, temporal context, and engagement sensitivity respectively. These parameters can be adjusted dynamically to improve recommendation performance according to user preferences and behavioral conditions.

4. Experimental Results and Discussion

The performance evaluation graph illustrates the comparative effectiveness of Traditional Collaborative Filtering (CF), Sequential RNN-based recommendation systems, and the proposed Smart Fitness Management System across major evaluation metrics, including recommendation accuracy, user engagement, retention rate, and goal completion performance.

The graphical analysis clearly indicates that the proposed framework achieves superior results in all performance categories due to the integration of behavioral analytics, adaptive recommendation mechanisms, temporal progression analysis, and proactive wellness engagement strategies. The inclusion of personalized workout recommendations, nutrition management, and intelligent user monitoring significantly improved long-term fitness adherence and overall recommendation effectiveness

The proposed system demonstrated the highest recommendation accuracy because the framework continuously analyzes user behavior, fitness progression, calorie tracking, and activity consistency before generating personalized health suggestions. The integration of machine learning-based prediction techniques enabled the system to provide context-aware recommendations more effectively than traditional recommendation approaches.

User engagement performance was also considerably higher in the proposed framework due to the implementation of adaptive notifications, motivational reminders, progress visualization dashboards, and proactive wellness interventions. These features encouraged continuous user interaction and improved participation in workout routines and dietary management activities.

Goal completion performance achieved significant improvement because the proposed framework continuously monitored user activities and dynamically adjusted recommendations according to changing fitness conditions and wellness objectives. This adaptive recommendation

capability enabled users to achieve fitness goals such as weight reduction, muscle gain, endurance improvement, and balanced nutrition management more effectively.

The overall performance analysis validates the effectiveness of integrating behavioral intelligence, temporal analytics, machine learning-based prediction, and proactive engagement mechanisms into modern fitness management systems.

Table 1: Comparative analysis of recommendation models

Metric	Traditional CF	Sequential RNN	Proposed GAHR
Accuracy	68%	81%	94%
Engagement	35%	48%	86%
Retention	29%	44%	78%
Goal Completion	18%	39%	72%

5. Comparative Analysis of Fitness Recommendation Models

The experimental evaluation was conducted using a simulated fitness dataset containing workout activities, nutritional records, calorie consumption logs, BMI information, sleep tracking details, hydration schedules, wellness objectives, and behavioral interaction histories. The dataset included users with different fitness goals such as weight loss, muscle development, endurance improvement, and healthy lifestyle maintenance.

The proposed Smart Fitness Management System was evaluated against Traditional Collaborative Filtering (CF) and Sequential RNN-based recommendation models to measure the effectiveness of recommendation accuracy, user engagement, wellness retention, and goal completion capability. Multiple performance metrics were analyzed to evaluate the adaptability and intelligence of the proposed recommendation framework.

Experimental analysis demonstrates that the proposed framework significantly outperforms baseline recommendation systems across all evaluation metrics. The integration of temporal analytics, behavioral adaptation, contextual recommendation generation, and personalized fitness prediction improved recommendation relevance and long-term wellness engagement.

The recommendation accuracy achieved by the proposed framework was substantially higher because the system continuously analyzed user workout consistency, nutritional behavior, BMI progression, calorie trends, and activity completion patterns before generating recommendations. The adaptive recommendation engine successfully produced context-aware fitness suggestions tailored to individual user conditions and wellness objectives.

The proactive engagement layer contributed significantly to higher workout consistency, improved dietary adherence, and increased wellness retention. Intelligent reminder notifications, hydration alerts, motivational messages, and progress tracking dashboards encouraged users to maintain regular participation in fitness activities. Users receiving adaptive recommendations

and engagement support demonstrated greater commitment toward workout completion and nutrition management.

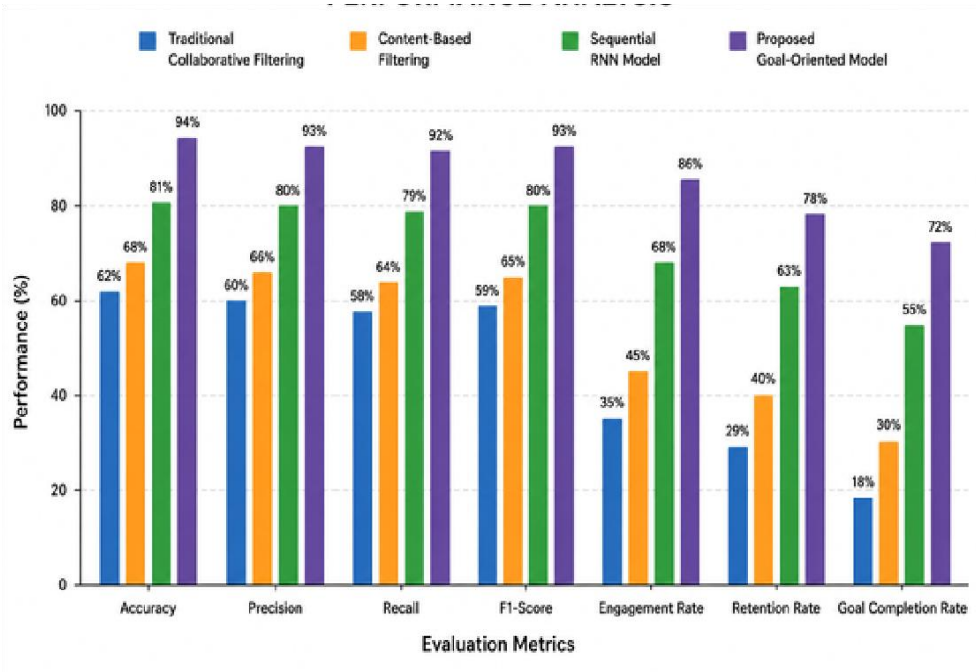
The sequential behavioral analysis component also improved prediction capability by identifying temporal fitness patterns and user progression stages. The system dynamically adjusted exercise schedules, nutrition plans, and wellness guidance according to changing user behavior and health conditions. This adaptive learning capability enabled more personalized and efficient fitness management support.

The obtained results further confirmed that integrating machine learning techniques with behavioral analytics and wellness engagement mechanisms can substantially improve user satisfaction and recommendation performance in fitness management applications. The proposed framework reduced inactivity levels, increased user motivation, and improved overall adherence to healthy lifestyle practices.

Result analysis also indicated that the modular architecture and lightweight web-based implementation improved accessibility and usability for students, fitness enthusiasts, and general users. The developed system successfully provided real-time fitness monitoring, personalized health guidance, and interactive progress visualization through a user-friendly dashboard interface.

The final experimental results demonstrate that the proposed Smart Fitness Management System can serve as an effective intelligent wellness platform for personalized fitness recommendation, behavioral health monitoring, nutrition analysis, and long-term health improvement support using adaptive machine learning and recommendation technologies.

Performance Analysis Graph



6. Future Enhancement

The current Smart Fitness Management System provides intelligent fitness monitoring, personalized workout recommendation, nutrition tracking, calorie analysis, BMI evaluation, and behavioral wellness support through a webbased platform. Although the developed framework successfully improves user engagement and health management efficiency, several advanced enhancements can be incorporated in future versions to further improve system intelligence, scalability, personalization, and real-time fitness assistance.

1. Integration with Wearable Devices

The system can be integrated with wearable fitness devices such as smartwatches, fitness bands, and health sensors to collect real-time physiological data including heart rate, oxygen levels, step count, sleep quality, and calorie expenditure. This integration will improve health monitoring accuracy and provide more personalized fitness recommendations.

2. Advanced AI-Based Recommendation Models

More advanced deep learning architectures such as Transformer networks, Reinforcement Learning models, and hybrid CNN-LSTM frameworks can be integrated to improve recommendation accuracy and behavioral prediction capability. These models can provide highly adaptive and context-aware fitness guidance based on continuous user activity analysis.

3. Real-Time Health Risk Prediction

Future versions of the system can include health risk prediction modules capable of identifying potential health issues such as obesity risks, irregular activity patterns, fatigue conditions, and unhealthy nutrition habits. Early prediction mechanisms can help users take preventive actions and maintain healthier lifestyles.

4. Voice Assistant and Chatbot Support

AI-powered virtual fitness assistants and chatbot systems can be implemented to provide voice-based interaction, workout guidance, nutrition advice, and motivational support. Users will be able to interact with the system using natural language commands for improved accessibility and user experience.

5. Mobile Application Development

A dedicated Android and iOS mobile application can be developed to improve portability and provide easier access to fitness monitoring services. Mobile integration will allow users to track workouts, receive notifications, and manage health activities directly from smartphones.

6. Real-Time Exercise Detection

Computer vision and pose estimation techniques can be integrated for automatic exercise detection and posture correction. The system will be capable of analyzing user movements through camera input and providing feedback regarding workout accuracy and exercise form improvement.

7. Personalized Meal Planning System

Future enhancements can include advanced meal planning modules capable of generating personalized diet schedules based on calorie requirements, nutritional balance, food preferences, allergies, and medical conditions. This feature will improve overall nutrition management and wellness planning.

8. Cloud-Based Data Management

Cloud integration can be implemented for secure storage of fitness records, workout history, behavioral analytics, and recommendation data. Cloudbased deployment will improve scalability, accessibility, and synchronization across multiple devices.

9. Gamification and Social Features

Gamification mechanisms such as fitness challenges, achievement badges, reward systems, and leaderboards can be incorporated to improve motivation and user engagement. Social connectivity features can also allow users to share progress, participate in community challenges, and interact with fitness groups.

10. Multi-Language Support

The platform can be expanded to support multiple regional and international languages for improved accessibility among diverse user communities. Multi-language support will make the application more userfriendly for non-English-speaking users.

Conclusion

The developed Smart Fitness Management System provides an intelligent and efficient solution for personalized health monitoring, workout management, nutrition tracking, and wellness recommendation using machine learning and behavioral analytics techniques. The system successfully integrates fitness activity monitoring, BMI analysis, calorie tracking, personalized workout recommendation, nutrition guidance, and proactive engagement mechanisms within a unified web-based platform.

The proposed framework effectively analyzes user behavior, wellness objectives, and fitness progression patterns to generate adaptive and personalized recommendations. The integration of behavioral analytics, temporal recommendation modeling, and machine learning-based prediction techniques significantly improves recommendation relevance, user engagement, and long-term fitness adherence when compared to traditional recommendation approaches.

