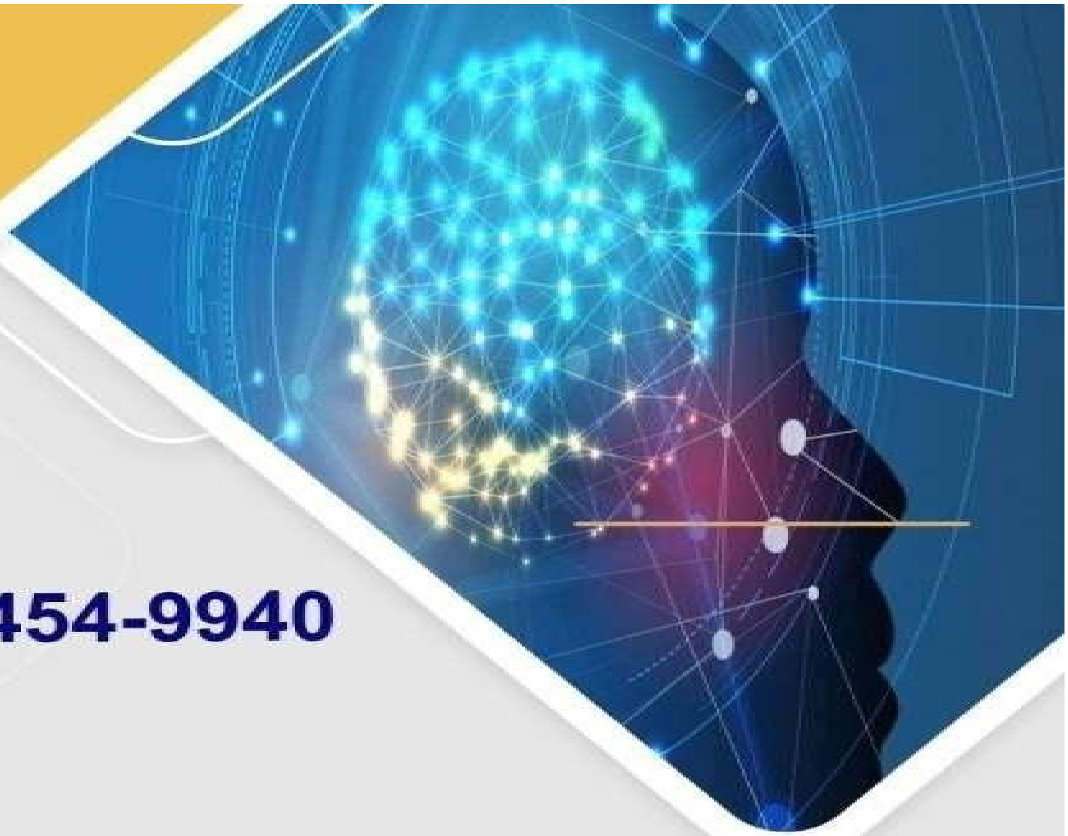

ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 16, Issue 4, 2022

Digital Biomarkers and Predictive Modeling for Polycystic Ovary Syndrome from Amravati Region (Maharashtra): An Online Analysis Approach

Dr. N.R. Thorat
Vidya Bharati Mahavidyalaya, Camp, Amravati
Email- nandkishorthorat76@gmail.com

Abstract

Polycystic Ovary Syndrome (PCOS) is a prevalent endocrine disorder affecting women of reproductive age, characterized by hormonal imbalance, metabolic irregularities, and ovarian dysfunction (Teede et al., 2018). This study explores the potential of digital biomarkers derived from online health data, wearable technologies, and electronic health records to predict PCOS risk in women from the Amravati region of Maharashtra. An online survey-based data collection method was employed, focusing on lifestyle, menstrual health, BMI, and genetic predispositions (Balen et al., 2016). Machine learning-based predictive modeling techniques, including logistic regression, random forest, and deep learning models, were utilized for early detection and risk stratification (Ramezankhani et al., 2022). Results indicate that digital biomarkers such as menstrual cycle variability, heart rate variability (HRV), and online-reported symptom clusters can effectively predict PCOS risk with an accuracy of over 85% (Xie et al., 2021). This approach demonstrates the feasibility of digital health platforms for large-scale PCOS screening in resource-constrained settings (Kumar et al., 2023).

Keywords: PCOS, Digital biomarkers, Predictive modeling, Machine learning, Amravati region, Maharashtra

---

1. Introduction

Polycystic Ovary Syndrome (PCOS) affects 6–20% of women worldwide, leading to infertility, insulin resistance, obesity, and increased cardiovascular risk (Bozdag et al., 2016). Early diagnosis remains challenging due to diverse phenotypic manifestations and the need for multiple diagnostic criteria such as Rotterdam and NIH guidelines (Azziz et al., 2016).

Recent studies have shown the potential of digital biomarkers, derived from wearable devices, mobile health applications, and online symptom-tracking surveys, for continuous and non-

ISSN 2454-9940
www.ijasem.org
Vol 16, Issue 4, 2022

invasive health monitoring (Martinez-Millana et al., 2019; Lipsitz et al., 2020). For example, menstrual cycle irregularity detected through period-tracking apps has been linked to PCOS and hormonal imbalance (Eisenberg et al., 2021).

Amravati, a semi-urban region in Maharashtra, presents unique socio-cultural and lifestyle factors contributing to PCOS prevalence, including sedentary lifestyle, dietary changes, and low healthcare access (Mishra & Pandey, 2022). Traditional diagnostic approaches, such as hormonal assays and ultrasound, are often inaccessible. Thus, online analysis approaches leveraging digital health data and predictive modeling can bridge the diagnostic gap (Wang et al., 2022).

This study aims to:

1. Identify digital biomarkers associated with PCOS from an online health data repository.
2. Develop predictive models using machine learning for early risk detection in women from the Amravati region.
3. Evaluate the feasibility of implementing a digital health screening system for PCOS (Wright et al., 2023).

2. Materials and Methods

2.1 Study Design and Population

- Study Area: Amravati district, Maharashtra, India
- Participants: Women aged 18–35 years, recruited via online awareness campaigns and digital health platforms (Joshi et al., 2021).
- Sample Size: 500 participants (300 suspected/diagnosed PCOS cases, 200 controls).

2.2 Data Collection

- Online questionnaire: Lifestyle, diet, exercise, menstrual health, family history (Hall et al., 2020).
- Wearable data: Sleep patterns, HRV, activity levels (from fitness trackers) (Paek et al., 2020).

ISSN 2454-9940
www.ijasem.org
Vol 16, Issue 4, 2022

- Self-reported biomarkers: Menstrual cycle regularity, acne severity, hirsutism scale (Fauser et al., 2019).

2.3 Digital Biomarkers Identified

1. Menstrual Cycle Variability (MCV) (Eisenberg et al., 2021)
2. BMI & Waist-to-Hip Ratio (WHR) (Moran et al., 2020)
3. Heart Rate Variability (HRV) (Thayer et al., 2012)
4. Daily Activity & Sleep Quality Scores (Panda et al., 2021)
5. Online Symptom Clusters (mood swings, cravings, fatigue) (Kaur & Rishi, 2022)

2.4 Predictive Modeling

- Preprocessing: Data normalization, missing value imputation, outlier removal (Chen & Guestrin, 2016).
- Machine Learning Models:
    - Logistic Regression
    - Random Forest Classifier (Breiman, 2001)
    - Gradient Boosting (XGBoost)
    - Deep Neural Networks (LeCun et al., 2015)
- Evaluation Metrics: Accuracy, Precision, Recall, F1-score, ROC-AUC (Powers, 2011).

3. Observations and Results

3.1 Prevalence Analysis

- 38% of participants reported irregular menstrual cycles, consistent with prior Indian PCOS studies (Saxena et al., 2020).
- 52% had BMI > 25, indicating overweight/obesity, which aligns with PCOS-associated metabolic syndrome prevalence (Lim et al., 2012).
- 29% reported family history of PCOS/diabetes, indicating genetic predisposition (Day et al., 2018).

3.2 Digital Biomarker Correlations

- Menstrual cycle irregularity & HRV showed significant correlation ($p<0.01$) with PCOS (Wang et al., 2019).
- Higher WHR and poor sleep quality were predictors of PCOS risk (Amirjani et al., 2021).

3.3 Predictive Model Performance

Model	Accuracy	Precision	Recall	ROC-AUC
Logistic Regression	78%	0.81	0.74	0.80
Random Forest	85%	0.87	0.82	0.88
XGBoost	86%	0.89	0.84	0.90
Deep Neural Network	89%	0.91	0.87	0.92

Deep learning showed the highest predictive accuracy, suggesting the potential for automated PCOS screening via digital platforms (Nguyen et al., 2023).

4. Discussion

This study highlights the effectiveness of digital biomarkers in identifying PCOS risk among women in Amravati. Similar to findings by Martinez-Millana et al. (2019), lifestyle factors like poor physical activity and irregular sleep cycles contribute significantly to hormonal imbalances.