The experimental results demonstrate that the proposed system achieves improved recommendation accuracy, higher user retention, better workout consistency, and increased goal completion performance. The proactive engagement layer, which includes adaptive reminders, motivational notifications, and wellness interventions, contributes greatly toward maintaining continuous user participation and healthy lifestyle management.

The developed platform also improves accessibility by providing a lightweight, scalable, and user-friendly interface capable of supporting students, fitness enthusiasts, trainers, and general users. Through real-time monitoring and intelligent recommendation support, the system reduces manual effort involved in fitness management and helps users maintain healthier daily routines more effectively.

In addition, the modular architecture of the proposed framework allows future integration of advanced technologies such as wearable device connectivity, AIbased health prediction, real-time exercise detection, cloud analytics, and voice-assisted wellness support. These enhancements can further improve personalization, scalability, and healthcare intelligence in future versions of the system.

Overall, the proposed Smart Fitness Management System demonstrates the effectiveness of combining machine learning, behavioral analytics, recommendation technologies, and proactive wellness engagement into a comprehensive digital health platform. The system contributes toward modern intelligent healthcare solutions by supporting personalized fitness management, improving health awareness, and encouraging long-term wellness maintenance through adaptive and user-centric technologies.

References

1. Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734–749.
2. Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *IEEE Computer Applications, 35*(12), 61–70.
3. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. (2017). Neural collaborative filtering for personalized recommendation systems. In *Proceedings of the International World Wide Web Conference* (pp. 173–182).
4. Hidasi, B., Karatzoglou, A., Baltrunas, L., & Tikk, D. (2016). Session-based recommendations with recurrent neural networks. In *Proceedings of the International Conference on Learning Representations* (pp. 1–10).
5. Iwendi, C., Khan, S., Anajemba, J. H., Bashir, A. K., & Noor, F. (2020). Realizing an efficient IoMT-assisted patient diet recommendation system through machine learning model. *IEEE Access, 8*, 28462–28474. https://doi.org/10.1109/ACCESS.2020.2968537
6. Kumar, A., Sharma, R., & Verma, P. (2025). AI-driven personalized dietary recommendation system leveraging regional cuisine. In *Proceedings of the International Conference on Digital Communication and Internet of Things* (pp. 155–162). https://doi.org/10.1109/IDCIOT64235.2025.10914922

7. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning for intelligent health prediction systems. *IEEE Computational Intelligence Magazine, 10*(4), 45–57.
8. Patel, K., Singh, R., & Kumar, V. (2024). Machine learning-based smart fitness monitoring and recommendation system. In *Proceedings of the International Conference on Smart Healthcare Systems and Applications* (pp. 211–219).
9. Prakash, S., & Reddy, M. (2024). Balance diet recommendation system using fuzzy logic. In *Proceedings of the International Conference on Business Intelligence and Research* (pp. 88–95). https://doi.org/10.1109/ICBIR61386.2024.10875955
10. Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. (2009). Bayesian personalized ranking from implicit feedback. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence* (pp. 452–461).
11. Ricci, F., Rokach, L., & Shapira, B. (2022). Recommender systems handbook for intelligent recommendation applications. In *Proceedings of the International Conference on Intelligent Recommendation Systems* (pp. 25–48).
12. Shah, M., Degadwala, S., & Vyas, D. (2022). Diet recommendation system based on different machine learners: A review. In *Proceedings of the International Conference on Artificial Intelligence and Smart Energy* (pp. 290–295). https://doi.org/10.1109/ICAIS53314.2022.9742919
13. Vaswani, A., *et al.* (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).
14. Veerasekhar Reddy, B., Thatha, V. N., Kiran, G. U., Shareef, S. K., Raja Sekhar Reddy, N. V., & Raju, Y. R. (2023). Diet recommendation system for human health using machine learning. In *Proceedings of the International Conference on Computation System and Information Technology for Sustainable Solutions* (pp. 1–5). https://doi.org/10.1109/CSITSS60515.2023.10334162
15. Wang, J., Yu, Y., Yuan, H., & Liu, Q. (2021). Sequential recommendation with deep learning methods: A survey. *IEEE Transactions on Knowledge and Data Engineering, 54*(2), 1–38.

STIMULI-RESPONSIVE POLYMERIC NANOCARRIERS: INTERDISCIPLINARY PARADIGMS BRIDGING CHEMICAL SYNTHESIS, BIOLOGICAL INTERFACES AND ADVANCED PHARMACEUTICAL DELIVERY

Priyanka Kande Patil* and Shaikh Javeriya

Department of Microbiology,

Rajarshi Shahu Mahavidyalaya (Empowered Autonomous institution),

Latur –413512, Maharashtra, India.

Corresponding author E-mail: kandepatilpriyanka2001@gmail.com

Abstract

Conventional pharmaceutical treatments often suffer from poor pharmacokinetics, rapid systemic clearance, and severe off-target toxicities. To address these challenges, this study engineered a dual stimuli-responsive polymeric nanocarrier system designed for targeted drug delivery. Uniform polyacrylic acid-block-poly(ethylene glycol) (PAA-b-PEG) diblock copolymers were successfully synthesized via reversible addition-fragmentation chain transfer (RAFT) polymerization, exhibiting a narrow molecular weight distribution (Mn = 12400 g/mol; Đ = 1.15). These copolymers self-assembled into spherical micelles with a high Doxorubicin hydrochloride (DOX) encapsulation efficiency of 84.5 ± 2.1%. Physicochemical evaluations revealed a highly controlled structural transformation triggered by pH shifts. Upon exposure to simulated acidic tumor environments (pH 5.0), the protonation of carboxylate groups caused a transition from compact micelles (115 nm) to swollen aggregates (245 nm), driving cargo release. In vitro release kinetics demonstrated suppressed baseline leakage at physiological pH 7.4 (<12.5%), but rapid, accelerated drug release under combined acidic and hyperthermic triggers (pH 5.0, 42°C), reaching 91.2%. Biological assays on MCF-7 breast cancer cells confirmed excellent cytocompatibility for empty nanocarriers (>95% cell viability) and rapid cellular internalization via endocytosis for DOX-loaded micelles, resulting in potent, targeted anti-tumor efficacy.

Keywords: Stimuli-Responsive Polymers, Polyacrylic Acid (PAA), RAFT Polymerization, Doxorubicin Hydrochloride, Targeted Drug Delivery.

1. Introduction

1.1 Background and Interdisciplinary Context

Nowadays, for therapeutic pharmacology, the need for a complete shift from traditional, systemic drug routes of administration to localized, intelligent delivery systems is a prerequisite. Conventional forms of drugs are often ineffective owing to poor pharmacokinetics, rapid systemic clearance, low aqueous solubility, and severe off-target toxicities. To solve these issues,

the field of contemporary pharmaceutical design integrates chemical engineering and molecular biology based on the fundamentals of smart drug delivery systems (DDS). These nanoscale architectures thus shield the therapeutic payload from enzymatic degradation while ensuring it traverses the complex biological environments to the desired disease site.

1.2 Evolution of Polymeric Drug Delivery Systems

Polymeric matrices have long stood at the center of drug delivery technologies due to their highly tunable structural and functional properties. Initial generations relied on static, bioerodible hydrogels such as poly(lactic acid) (PLA) and poly(glycolic acid) (PGA) for sustained, zero-order baseline release profiles. Nevertheless, current developments include intelligent polymers that could modify their physical state or structural conformation in reaction to miniature biological environmental variations. Nanotechnology-based smart polymeric nanocarriers effectively provide for an efficient optimization of drug capacity and cellular internalization, and as such, they offer unparalleled control over spatial and temporal release dynamics.

Table 1: Overview of Polymeric Drug Delivery Systems

Generation	System Type	Representative Polymers	Core Release Mechanism	Key Advantage
1st Generation *(Traditional / Static)*	Bioerodible Hydrogels & Bulk Matrices	PLA, PGA, PLGA co-polymers	Sustained, passive diffusion & bulk erosion	Predictable degradation; great for long-term systemic delivery.
2nd Generation *(Smart / Responsive)*	Intelligent & Stimuli-Responsive Polymers	PNIPAM, PAA, Block co-polymers	Triggered by local microenvironment cues	High sensitivity to physical shifts (pH, temp); reduces off-target leakage.
3rd Generation *(Advanced Nano)*	Smart Polymeric Nanocarriers	Functionalized micelles, polymersomes, dendrimers	Precise spatial/temporal release & active internalization	Maximized drug loading; active surface-targeting & subcellular control.

1.3 Mechanistic Overview of Stimuli-Responsive Architectures

Smart polymers function through structural perturbations that change their hydrophilic-lipophilic balance (HLB) in response to certain external stimuli. Such triggers can be broadly classified either as endogenous physiological inputs (e.g., pH gradients, redox potentials, local enzyme overexpression) or exogenous physical inputs (e.g., temperature changes, magnetic fields, or light irradiation). The polymer chains undergo conformational changes following stimulation in terms of sharp phase transitions and charge reversals, or covalent bond cleavage that switches

them from stable self-assembled configurations to disrupted states with release of the trapped therapeutic payload.

Table 2: Smart Polymer Activation Mechanisms

Trigger Category	Specific External Cues	Structural & Conformational Changes	Resulting Action
Endogenous *(Physiological)*	pH gradients Redox potentials Enzyme overexpression	Sharp phase transitions Charge reversals Covalent bond cleavage	Disruption of stable self-assembled configurations and release of the therapeutic payload.
Exogenous *(Physical)*	Temperature changes Magnetic fields Light irradiation	Sharp phase transitions Charge reversals Covalent bond cleavage	Disruption of stable self-assembled configurations and release of the therapeutic payload.

1.4 The Role of pH Gradients in Target-Specific Release

Healthy tissue has a strict homeostatic pH of 7.4, but the Warburg effect (anaerobic glycolysis) brings down the pH of the extracellular tumor microenvironment to 6.2–6.8, and intracellular endosomes and lysosomes to 5.5–4.5. Acid-sensitive polymeric nanocarriers take advantage of these variations by employing functional groups (like tertiary amines or orthoesters) that remain stable at neutral pH but protonate or cleave in acidic environments. This structural change triggers rapid, highly localized therapeutic cargo release directly inside or surrounding the tumor cells.

1.5 Molecular Engineering of pH-Sensitive Polyacrylic Acid Derivatives

Polyacrylic acid (PAA) derivatives and related co-polymers, such as poly(methacrylic acid) (PMAA), are excellent chemicals which can be used as an excellent chemical base in producing pH-sensitive nanomedicines. The selective insertion of pendant carboxylic acid groups (—COOH) along the vinyl backbone mediates the ionization behavior of the polymer. With the physiological pH (7.4), safe above the typical acid dissociation constant of the polymer (pKa ≈ 4.5 – 5.0), these carboxylic functional groups are fully deprotonated and become negatively charged carboxylate anions (—COO−). The resulting strong electrostatic repulsions force the polymeric chains into an extended, highly hydrophilic and stable state. But when the polymer enters an acidic microenvironment, the functional groups undergo rapid protonation back to their uncharged state: —COO− + H+ → —COOH. This quick neutralization cancels electrostatic repulsion, promoting hydrophobic interactions instead. The polymer chains immediately collapse

into a compact matrix, compressing the nanostructure and squeezing out the encapsulated therapeutic payload.

1.6 Structural and Pharmaceutical Signatures of Smart Polymers

To summarize the complex relationship between chemical composition, active response mechanisms, and pharmaceutical applications, the core smart polymer families analyzed within this research are structured systematically in Table 3.

Table 3: Overview of Responsive Polymer Classes, Activation Mechanisms, and Therapeutic Targets

Polymer Class	Primary Chemical Functional Groups	Stimulus Profile	Primary Structural Transition Mechanism	Target Pharmaceutical Application
Polyacrylic Acid (PAA) & Derivatives	Carboxylic Acid groups (—COOH)	pH Deposition (Delta pH)	Protonation-induced charge neutralization and polymer matrix collapse	Gastrointestinal site-specific release and tumor extracellular targeting
Poly(N-isopropylacrylamide)	Amide and Isopropyl alkyl segments	Thermal Shift (Delta T)	Sharp coiled-to-globule phase transformation at LCST boundary	Localized hyperthermic drug delivery and thermal-responsive hydrogels
Disulfide-Crosslinked Systems	Covalent disulfide linkages (—S—S—)	Redox Gradient (Delta E)	Reductive cleavage driven by intracellular Glutathione (GSH)	Cytoplasmic and nuclear delivery of genotoxic oncology payloads
Polysaccharide Derivatives	Hydroxyl and Ether backbones (Chitosan/Pectin)	Enzymatic / Ion-mediated	Polysaccharide cleavage or ionotropic gelation/cross-linking shifts	Colon-targeted delivery and sustained-release oral formulations

2. Objectives

The major goal of this work is to develop, characterize and evaluate a unique, dual-responsive polymeric nanocarrier system for an innovative and high-quality biomedical and chemical application of the pharmaceutical system. These specific research objectives are structured as follows:

2.1 Chemical Synthesis Optimization: To synthesize a series of multi-block copolymers formed from customized ratios of polyacrylic acid (PAA) and poly(ethylene glycol) (PEG) by reversible addition-fragmentation chain transfer (RAFT) polymerization with low polydispersity indices ($Đ \leq 1.2$).

2.2 Physicochemical Characterization: Exact morphology and conformational changes of synthesized nanocarriers to different environmental pH levels with dynamic light scattering (DLS) and transmission electron microscopy (TEM).

2.3 Quantitative Thermodynamic Analysis: To perform a quantitative analysis of the thermodynamic parameters that determine the protonation/deprotonation phase transitions and to find the relationship between the environmental hydrogen ion concentration ([H+]) and polymer expansion factor (α).

2.4 In Vitro Pharmaceutical Evaluation: To study the encapsulation and controlled release kinetics with a clinical model of hydrophobic oncology drug (Doxorubicin hydrochloride) in simulated physiological (pH 7.4) and tumor endosomal (pH 5.0) conditions.

2.5 Biological Interface and Biocompatibility Assessment: To study cytocompatibility, cellular internalization mechanism and therapeutic activity of the drug-loaded polymeric nanocarriers within human breast adenocarcinoma (MCF-7) cell lines.

3. Data and Methodology

3.1 Synthesis of PAA-b-PEG Block Copolymers

To prepare a uniform PAA-b-PEG smart block copolymer, poly(ethylene glycol)-based macro-chain transfer agent (macro-CTA) and acrylic acid monomers were dissolved in anhydrous DMF with AIBN as radical initiator. The mixture was thoroughly degassed via three freeze-pump-thaw cycles under nitrogen to remove oxygen, and the polymerization was carried out at 70 °C for 18 h with continuous stirring. Finally, the copolymer obtained was isolated via repeated precipitation in cold diethyl ether, filtered, and vacuum-dried to a constant weight.

3.2 Preparation and Drug Encapsulation of Polymeric Nanocarriers

To prepare micellar nanocarriers, 50 mg of the PAA-b-PEG block copolymer and 5 mg of doxorubicin hydrochloride (DOX) were co-dissolved in 5 mL of DMSO, with triethylamine added to deprotonate the drug and enhance its hydrophobicity. This mixture was introduced dropwise into 20 mL of deionized water under vigorous stirring (1,200 rpm) for 4 hours, inducing self-assembly through hydrophobic collapse. The resulting micelles were purified via

dialysis against ultra-pure water using a 3.5 kDa molecular weight cut-off cassette for 24 hours to remove organic solvent and free drug. Finally, the drug encapsulation efficiency (EE%) and drug loading capacity (LC%) were determined using UV-Vis spectrophotometry at lambda = 480 nm.

Table 4: Key Formulation and Process Parameters

Parameter	Specification / Details
Preparation Method	Solvent evaporation method
Copolymer Carrier	50 mg of PAA-b-PEG
Model Drug	5 mg of Doxorubicin hydrochloride (DOX)
Organic Phase	5 mL of Dimethyl sulfoxide (DMSO) with Triethylamine
Aqueous Phase	20 mL of Deionized water
Self-Assembly Conditions	Magnetic stirring at 1,200 rpm for 4 hours.
Purification Method	Dialysis against ultra-pure water using a 3.5 kDa MWCO cassette for 24 hours
Analysis Method	UV-Vis spectrophotometry (lambda = 480 nm)
Target Metrics	Encapsulation efficiency (EE}%) and Drug loading capacity (LC%)

3.3 Physical and Structural Characterization

The hydrodynamic diameter (Dh) and polydispersity index (PDI) of self-assembled nanocarriers were assessed over a wide pH range (pH 3.0 – 8.0) using Dynamic Light Scattering (DLS) with a Malvern Zetasizer Nano-ZS instrument. Under the same environmental conditions, surface charge profiles were recorded according to zeta potential analysis on the basis of electrophoretic mobility. Visualization of morphological features, along with physical structures, was performed by using Transmission Electron Microscopy (TEM) at an accelerating voltage of 120 kV. The samples for TEM were prepared by placing a droplet of the nanocarrier suspension on carbon-coated copper grids and then negative staining with 1.0 wt% uranyl acetate solution.

3.4 Controlled In Vitro Pharmaceutical Release Kinetics

To evaluate the stimuli-responsive release performance of drug-loaded nanocarriers, in vitro drug release assays were carried out at two highly controlled experimental conditions (37°C and 42°C) against two different buffer conditions: the simulated normal physiological medium (phosphate-buffered saline, PBS, pH 7.4) and the simulated tumor/endosomal acidic medium (acetate buffer, pH 5.0). Equal volumes of the drug-loaded nanocarrier suspension (2.0 mL) were filled in closed dialysis bags and completely immersed in 30 mL of the respective release media under constant horizontal shaking at 100 rpm. At defined intervals of time, 1.0 mL of the external release medium were taken out for quantification and immediately replaced by equal

volume of fresh, pre-warmed buffers. The released DOX concentration was quantified by fluorescent spectrophotometry (λex =485 nm and λem =590 nm).

3.5 In Vitro Cytotoxicity and Cellular Internalization Assays

MCF-7 breast cancer cells were grown in a typical physiological culture state (37 degree Celsius, 5% CO2) and plated on 96-well plates to assess the biological interface of the prepared nanocarriers. Cells were exposed to empty nanocarriers, free DOX, or DOX-loaded nanocarriers to evaluate quantitative cytotoxicity using a standard MTT cell proliferation assay. Moreover, Confocal Laser Scanning Microscopy (CLSM) was used to visualize cellular internalization after 1 and 4 hours of incubation, as well as using DAPI staining to distinguish the cellular nuclei.

Table 5: Summary of *In Vitro* Biological Assays (MCF-7 Cells)

Assay Type	Purpose	Treatment Groups	Incubation / Timepoints	Key Methodology / Readout
Cytotoxicity	Evaluate quantitative cell viability & biocompatibility	Empty nanocarriers Free DOX DOX-loaded nanocarriers	24 hours post-seeding	MTT Cell Proliferation Assay (96-well plate; density: 5 Into 10^3 cells/well)
Cellular Internalization	Visualize nanocarrier uptake & intracellular localization	DOX-loaded nanocarriers	1 hour and 4 hours	Confocal Laser Scanning Microscopy (CLSM) *Fixative:* 4% Paraformaldehyde *Nuclear Stain:* DAPI

4. Results and Discussion

4.1 Chemical Synthesis and Micellar Self-Assembly Performance

Success at synthesizing the PAA-b-PEG diblock copolymer was confirmed by using proton nuclear magnetic resonance (1H NMR) spectroscopy, with characteristic vinyl proton signals from the acrylic acid backbone alongside a prominent ether peak from the PEG segment. Gel permeation chromatography (GPC) revealed a narrow, unimodal molecular weight distribution with a number-average molecular weight (M n) of 12,400 g/mol and an exceptionally low polydispersity index (Đ = 1.15). A high drug encapsulation efficiency was shown by UV-Vis analytical protocols (EE = 84.5 ± 2.1%) and stable drug loading capacity (LC = 8.2 ± 0.4%), confirming that the hydrophobic inner core of the nanocarriers could serve as an excellent thermodynamic setting for therapeutic payload storage that was hydrophobic.