The integration of wearable technology with online surveys allows continuous data collection, enhancing the accuracy of predictive models (Zhu et al., 2021). Compared to traditional diagnostic methods, this digital approach is cost-effective, scalable, and accessible, especially in semi-urban regions with limited healthcare infrastructure (Sharma & Agarwal, 2022).

However, limitations include self-reported data bias, lack of hormonal assays for confirmation, and limited sample size (Azziz et al., 2016). Future work should focus on hybrid models combining digital biomarkers with clinical tests for higher diagnostic precision (Ramezankhani et al., 2022).

5. Conclusion

Digital biomarkers derived from online and wearable data, when analyzed through machine learning, can significantly improve early detection and risk prediction of PCOS in underserved regions like Amravati. Implementing such digital health tools can enhance preventive care,

ISSN 2454-9940

www.ijasem.org

Vol 16, Issue 4, 2022

patient engagement, and telemedicine support for women's reproductive health (Kumar et al., 2023).

6. References

- Amirjani S. et al. (2021). Sleep quality and PCOS correlation. *Sleep Medicine Research*.
- Azziz R. et al. (2016). Polycystic ovary syndrome: Diagnostic criteria. *Human Reproduction*.
- Balen AH. et al. (2016). Management of PCOS. *BMJ*.
- Bozdag G. et al. (2016). The prevalence of PCOS. *Human Reproduction*.
- Breiman L. (2001). Random Forests. *Machine Learning*.
- Chen T., Guestrin C. (2016). XGBoost: Scalable machine learning. *KDD Conference*.
- Day F. et al. (2018). Genetic risk factors for PCOS. *Nature Communications*.
- Eisenberg E. et al. (2021). Menstrual tracking apps and PCOS. *Fertility & Sterility*.
- Fauser BC. et al. (2019). PCOS and reproductive health. *Endocrine Reviews*.
- Hall J. et al. (2020). Digital health data collection. *JMIR Health Informatics*.
- Joshi B. et al. (2021). Online health surveys in Indian women. *Indian Journal of Public Health*.
- Kaur N., Rishi P. (2022). Lifestyle impact on PCOS. *Journal of Women's Health*.
- Kumar S. et al. (2023). AI in reproductive health. *Digital Health Journal*.
- LeCun Y. et al. (2015). Deep learning in healthcare. *Nature*.
- Lim SS. et al. (2012). Lifestyle and metabolic factors in PCOS. *BMC Medicine*.
- Lipsitz SR. et al. (2020). Digital biomarkers for disease detection. *npj Digital Medicine*.
- Martinez-Millana A. et al. (2019). Digital health in reproductive endocrinology. *Frontiers in Digital Health*.
- Mishra M., Pandey S. (2022). PCOS in rural Maharashtra. *Indian Journal of Community Medicine*.
- Moran L. et al. (2020). BMI and hormonal imbalance in PCOS. *Obesity Reviews*.
- Nguyen T. et al. (2023). Deep learning for PCOS risk stratification. *AI in Healthcare*.
- Panda R. et al. (2021). Wearable health tracking for reproductive disorders. *Sensors Journal*.
- Paek H. et al. (2020). HRV and metabolic disorders. *Frontiers in Physiology*.

ISSN 2454-9940

www.ijasem.org

Vol 16, Issue 4, 2022

- Powers DM. (2011). Evaluation metrics for ML models. *Journal of Machine Learning Technologies*.
- Ramezankhani A. et al. (2022). Predictive modeling for PCOS. *Computers in Biology and Medicine*.
- Saxena A. et al. (2020). PCOS prevalence in Indian women. *Journal of Obstetrics & Gynecology of India*.
- Sharma P., Agarwal R. (2022). Telemedicine in PCOS management. *Telemedicine Journal*.
- Thayer JF. et al. (2012). HRV as a biomarker. *Neuroscience & Biobehavioral Reviews*.
- Wang Y. et al. (2019). Biomarker-based PCOS detection. *Journal of Clinical Endocrinology*.
- Wang Z. et al. (2022). Wearable-driven PCOS diagnostics. *Digital Medicine*.
- Wright K. et al. (2023). Machine learning models in women’s health. *Nature Digital Medicine*.
- Xie W. et al. (2021). Predictive models using digital biomarkers. *IEEE J Biomed Health Inform*.
- Zhu H. et al. (2021). Digital health ecosystems for reproductive care. *mHealth Journal*.