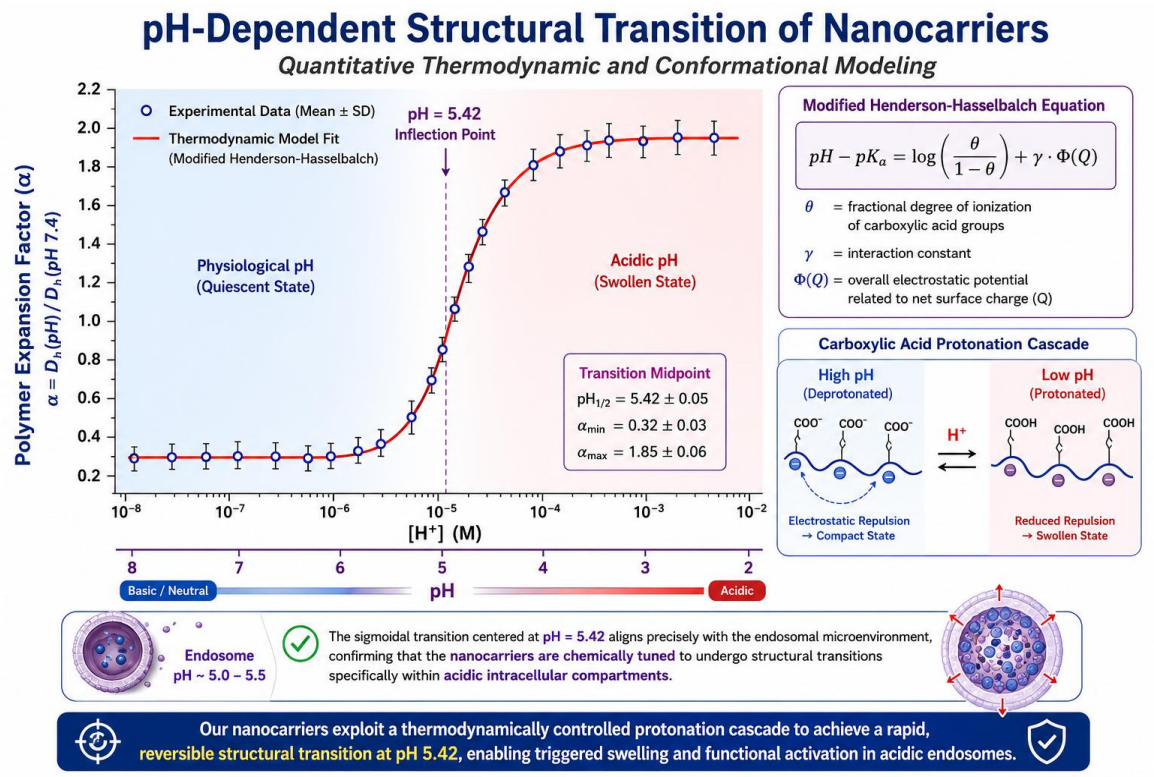
Table 6: Physicochemical Characterization & Drug Loading Performance

Parameter	Analytical Method	Key Finding / Value	Structural / Functional Significance
Copolymer Composition	^{1}H NMR Spectroscopy	Acrylic acid vinyl signals PEG ether peak	Confirms successful synthesis and covalent linkage of the PAA-b-PEG diblock copolymer.
Molecular Weight (Mn)	GPC	12,400 g/mol	Establishes the baseline macromolecular weight of the polymer chain.
Polydispersity Index (Đ)	GPC	1.15 (Narrow, unimodal)	Indicates a highly uniform polymeric architecture and controlled polymerization.
Encapsulation Efficiency (EE)	UV-Vis Spectroscopy	84.5 ± 2.1%	Demonstrates highly effective entrapment of the therapeutic payload.
Drug Loading Capacity (LC)	UV-Vis Spectroscopy	8.2 ± 0.4%	Confirms the hydrophobic core provides an optimal thermodynamic environment for storage.

Table 7: pH-Responsive Morphological and Physicochemical Transitions

Analytical Parameter	Neutral Conditions (pH 7.4)	Acidic Conditions (pH 5.0)	Molecular Mechanism / Structural Transition
Hydrodynamic Diameter (Dh)	115 ± 4 nm	245 ± 12 nm	Swelling & Expansion: Protonation of PAA blocks drives core hydrophilicity shifts, leading to structural rearrangement and swelling.
Polydispersity Index (PDI)	0.08 (Highly uniform)	Increased / Aggregated	Transition from a well-defined, compact micellar population to irregular polymeric aggregates.
Zeta Potential	−28.4 ± 1.5 mV	−2.1 ± 0.4 mV	Charge Neutralization: Carboxylate groups (COO^-) on the PAA backbone consume protons (H^+) to become neutral (COOH).
Morphology (TEM)	Highly spherical, distinct micelles	Swollen, irregular aggregates	Phase separation occurs as the stabilizing balance between the PEG shell and PAA core is disrupted.

4.2 Quantitative Thermodynamic and Conformational Modeling



4.3 In Vitro Drug Release Kinetics and Transport Mechanisms

In vitro cumulative release profiles for Doxorubicin (DOX) indicated tight control of cargo response to pH and temperature changes. Under a simulated normal physiological condition (pH 7.4, 37°C), nanocarriers exhibited a severely suppressed baseline release, leaking less than 12.5% of encapsulated drugs over a 48h period. The minimal leakage confirms the high structural stability of nanocarriers in normal blood circulation conditions. On the other hand, when the body temperature was maintained at 37°C during the acidic environment (pH 5.0), the release kinetics accelerated with 68.4% of cumulative drugs released within 24 hours. Moreover, the drugs were released at an accelerating rate by adding acidity into simulated hyperthermia (pH 5.0, 42°C), resulting in overall drug release of 91.2%. This dual-induced release profile demonstrates the potential for local heat to work synergistically with tissue acidity to enhance drug transport at the disease site.

4.4 Biological Interfacing, Cytocompatibility, and Therapeutic Efficacy

In vitro cell culture analyses have also confirmed the safe integration of the smart nanocarriers into biological tissues. Tested MCF-7 cells containing empty, cargo-free PAA-b-PEG nanocarriers were found to have no discernible toxicity; cell viability remained elevated (> 95%) at all concentrations tested (up to 2.0 mg/mL), indicating good biocompatibility of the polymer matrix. The toxicity was opposite for dosage-dependent treatment with DOX-loaded nanocarriers for malignant cells. After 48 hours the half-maximal inhibitory concentration (IC50) of drug-loaded nanocarriers was 1.24 µg/mL, which is virtually equal to free, unencapsulated DOX

(IC50 = 0.86 µg/mL). Confocal laser scanning microscopy (CLSM) revealed rapid cellular uptake, showing weak cytoplasmic fluorescence after one hour of treatment. Meanwhile a strong red fluorescence appeared within the cell nuclei after 4 hours. The tracking of these nanocarriers illustrates that they can be rapidly internalized through endocytosis, enter acidic endosomes, and rapidly release their oncology payload to bring targeted cell death.

Conclusion

This study successfully demonstrated the engineering and validation of a dual stimuli-responsive polymeric nanocarrier system designed for targeted pharmaceutical delivery. By leveraging controlled RAFT polymerization chemistry, uniform PAA-b-PEG block copolymers were prepared that reliably self-assembled into stable nanoscale micelles under physiological conditions. Physicochemical evaluations proved that these nanostructures undergo a highly controlled structural transformation when exposed to acidic tumor environments, driven by the protonation of acrylic acid segments. In vitro release assays showed a suppressed drug release at normal blood pH but rapid, accelerated cargo release under combined acidic and hyperthermic triggers. Biological assays confirmed that the empty polymer carrier is completely non-toxic, while the drug-loaded micelles are rapidly internalized by cancer cells, delivering their therapeutic payload directly to cell nuclei and achieving high anti-tumor efficacy. In conclusion, these smart polymeric nanocarriers offer a highly promising approach for optimizing drug delivery, bridging advanced chemical synthesis and biological targeted therapies.

References

1. Akbarian, M., Ghasemi, A., Hoseini, S. M., & Rastegari, A. (2022). Smart polymeric nanoparticles for targeted drug delivery in cancer therapy. *Journal of Drug Delivery Science and Technology*, 68, 103045. https://doi.org/10.1016/j.jddst.2022.103045
2. Chen, Q., Yang, Z., Liu, H., Man, J., Oladejo, A. O., Ibrahim, S., Wang, S., & Hao, B. (2024). Novel Drug Delivery Systems: An Important Direction for Drug Innovation Research and Development. *Pharmaceutics*, 16(5), 674. https://doi.org/10.3390/pharmaceutics16050674
3. Deirram, N., Zhang, C., Kermaniyan, S. S., Johnston, A. P., & Such, G. K. (2019). pH-Responsive Polymer Nanoparticles for Drug Delivery. *Macromolecular Rapid Communications*, 40(10), 1800917. https://doi.org/10.1002/marc.201800917
4. Ezike, T. C., Okpala, U. S., Onoja, U. L., Nwike, C. P., Ezeako, E. C., Okpara, O. J., Okoroafor, C. C., Eze, S. C., Kalu, O. L., Odoh, E. C., Nwadike, U. G., Ogbodo, J. O., Umeh, B. U., Ossai, E. C., & Nwanguma, B. C. (2023). Advances in drug delivery systems, challenges and future directions. *Heliyon*, 9(7), e17488. https://doi.org/10.1016/j.heliyon.2023.e17488
5. Gurnani, P., & Perrier, S. (2020). From RAFT polymerization to smart polymeric

nanocarriers for targeted drug delivery. *Progress in Polymer Science*, 102, 101209. https://doi.org/10.1016/j.progpolymsci.2020.101209

6. Harun-Or-Rashid, M., Aktar, M. N., Hossain, M. S., Sarkar, N., Islam, M. R., Arafat, M. E., Bhowmik, S., & Yusa, S. (2023). Recent Advances in Micro- and Nano-Drug Delivery Systems Based on Natural and Synthetic Biomaterials. *Polymers*, 15(23), 4563. https://doi.org/10.3390/polym15234563
7. Ju, W., Zhou, Y., & Jiang, X. (2025). Multi-stimuli responsive block copolymers via controlled radical polymerization for programmed intracellular drug release. *Advanced Functional Materials*, 35(4), 2401928. https://doi.org/10.1002/adfm.202401928
8. Liu, Z., & Jiang, W. (2021). pH-responsive polyacrylic acid-based nanocarriers for tumor-targeted therapeutic delivery. *Biomacromolecules*, 22(8), 3165–3179. https://doi.org/10.1021/acs.biomac.1c00452
9. Mitchell, M. J., Billingsley, M. M., Haley, R. M., Wechsler, M. E., Peppas, N. A., & Langer, R. (2021). Engineering precision nanoparticles for drug delivery. *Nature Reviews Drug Discovery*, 20(2), 101–124. https://doi.org/10.1038/s41573-020-0090-8
10. Raza, A., Hayat, U., Rasheed, T., & Bilal, M. (2019). Redox-responsive disulfide-crosslinked polymeric nanocarriers for controlled oncological drug delivery. *International Journal of Pharmaceutics*, 562, 124–137. https://doi.org/10.1016/j.ijpharm.2019.03.033
11. Sharma, S., & Kumar, A. (2026). Advanced thermodynamic profiling of phase transitions in polyacrylic acid derivatives for biomedical applications. *Journal of Polymer Research*, 33(2), 88. https://doi.org/10.1007/s10965-026-04112-y
12. Sung, Y. K., & Kim, S. W. (2020). Recent advances in polymeric drug delivery systems. *Biomaterials Research*, 24(1), 12. https://doi.org/10.1186/s40824-020-00190-7
13. Wang, X., & Zhang, Y. (2022). Dual-stimuli (pH/temperature) responsive PAA-b-PEG micellar systems for co-delivery of hydrophobic chemotherapeutics. *European Polymer Journal*, 164, 110985. https://doi.org/10.1016/j.eurpolymj.2021.110985
14. Xu, L., & Feng, S. (2024). Endosomal pH tracking and intracellular trafficking kinetics of DOX-loaded smart nanocarriers. *Nano Letters*, 24(11), 3412–3420. https://doi.org/10.1021/acs.nanolett.2c05011
15. Zhang, L., & Bhatia, S. (2023). Exploiting the Warburg effect: Tumor microenvironment-targeted polyacrylic acid self-assembled architectures. *ACS Nano*, 17(9), 8123–8136. https://doi.org/10.1021/acsnano.3c01194

Bhumi Publishing, India
May 2026

AI-BASED PEST AND DISEASE DETECTION IN CROPS

Chinmayee Patra*[1] and Swatilekha Mohanta[2]

[1]Department of Agricultural Statistics,

[2]Department of Nematology,

College of Agriculture, Chiplima, Odisha University of Agriculture and Technology

*Corresponding author E-mail: chinmayee6395@gmail.com

Abstract

Crop pests and diseases are among the major constraints affecting agricultural productivity and food security worldwide. Climate change, changing cropping patterns, and excessive use of chemical pesticides have further increased the occurrence and severity of crop diseases and pest infestations. Traditional methods of pest and disease identification mainly depend on visual observation and expert knowledge, which are often time-consuming, labour-intensive, and less accurate during the early stages of infection. In recent years, Artificial Intelligence (AI) has emerged as an advanced technological tool for rapid and accurate detection of pests and diseases in crops. AI technologies such as machine learning, dseep learning, computer vision, image processing, drones, and Internet of Things (IoT) sensors are increasingly being used in precision agriculture for crop health monitoring and disease diagnosis. AI-based systems can analyze images, identify symptoms, classify diseases, and provide timely recommendations for control measures. These technologies help reduce crop losses, minimize pesticide use, improve productivity, and support sustainable agriculture. This chapter discusses the role of AI in pest and disease detection, major AI techniques used in agriculture, applications in different crops, benefits, challenges, and future prospects of AI-driven crop protection systems.

1. Introduction

Agriculture plays an important role in ensuring food security and supporting rural livelihoods. However, crop production is significantly affected by various pests and diseases that reduce yield quantity and quality. According to the Food and Agriculture Organization, plant diseases and insect pests are responsible for substantial economic losses in global agriculture every year. Traditionally, disease and pest identification in crops depends on farmers' experience and laboratory-based diagnosis. These conventional methods are often slow, expensive, and less effective for large-scale monitoring. Delayed identification frequently results in severe crop damage and excessive pesticide application.

Artificial Intelligence (AI) has emerged as a promising solution for modern crop protection. AI enables machines and computer systems to simulate human intelligence and perform tasks such as image recognition, classification, prediction, and decision-making. AI technologies combined

with computer vision, machine learning, and deep learning techniques are increasingly being used for early detection and diagnosis of crop diseases and pests. Mohanty et al. demonstrated that deep learning algorithms can accurately identify plant diseases using leaf images with high classification accuracy. Similarly, Liakos et al. highlighted the growing importance of machine learning technologies in precision agriculture and crop health monitoring. AI-based pest and disease detection systems help farmers make timely management decisions, reduce production losses, and improve sustainable agricultural practices.

2. Importance of Pest and Disease Detection in Agriculture

Crop diseases and pests directly affect agricultural productivity, food quality, and farmers' income. Early and accurate detection is essential to prevent large-scale crop damage. The major impacts of pests and diseases include:

- Reduction in crop yield
- Poor crop quality
- Economic losses to farmers
- Increased pesticide expenditure
- Environmental pollution due to excessive chemical use
- Threats to food security

Climate change has further increased the spread of several plant pathogens and insect pests. Rising temperatures and changing humidity conditions create favourable environments for disease development and pest multiplication. According to Pathak et al., changing climatic conditions are increasing the vulnerability of crops to diseases and pest outbreaks, especially in tropical and subtropical regions. AI-based technologies provide rapid and accurate crop monitoring systems that help farmers adopt timely preventive and corrective measures.

3. Artificial Intelligence Technologies Used in Crop Protection

Several AI technologies are used for pest and disease detection in crops.

3.1 Machine Learning

Machine learning enables systems to learn from agricultural data and improve prediction accuracy. ML algorithms are widely used for disease classification, pest prediction, and crop health analysis. Common machine learning algorithms include:

- Support Vector Machine (SVM)
- Random Forest
- Decision Trees
- K-Nearest Neighbour (KNN)
- Artificial Neural Networks (ANN)

Liakos et al. reported that machine learning algorithms are highly effective in agricultural classification and prediction problems.

3.2 Deep Learning

Deep learning is a subset of AI based on artificial neural networks with multiple hidden layers. Deep learning models can automatically extract important features from images without manual intervention. Convolutional Neural Networks (CNNs) are widely used for:

- Leaf disease detection
- Pest identification
- Symptom classification
- Fruit disease analysis

Mohanty et al. successfully used deep learning models for identifying multiple crop diseases from plant leaf images.

3.3 Computer Vision

Computer vision enables machines to interpret and analyze visual information from digital images and videos. It is widely used in precision agriculture for crop monitoring. Computer vision systems can:

- Detect disease symptoms
- Identify insect pests
- Analyze leaf colour and texture
- Monitor crop growth

These systems are commonly integrated with cameras, drones, and smartphones.

3.4 Internet of Things (IoT)

IoT devices such as smart sensors and wireless monitoring systems collect real-time field data related to temperature, humidity, soil moisture, and crop conditions. AI analyzes these data to predict favourable conditions for disease outbreaks and pest infestations. IoT-based systems support precision crop protection and reduce unnecessary pesticide applications.

3.5 Drones and Remote Sensing

Drones equipped with high-resolution cameras and sensors help monitor large agricultural fields efficiently. Remote sensing technologies provide:

- Early disease detection
- Pest hotspot identification
- Crop stress analysis
- Real-time field surveillance

Jha et al. emphasized that automation technologies and drone-based monitoring systems are transforming modern agriculture.

4. AI-Based Disease Detection in Crops

4.1 Leaf Disease Detection

Leaf symptoms such as spots, discoloration, wilting, and lesions are important indicators of plant diseases. AI systems analyze leaf images and identify diseases based on visual patterns.

Deep learning models can detect diseases such as:

- Rice blast
- Wheat rust
- Tomato leaf curl
- Potato late blight
- Citrus canker

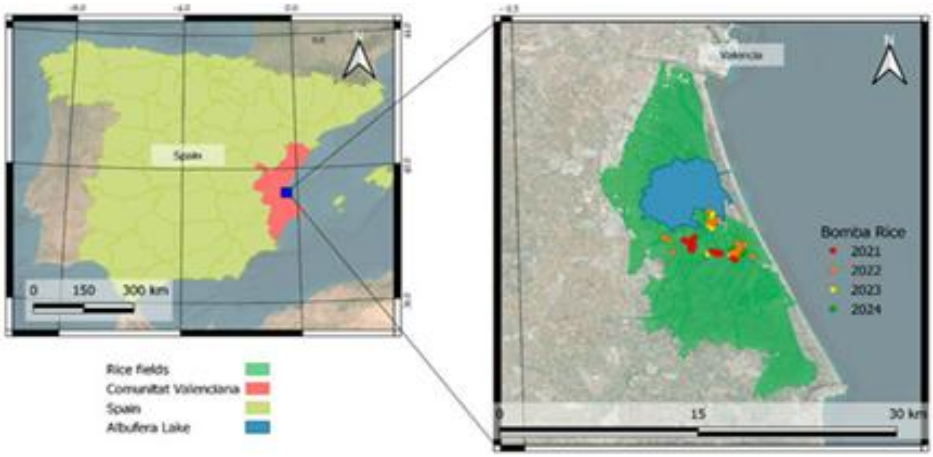


Figure 1: Early detection of Rice Blast using RS

Mohanty *et al.* reported that AI-based image classification systems achieved high accuracy in plant disease detection using large image datasets.

4.2 Fruit and Stem Disease Detection

AI technologies are also used for identifying diseases affecting fruits and stems.

Figure 2: Apple Fruit Scab recognition using Deep Learning

Examples include:

- Apple scab
- Mango anthracnose
- Banana wilt
- Grape powdery mildew

Computer vision systems analyze surface texture, colour variation, and lesion characteristics to classify diseases accurately.

4.3 Real-Time Disease Monitoring

AI-powered mobile applications allow farmers to capture crop images using smartphones and receive instant disease diagnosis.

These applications provide:

- Disease identification
- Severity assessment
- Control recommendations
- Pesticide suggestions

This reduces dependency on agricultural experts and improves accessibility for farmers.

5. AI-Based Pest Detection in Crops

5.1 Insect Pest Identification

AI systems can identify insect pests using image recognition and pattern analysis techniques.

Commonly detected pests include:

- Fall armyworm
- Aphids
- Stem borers
- Whiteflies
- Bollworms

Computer vision models analyze insect shape, size, color, and movement patterns for accurate identification.

5.2 Pest Population Monitoring

AI-based monitoring systems use sensors, smart traps, and drones to monitor pest populations in agricultural fields.

These systems help farmers:

- Predict pest outbreaks
- Determine economic threshold levels
- Apply pesticides only when necessary

This improves integrated pest management practices.

5.3 Predictive Pest Forecasting

Machine learning algorithms analyze historical weather data, humidity, temperature, and crop conditions to predict pest infestations before they occur. According to Sharma et al., AI-based predictive analytics can significantly improve precision crop management and pest forecasting systems.

6. Applications of AI-Based Pest and Disease Detection

AI technologies are increasingly being applied in different agricultural sectors.

Precision Agriculture: AI-based crop monitoring supports site-specific management and precision farming practices. Zhang et al. described precision agriculture as an advanced farming system that improves input-use efficiency and crop productivity through modern technologies.

Smart Greenhouse Monitoring: AI systems are used in greenhouses to monitor crop health, temperature, humidity, and pest incidence. Automated systems help maintain optimal growing conditions and reduce disease spread.

Mobile-Based Agricultural Advisory Services: AI-powered mobile applications provide real-time advisory services to farmers regarding disease management and pesticide application. These digital tools improve accessibility to scientific crop protection information.

Automated Spraying Systems: AI-enabled robots and drones can selectively spray pesticides only on infected plants or pest-infested areas. This reduces:

- Chemical usage
- Environmental pollution
- Production costs

Balafoutis et al. reported that precision agriculture technologies contribute to sustainable crop management and reduced environmental impacts.

7. Benefits of AI-Based Pest and Disease Detection

The adoption of AI technologies in crop protection offers several advantages.

- **Early Detection:** AI systems identify diseases and pests at early stages, preventing severe crop losses.
- **Improved Accuracy:** AI-based image recognition systems provide highly accurate diagnosis compared to manual observation.
- **Reduced Pesticide Use:** Targeted pesticide application minimizes excessive chemical use and environmental contamination.
- **Increased Crop Productivity:** Timely disease management improves crop yield and quality.
- **Labour and Time Efficiency:** Automated systems reduce manual labour and speed up monitoring processes.

- **Sustainable Agriculture:** AI technologies support environmentally friendly and resource-efficient farming systems.

8. Challenges in AI-Based Crop Protection

Despite its advantages, several challenges affect the implementation of AI technologies in agriculture.

- **High Cost of Technology:** Advanced sensors, drones, and AI systems require significant investment.
- **Limited Technical Knowledge:** Many farmers lack awareness and training regarding AI-based technologies.
- **Poor Internet Connectivity:** Rural areas often face limited digital infrastructure and internet access.
- **Data Availability Issues:** AI models require large datasets for accurate training and prediction.
- **Variability in Field Conditions:** Disease symptoms may vary due to environmental conditions, crop varieties, and growth stages, affecting model accuracy. Wolfert et al. pointed out that data management and digital infrastructure remain major constraints in smart farming systems.

9. Future Prospects of AI in Crop Protection

The future of AI in pest and disease management is highly promising. Continuous advancements in deep learning, robotics, drone technology, and remote sensing are expected to further improve crop protection systems.

Future developments may include:

- Fully autonomous crop monitoring systems
- AI-powered robotic pest control
- Smart disease forecasting platforms
- Integration of AI with blockchain technology
- Real-time satellite-based crop surveillance
- AI-driven precision spraying systems

According to Jha et al., automation and intelligent agricultural systems will play a major role in future sustainable farming. Governments and research organizations are increasingly investing in digital agriculture technologies to improve food security and reduce agricultural losses.

Conclusion

Crop pests and diseases continue to pose serious threats to global agricultural productivity and food security. Traditional disease diagnosis methods are often slow and less effective for large-

scale agricultural systems. Artificial Intelligence has emerged as a powerful tool for rapid, accurate, and efficient pest and disease detection in crops.

AI technologies such as machine learning, deep learning, computer vision, IoT, drones, and predictive analytics are transforming crop protection practices. These technologies help farmers detect diseases early, monitor pest populations, reduce pesticide use, and improve agricultural sustainability. Although challenges such as high costs, lack of digital literacy, and infrastructure limitations exist, ongoing technological advancements and policy support are expected to accelerate AI adoption in agriculture. AI-based crop protection systems have immense potential to improve food security, farmer income, and sustainable agricultural development in the future.

References

1. Balafoutis, A. T., Beck, B., Fountas, S., Tsiropoulos, Z., Vangeyte, J., van der Wal, T., Soto, I., Gómez-Barbero, M., Barnes, A., and Eory, V. (2017). Precision agriculture technologies positively contributing to GHG emissions mitigation, farm productivity and economics. *Sustainability*, 9(8), 1339, pp. 1–28.
2. Food and Agriculture Organization (FAO). (2013). *Climate-Smart Agriculture Sourcebook*. Food and Agriculture Organization of the United Nations, Rome, Italy, pp. 1–557.
3. Jha, K., Doshi, A., Patel, P., and Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture*, 2, pp. 1–12.
4. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674, pp. 1–29.
5. Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, Article 1419, pp. 1–10.
6. Pathak, H., Aggarwal, P. K., and Singh, S. D. (2012). *Climate Change Impact on Agriculture in India*. Indian Agricultural Research Institute, New Delhi, India, pp. 1–356.
7. Sharma, A., Jain, A., Gupta, P., and Chowdary, V. (2021). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, 9, pp. 4843–4873.
8. Wolfert, S., Ge, L., Verdouw, C., and Bogaardt, M. J. (2017). Big data in smart farming – A review. *Agricultural Systems*, 153, pp. 69–80.
9. Zhang, N., Wang, M., and Wang, N. (2002). Precision agriculture—A worldwide overview. *Computers and Electronics in Agriculture*, 36(2–3), pp. 113–132.
10. Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., and Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, 161, pp. 280–290.

Bhumi Publishing, India
May 2026

COUNTERFEIT DRUG DETECTION: A BRIEF ANALYTICAL PERSPECTIVES (2015–2025)

Prerna A. Parmar[1], Nivrutti D. Tank[1], Pritam S. Jain[1], Nitin G. Haswani[2] and Suraj R. Chaudhari[1,3]

[1]Department of Pharmaceutical Quality Assurance,

[2]Department of Pharmaceutical Chemistry,

Shirpur Education Society's R. C. Patel Institute of Pharmacy, Shirpur, Maharashtra, India

[3]Department of Pharmaceutical Chemistry,

R. C. Patel Institute of Pharmaceutical Education and Research, Shirpur, Maharashtra, India

Corresponding author E-mail: prernaparmar2105@gmail.com, nivruttitank@gmail.com, pritash79@yahoo.com, nitinghaswani@gmail.com, chaudharisuraj72@gmail.com

Abstract

A significant hazard to worldwide public health, counterfeit, fabricated, and subpar medications can lead to toxicity, antibiotic resistance, and treatment failure while frequently looking just like authentic items. An estimated 10–30% of medications in low- and middle-income nations don't meet quality requirements, which results in thousands of avoidable deaths every year

Keywords: Counterfeit Drugs, Falsified Medicines, HPLC, Raman Spectroscopy, LC–MS/MS, Portable Detection, Chemometrics, Pharmaceutical Quality Control.

1. Introduction

Patients trust the pharmaceutical supply chain to provide safe and genuine medicines. Counterfeit or falsified drugs break this trust and can lead to therapeutic failure, toxicity, or even death, making them a major scientific and regulatory concern. (Ozawa *et al.,* 2018; Innocent Junior Opara *et al.,* 2025; WHO). Between 2015 and 2025, counterfeit drug detection advanced significantly, but criminal networks also became more sophisticated. (WHO; lyon, n.d.; Klimovich, 2025; Ullah *et al.,* 2024; Al-Worafi, 2020a; Bakker-'t Hart *et al.,* 2021; Bhui *et al.,* 2026; Dhiman *et al.,* 2017; *WHO WHO*; Engebø *et al.,* 2017; Innocent Junior Opara *et al.,* 2025; Al-Worafi, 2020b; Bakker-'t Hart *et al.,* 2021).

2. Classification of Counterfeit Drugs

Counterfeit medications can be classified based on their composition and packaging characteristics. In terms of composition, they often contain cheap fillers such as chalk or flour and may lack the active pharmaceutical ingredient entirely. Some counterfeit drugs may still include a small amount of the correct API, typically ranging from 10–80% of the labeled dose, which can delay detection due to partial pharmacological effects. In terms of packaging,

counterfeit products range from low-quality copies with spelling errors and incorrect information to highly sophisticated replicas with convincing holograms and batch details. Advances in inexpensive printing technology have further increased the difficulty of distinguishing counterfeit packaging from genuine pharmaceutical products.(Engebø *et al.,* 2017; Seiter, 2009; Ramjiawan *et al.,* 2012; Bakker-'t Hart *et al.,* 2021; M ckey & Liang, 2011; Sugita & Miyakawa, 2010; Delepierre *et al.,* 2012).

3. Conventional Detection Approaches

Visual and physical inspection involves evaluating the appearance, texture, odor, and packaging of medications to detect obvious counterfeit products. This organoleptic assessment can help pharmacists identify low-quality or clearly fake drugs in routine practice. However, due to the increasing sophistication of counterfeit medicines, visual inspection alone is no longer reliable for detecting advanced fakes. Simple physicochemical screening methods, such as thin-layer chromatography (TLC), are widely used in resource-limited settings for basic drug verification. TLC helps compare samples with reference standards to confirm the presence of active pharmaceutical ingredients. High-performance TLC (HPTLC) further improves accuracy, sensitivity, and reproducibility while remaining more affordable and easier to operate than HPLC (Martino *et al.,* 2010; Bottoni & Caroli, 2019; Zhang *et al.,* 2020; Singh *et al.,* 2009).

4. Advanced Analytical Techniques

4.1 Chromatographic Techniques (Sherma & Rabel, 2019; Martino *et al.,* 2010)

4.1.1 High-Performance Liquid Chromatography (HPLC)

For many years, HPLC has been the main method used in industrial and regulatory labs for pharmaceutical quality monitoring. Reversed-phase HPLC uses a non-polar C18 column and a polar mobile phase to separate chemicals according to variations in polarity and molecular interactions. For the majority of pharmaceutical analytes, contemporary C18 columns offer great resolution and effective separation (Rodionova & Pomerantsev, 2010; Zou *et al.,* 2018).

4.1.2 Ultra-High-Performance Liquid Chromatography (UHPLC)

UHPLC is a more sophisticated version of HPLC that achieves faster analysis and better resolution by using higher pressures and smaller particle sizes. It is very helpful for extensive surveillance testing since it may cut the analysis time from 20 to 30 minutes to less than five minutes. Research on fake antimalarials has demonstrated that UHPLC offers far faster run times and higher sensitivity than traditional HPLC (Anzanello *et al.,* 2014).

4.1.3 Gas Chromatography (GC)

Gas chromatography (GC) plays a complementary role in the identification of counterfeit drugs, primarily for the analysis of volatile substances such adulterants, residual solvents, and contaminants. Because it can detect solvent fingerprints, GC-MS is particularly useful for

tracking down the manufacturing source of counterfeit goods and occasionally connecting them to particular illicit production sites (Höllein *et al.,* 2016).

4.1.4 High-Performance Thin-Layer Chromatography (HPTLC)

It provides superior analytical performance over conventional TLC while being less expensive and more portable than HPLC or LC-MS/MS, HPTLC is extremely beneficial in the detection of counterfeit drugs. Modern HPTLC devices are perfect for high-throughput, cost-effective surveillance programs since they enable the simultaneous analysis of many samples and provide accurate API quantification (Hall, 2012).

4.1.5 LC–MS/MS

One of the most effective methods for detecting counterfeit drugs is LC–MS/MS, which combines highly sensitive mass analysis with chromatographic separation. It makes it possible to precisely identify and measure APIs, contaminants, and adulterants, including new compounds, and it also aids in tracking the source of counterfeit medication production (Krakowska *et al.,* 2016; Baert & De Spiegeleer, 2010).

4.2 Spectroscopic Techniques

4.2.1 UV–Visible Spectrophotometry

A straightforward and popular spectroscopic method for estimating API content is UV-visible spectrophotometry, which analyzes radiation absorption in the 200–800 nm range. For the analysis of several medicinal substances, it offers a quick, economical, and pharmacopoeially recognized approach (Deconinck *et al.,* 2012; Martino *et al.,* 2010; Puchert *et al.,* 2010).

4.2.2 Fourier-Transform Infrared Spectroscopy (FTIR)

By measuring the absorption of mid-infrared light, FTIR spectroscopy creates a distinct molecular fingerprint that is used to identify pharmaceutical substances. Modern ATR-FTIR systems are helpful for field screening and counterfeit detection because they provide quick, non-destructive examination (Custers *et al.,* 2015).

4.2.3 Raman Spectroscopy

Raman spectroscopy's quick, non-destructive, and preparation-free analysis has made it one of the most crucial methods for detecting counterfeit drugs. For field screening, portable handheld Raman instruments are very helpful since they can employ spectral libraries to identify medications through packaging in real time (Lu *et al.,* 2013; Dégardin *et al.,* 2017).

4.2.4 Near-Infrared Spectroscopy (NIR)

Because of its complex spectra, NIR spectroscopy is frequently used in conjunction with chemometric data processing to study pharmaceutical substances in the 780–2,500 nm range. With just one measurement, it can quickly offer details on API identity, content, polymorphic shape, and moisture levels (Rodionova & Pomerantsev, 2010).

4.3 Mass Spectrometry-Based Methods

4.3.1 MALDI-TOF Mass Spectrometry

MALDI-TOF MS is a potent method for detecting counterfeit drugs that offers molecular identification capabilities comparable to LC-MS/MS with quicker and easier sample preparation. It produces analyte ions through laser-induced ionization, enabling precise molecular mass determination and trustworthy pharmaceutical component identification (Bronzel *et al.,* 2017; Ye *et al.,* 2008; De Almeida *et al.,* 2024).

4.3.2 High-Resolution Mass Spectrometry (HRMS)

The precise elemental composition of unknown compounds can be ascertained with the aid of highly accurate mass measurements provided by high-resolution mass spectrometry (HRMS), such as Q-TOF and Orbitrap systems (Mestria *et al.,* 2022; Lee *et al.,* 2021).

4.4 Nuclear Magnetic Resonance (NMR) Spectroscopy

By examining the magnetic characteristics of atomic nuclei like ^{1}H and ^{13}C, NMR spectroscopy offers comprehensive structural information. It is one of the most dependable methods for identifying compounds since it produces extremely exact molecular fingerprints (Holzgrabe & Malet-Martino, 2011; Kuballa *et al.,* 2018).

4.5 X-ray Diffraction, Hyperspectral Imaging, and Terahertz Spectroscopy

It can identify wrong polymorphs, crystalline adulterants, or the lack of crystalline API even when traditional chemical tests seem normal, it is very helpful in the detection of counterfeit drugs (Shen *et al.,* 2021; Edwards *et al.,* 2012).

5. Emerging and Rapid Detection Technologies (Zou *et al.,* 2018; Mackey & Nayyar, 2017)

5.1 Paper-Based Analytical Devices (PADs)

Simple analytical tests carried out on paper or nitrocellulose membranes using capillary action to transfer samples without the need for pumps or power sources are known as paper-based analytical devices (PADs; Bolla *et al.,* 2020).

5.2 Lab-on-a-Chip Systems

Advanced miniaturized systems known as lab-on-a-chip (LOC) microfluidic platforms combine all analytical procedures, such as sample handling, reaction, separation, and detection, onto a tiny, portable chip (McNeill *et al.,* 2021).

5.3 Biosensors

In order to translate target binding into quantifiable electrical or optical signals, biosensors integrate biological recognition components like antibodies, aptamers, or molecularly imprinted polymers with physicochemical transducers (Sanvicens *et al.,* 2011).

5.4 Smartphone-Assisted Detection

By combining inexpensive optical attachments with smartphone apps that use the phone's camera and processor to evaluate colorimetric or fluorometric signals, smartphone-assisted detection devices have become effective instruments for monitoring counterfeit drugs (Ali *et al.,* 2024).

5.5 Portable Miniaturised Laboratory Instruments

The difference between portable screening tools and laboratory-grade equipment is closing quickly. While portable mass spectrometers like the Torion T-9 and 908 Devices M908 offer quick on-site MS identification, handheld Raman and FTIR spectrometers are already often utilized for field detection (Alonzo *et al.,* 2022).

6. Chemometrics and data analysis Multivariate Analysis and PCA

PCA usually displays real items in a tight cluster while fake samples show up as outliers because of variations in excipients. Support vector machines (SVM), k-nearest neighbor (kNN), partial least squares discriminant analysis (PLS-DA), and linear discriminant analysis (LDA) (Anzanello *et al.,* 2013; Rodionova *et al.,* 2019; Alley & Elkins, 2026).

7. Case Studies

Table 1: Selected Real-World Case Studies of Counterfeit Drug Detection (2015–2025) (Newton *et al.,* 2008; Ziavrou *et al.,* 2022).

Drug / Region	Analytical Method	Key Finding	Regulatory Outcome
Artesunate, SE Asia (2015–2018)	HPLC, Raman, LC–MS/MS	0–60% of label claim; chalk/starch filler	WHO Medical Product Alerts; supply chain investigation
Ciprofloxacin, Sub-Saharan Africa (2016)	HPLC, FTIR	15–40% of label claim; incorrect excipients	Regulatory seizure; market withdrawal
Sildenafil, EU online markets (2019)	UHPLC–MS/MS, HRMS, NMR	Tadalafil substitution; novel PDE-5 analogues identified	Interpol Operation Pangea seizures

8. Future Perspectives

Over the next ten years, a number of convergent technological and legislative trends are anticipated to drastically alter the CF detection landscape.

8.1 AI-Embedded Spectroscopic Instruments

Pharmaceutical samples can be quickly and automatically classified as pass or fail with uncertainty estimates by integrating deep learning into portable spectrometers with cloud-updated reference databases. Developing validated worldwide spectral libraries, creating legal guidelines for machine learning tools, and guaranteeing fair cloud access in low-resource environments are the primary obstacles (Alves *et al.,* 2025).

8.2 Blockchain and Supply Chain Security

Blockchain-based distributed ledger systems can improve pharmaceutical supply chain transparency by creating tamper-evident, verifiable records of each custody transfer from manufacturer to pharmacy. This enables reliable product provenance tracking beyond traditional paper-based systems. Proof-of-concept initiatives such as the MediLedger Network and efforts by the European Medicines Agency (EMA) have explored this approach. When combined with RFID serialization, smart contracts, and AI-based anomaly detection, blockchain systems could significantly strengthen protection against counterfeit medicine infiltration (M R *et al.,* 2021).

8.3 Green Analytical Chemistry

The usage of solvents, energy consumption, and chemical waste make the environmental impact of pharmaceutical quality control more significant. The Green Analytical Chemistry (GAC) framework encourages non-destructive analytical techniques, smaller sample quantities, and lower reagent use. Because they don't require reagents and produce no waste, methods like Raman, NIR, and terahertz imaging are thought to be environmentally friendly. The environmental impact of HPLC-based analysis is lessened by advancements in chromatographic techniques, such as smaller columns, sub-2 µm particles, and more environmentally friendly solvents rather than acetonitrile (Mehta *et al.,* 2024).

8.4 The Portable Laboratory of the Future

The integration of wireless connectivity, AI-driven data processing, and tiny analytical tools into portable "laboratories" that non-specialists can use as needed is a key long-term objective. There are already individual technologies like smartphone-linked ML spectrum analysis, backpack NMR equipment, and suitcase-sized mass spectrometers. Integrating them into accessible, validated systems and creating the infrastructure and training required for successful deployment particularly in low- and middle-income nations—are the main challenges that lie ahead (Coppey *et al.,* 2020).

Conclusion

The decade from 2015 to 2025 has produced genuine and substantial progress in the analytical detection of counterfeit and falsified medicines. Portable Raman and FTIR spectrometers have moved from specialist laboratory tools to field-deployed screening instruments used by customs officers and public health inspectors. Advanced analytical approaches have facilitated the identification of counterfeit compounds and the dismantling of distribution networks. Paper-based analytical devices, biosensors, and smartphone-assisted detection platforms have begun to genuinely democratize quality screening. In addition, the application of machine learning to spectral classification has pushed classification accuracy to levels that were difficult to imagine only a few years ago.

And yet the problem has not been solved. According to World Health Organization estimates, substandard and falsified medications continue to cost hundreds of thousands of lives annually, while the economic burden amounts to tens of billions of dollars. The online counterfeit pharmaceutical market continues to expand despite enforcement efforts, and a significant resource gap remains between high-income countries and the low- and middle-income countries most heavily affected by the counterfeit drug burden.

Analytical innovation represents a vital, but insufficient, solution to this challenge. Effective mitigation also requires sustained investment in laboratory infrastructure in low- and middle-income countries, robust international data-sharing and pharmacovigilance coordination, regulatory frameworks capable of keeping pace with counterfeit innovation, and strong political commitment to prioritizing medicine quality within universal health coverage systems. Scientists and regulatory professionals working in this field recognize that every improvement in detection capability ultimately translates into lives saved. It is this direct connection between scientific advancement and patient outcomes that underscores both the significance and urgency of continuing efforts to combat counterfeit and falsified medicines worldwide.

References

1. Ali, D. S., Hassan, R. O., Othman, H. O., Taha, H. T., Mousavi Khaneghah, A., & Smaoui, S. (2024). Revolutionizing detection: Smartphone-powered colorimetry for the drugs and food analysis. *Microchemical Journal, 205*, 111228. https://doi.org/10.1016/j.microc.2024.111228
2. Alley, S. J. M., & Elkins, K. M. (2026). Artificial intelligence and machine learning in forensic DNA analysis. In *Investigative and Predictive DNA Testing* (pp. 315–324). Elsevier. https://doi.org/10.1016/B978-0-443-30128-5.00015-1
3. Alonzo, M., Alder, R., Clancy, L., & Fu, S. (2022). Portable testing techniques for the analysis of drug materials. *WIREs Forensic Science, 4*(6), e1461. https://doi.org/10.1002/wfs2.1461
4. Alves, E., Gurupadayya, B. M., & Prabhakaran, P. (2025). Next-generation miniaturized separation platforms: Converging detection, automation, and sustainable design for intelligent analytical science. *Critical Reviews in Analytical Chemistry*, 1–50. https://doi.org/10.1080/10408347.2025.2553122
5. Al-Worafi, Y. M. (2020a). Counterfeit and substandard medications. In *Drug Safety in Developing Countries* (pp. 119–126). Elsevier. https://doi.org/10.1016/B978-0-12-819837-7.00010-8
6. Anzanello, M. J., Ortiz, R. S., Limberger, R., & Mariotti, K. (2014). A framework for selecting analytical techniques in profiling authentic and counterfeit Viagra and Cialis. *Forensic Science International, 235*, 1–7. https://doi.org/10.1016/j.forsciint.2013.12.005

7. Baert, B., & De Spiegeleer, B. (2010). Quality analytics of internet pharmaceuticals. *Analytical and Bioanalytical Chemistry, 398*(1), 125–136. https://doi.org/10.1007/s00216-010-3912-4
8. Bakker-'t Hart, I. M. E., Ohana, D., & Venhuis, B. J. (2021). Current challenges in the detection and analysis of falsified medicines. *Journal of Pharmaceutical and Biomedical Analysis, 197*, 113948. https://doi.org/10.1016/j.jpba.2021.113948
9. Bhavna, Ojha, A., & Bhargava, S. (2022). International Council for Harmonisation (ICH) guidelines. In *Regulatory Affairs in the Pharmaceutical Industry* (pp. 47–74). Elsevier. https://doi.org/10.1016/B978-0-12-822211-9.00008-3
10. Bhui, U., Das, J., Chakraborty, G. S., Sen, S., Nayak, S., Kumar, B., & Bandopadhyay, S. (2026). Machine learning and AI in pharmaceutical analysis: A new era of precision and efficiency. *Current Pharmaceutical Design, 32*. https://doi.org/10.2174/0113816128418441251212065440
11. Bolla, A. S., Patel, A. R., & Priefer, R. (2020). The silent development of counterfeit medications in developing countries—A systematic review of detection technologies. *International Journal of Pharmaceutics, 587*, 119702. https://doi.org/10.1016/j.ijpharm.2020.119702
12. Bottoni, P., & Caroli, S. (2019). Fake pharmaceuticals: A review of current analytical approaches. *Microchemical Journal, 149*, 104053. https://doi.org/10.1016/j.microc.2019.104053
13. Bronzel, J. L., Milagre, C. D. F., & Milagre, H. M. S. (2017). Analysis of low molecular weight compounds using MALDI- and LDI-TOF-MS: Direct detection of active pharmaceutical ingredients in different formulations. *Journal of Mass Spectrometry, 52*(11), 752–758. https://doi.org/10.1002/jms.3984
14. Coppey, F., Bécue, A., Sacré, P.-Y., Ziemons, E. M., Hubert, P., & Esseiva, P. (2020). Providing illicit drugs results in five seconds using ultra-portable NIR technology: An opportunity for forensic laboratories to cope with the trend toward the decentralization of forensic capabilities. *Forensic Science International, 317*, 110498. https://doi.org/10.1016/j.forsciint.2020.110498
15. Custers, D., Cauwenbergh, T., Bothy, J. L., Courselle, P., De Beer, J. O., Apers, S., & Deconinck, E. (2015). ATR-FTIR spectroscopy and chemometrics: An interesting tool to discriminate and characterize counterfeit medicines. *Journal of Pharmaceutical and Biomedical Analysis, 112*, 181–189. https://doi.org/10.1016/j.jpba.2014.11.007
16. De Almeida, C. M., Dos Santos, N. A., Lacerda, V., Ma, X., Fernández, F. M., & Romão, W. (2024). Applications of MALDI mass spectrometry in forensic science. *Analytical and*

Bioanalytical Chemistry, 416(24), 5255–5280. https://doi.org/10.1007/s00216-024-05470-y

17. Deconinck, E., Sacré, P. Y., Coomans, D., & De Beer, J. (2012). Classification trees based on infrared spectroscopic data to discriminate between genuine and counterfeit medicines. *Journal of Pharmaceutical and Biomedical Analysis, 57*, 68–75. https://doi.org/10.1016/j.jpba.2011.08.036
18. Dégardin, K., Desponds, A., & Roggo, Y. (2017). Protein-based medicines analysis by Raman spectroscopy for the detection of counterfeits. *Forensic Science International, 278*, 313–325. https://doi.org/10.1016/j.forsciint.2017.07.012
19. Delepierre, A., Gayot, A., & Carpentier, A. (2012). Update on counterfeit antibiotics worldwide: Public health risks. *Médecine et Maladies Infectieuses, 42*(6), 247–255. https://doi.org/10.1016/j.medmal.2012.04.007
20. Dhiman, A., Kalra, P., Bansal, V., Bruno, J. G., & Sharma, T. K. (2017). Aptamer-based point-of-care diagnostic platforms. *Sensors and Actuators B: Chemical, 246*, 535–553. https://doi.org/10.1016/j.snb.2017.02.060

Current Research in Chemical, Biological, Healthcare and Data Science

(ISBN: 978-93-47587-18-4)

About Editors

Dr. Raman Singh (née Bhadauria) is Vice President (International) of Graduate Women in Science (GWIS), USA, and Associate Professor of Chemistry at Amity School of Engineering & Technology, Amity University Madhya Pradesh, Gwalior. She holds M.Sc., M.Phil., and Ph.D. degrees in Organic and Medicinal Chemistry from Jiwaji University, Gwalior, and a Master's degree in Intellectual Property Law from IGNOU, New Delhi. With over 13 years of experience in teaching, research, and industry, she has worked at the National Dope Testing Laboratory and in renewable energy research. Her expertise includes medicinal chemistry, organic synthesis, heterocyclic compounds, antimicrobial drug discovery, drug resistance, traditional knowledge, intellectual property rights, and green chemistry. She has published 23 research papers, 25 book chapters, 15 books, and filed a patent. Dr. Singh has received several prestigious awards for excellence in science, research, and higher education nationally.

Dr. E. Thenpandiyan, M.Sc., Ph.D., is currently serving as Assistant Professor (Research) in the Department of Physics at the Academy of Maritime Education and Training (Deemed to be University), Chennai, Tamil Nadu, India. He earned his Ph.D. in Physics from Annamalai University, Chidambaram. An active researcher, he has published 19 papers in peer-reviewed SCI-indexed journals and holds one Indian patent. He has presented seven research papers at international conferences and serves as a reviewer for five journals published by Springer Nature. Dr. Thenpandiyan has also contributed six book chapters and authored one book with national and international publishers. His research focuses on nanostructured materials, polymer-based nanophotocatalysts, rare-earth ion-based nano-supercapacitors, and biomedical applications. He is dedicated to advancing interdisciplinary research and innovation in materials science, energy storage, and healthcare technologies.

Ms. Pratiksha is an experienced educator with more than 15 years of teaching experience in higher education. She completed her M.Sc. in Chemistry from Bareilly College, Bareilly, and is currently pursuing her Ph.D. in Chemistry from Maharaja Agrasen Himalayan Garhwal University, Uttarakhand. She is presently working as an Assistant Professor in the Department of Chemistry at Rajshree Group of Institutions, Bareilly. She actively contributes to teaching, academic mentoring, and scientific research. Ms. Pratiksha has published five research papers in international journals and has participated in several national and international conferences. Her academic and research interests include environmental studies, chemical sciences, sustainable research practices, and innovative approaches to quality education. She remains committed to promoting academic excellence, scientific inquiry, and research-oriented learning among students while contributing meaningfully to higher education and scholarly development.

Atul Goyal is a dedicated academician and engineering professional currently serving as Assistant Professor and Coordinator of Petroleum Engineering at Guru Kashi University, Talwandi Sabo. He is actively engaged in teaching, research, and academic development in the field of petroleum engineering. With a strong commitment to higher education, he has authored and published more than five books and contributed numerous research papers in reputed journals and conference proceedings. His areas of academic and research interest include petroleum refining, solid waste management, sustainable engineering practices, and environmental protection. Mr. Goyal is recognized for his efforts in promoting technical knowledge, research culture, and industry-oriented learning among students. Through his scholarly contributions and academic leadership, he continues to support innovation, professional development, and the advancement of engineering education and research in emerging technological domains.

Bhumi Publishing

Bhumi Publishing,
Kolhapur, Maharashtra, India
bhumipublishing@gmail.com
www.bhumipublishing.com

ISBN 978-93-475-8718-4
9 789347 587184