PERINATAL JOURNAL

Original Article

Perinatal Journal 2025; 33(2):624-631

https://doi.org/10.57239/prn.25.03320068

Unsupervised clustering of obstructed labor risk: A novel pattern recognition model using CPD indicators and fetal positional metrics

Hajar Suliman[1], Awadalla Abdelwahid[1]*, Elsadig Shiekedien[2], Ahazeej Gurashi[3], Mandour Mohamed Ibrahim[4]

[1]Department of Obstetrics and Gynecology, Alneelain University, Khartoum, Sudan
[2]Assistant Professor, Department of Obstetrics and Gynecology, Faculty of Medicine, University of Tabuk, Saudi Arabia
[3]Department of Obstetrics and Gynecology, Alhayat National Hospital, Khamis Mushait, Saudi Arabia
[4]College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia

Abstract

Obstructed labor is a leading source of maternal and neonatal morbidity in low-resource and internally displaced populations. Although clinically relevant, predictive modeling for obstructed labor is not fully developed. To construct and validate an unsupervised clustering model that stratifies obstructed labor risk based on interpretable and context-sensitive features. A K-means clustering algorithm was performed on a real-world dataset containing second-stage cesarean deliveries from Alhasahisa Teaching Hospital. The input features were fetal station, valvular swelling, molding, and cephalopelvic disproportion indicators. Visualization and internal validation using Principal Component Analysis (PCA) and radar charts. Three distinct clusters were classified as follows: Cluster A (low risk): Station 0/+1, no swelling, Apgar ~8.2, NICU ~8.8% Cluster B (moderate risk): Station -1, moderate swelling, Apgar ~6.1, NICU ~26.5% Cluster C (high risk): Station -2/-3, severe swelling, Apgar ~4.3, NICU ~53.1%. Model had distinct separation and interpretability without labeled outcomes. Conclusion. Thus, this is the first-of-its-kind unsupervised model for stratification of obstructed labor risk based on labor-specific features. It provides a scalable, ethical decision support tool for early triage in low-resource settings, and is consistent with global maternal health priorities.

Keywords: Obstructed labor, Unsupervised learning, Clustering model, Cephalopelvic disproportion, Maternal health, Ethical AI

Introduction

Obstructed labor continues to be a leading cause of maternal and neonatal morbidity, particularly among low-resource and forcibly resettled societies. Obstructed labor, defined as the failure of the fetus to go through the birth canal when the uterus is contracting intensely, is often preventable but is underdiagnosed until complications emerge [1]. Older diagnostic approaches strongly depend on subjective clinical judgment that can be delayed and/or not available in overburdened healthcare systems. This gap points to a pressing need for tools which predict early and ethically for obstructed labor risk. Artificial intelligence (AI) and machine learning (ML) developments have altered maternal health analytics in recent years. Supervised models have been developed for predicting preterm birth [2], labor duration [3], and delivery complications [4]. Yet obstructed labor is underrepresented in the predictive.

Modeling Literature, and no ML models exist for obstructed labor in this area of focus, specifically anatomical and positional indicators of obstructed labor. This is especially worrying as obstructed labor during second-stage deliveries is associated with high rates of cesarean section and NICU admissions [1]. Cephalopelvic Disproportion (CPD), fetal malposition, and pelvic parameters are crucial aspects of obstructed labor pathophysiology. This is where studies have examined for CPD prediction using MRI-based pelvimetry and anthropometric modeling [5], but which also requires complex resources and is not practical in low-income environments. Fetal position, station, and valvular swelling — easily visible in labor — offer a less abstract risk stratification opportunity. However, these features have not been included in unsupervised learning architectures. We

Correspondence: Awadalla Abdelwahid, Department of Obstetrics and Gynecology, Alneelain University, Khartoum, Sudan., **e-mail:** Awadalla Abdelwahid,
Received: October 11, 2025 **Accepted:** November 23, 2025

present a novel K-means clustering model that detects potential latent risk clusters for obstructed labor based on CPD indicators, fetal posture and maternal pelvic parameters. Instead of requiring labeled outcomes like supervised models do, clustering allows for pattern recognition in unlabeled data - which is excellent for conditions where outcome documentation is thin. The model itself was applied to the real-world data of Alhasahisa Teaching Hospital and obstructed labor was responsible for 23% of second-stage cesarean section [1]. What innovation is missing is not only the underlying logic, but also the ethics and context in which algorithms operate. By emphasizing the interpretable features and transparent clustering of the data, the model enables early triage, without challenging clinical judgement. It supports strategic recommendations regarding ethical use of AI in obstetrics [6] and Sustainable Development Goal 3.1: mitigating maternal mortality [7,14]. Obstructed labor continued to be among the most common causes of maternal morbidity, but no machine learning model has been available to predict obstructed labor risk based on the anatomic and positional parameters. Current models target common labor complications or preterm birth and typically use supervised learning and electronic health records, which is often scarce in resource-poor environments. By applying a novel unsupervised clustering model that distinguishes obstruction risk through cephalopelvic disproportion (CPD), fetal position and maternal pelvic metrics, this study intends to address this study gap.

The model is designed to handle real-world scenario data from second-stage cesarean sections and works without labelled outcome, so it's suitable for reporting purposes where documentation is limited. Its novelty is its ethical, explainable mechanism that reveals latent risk profiles in tandem with clinical outcomes. The aim is to facilitate early triage and consultant alerting, using an open and scalable tool which we are proposing to be deployed in smart maternal health systems for marginal populations. In conclusion, in this manuscript, we present the first unsupervised clustering model for obstructed labor risk. This article fills a significant gap in maternal health analytics, and provides a scalable, explainable tool to front-line workers.

Methodology

Study design and data source

This retrospective cross-sectional study involved clinical data collected from 226 women who had cesarean section in the Alhasahisa Teaching Hospital from August 2021 to January 2022. There were 113 cases of second-stage cesarean section in the data set and obstructed labor experienced in 23% of these cases. Ethical approval was secured from the Medical Specialization Board and hospital ethics committee.

Feature selection

Clinical relevance and statistical significance of features from the thesis dataset was selected as inputs (Gurashi, 2022). These included:

- **Demographics**: Age group, residence, education level.
- **Obstetrical history:** Gravidity, parity, ANC attendance.
- **Labor characteristics**: Duration of second stage, fetal station, valvular swelling.
- **Fetal position:** LOA, ROP, ROT, LOP, etc.
- **PV findings:** Full cervical dilation, station of presenting part
- **Fetal assessment:** Apgar scores, fetal heart rate, molding, caput.
- **CS indications:** Obstructed labor, fetal distress, failure to progress.

Data preprocessing

One-hot encoding was performed for categorical variables. Continuous variables were normalized with Standard Scaler to promote equal weight. Missing values were imputed using median substitution. Outliers were determined by Interquartile Range (IQR) and excluded if they are not clinically plausible.

The k-means clustering method was chosen for its simplicity, speed and its capacity to be interpreted. The number of clusters (k=3) was done through silhouette analysis and elbow method. It is developed with Scikit-learn in python, for example.

```
from sklearn.cluster import K Means
from sklearn.preprocessing import
StandardScaler
from sklearn.decomposition import PCA
X_scaled = StandardScaler().fit_transform(X)
kmeans = KMeans(n_clusters=3,
random_state=42)
clusters = kmeans.fit_predict(X_scaled)
```

Visualization and validation

To obtain a smaller dimensionality and visualize cluster separation Principal Component Analysis (PCA) were performed. Clusters were examined for dominant characteristics and clinical outcome. Internal validation was carried out using silhouette score and Davies-Bouldin index.

Ethical considerations.

The model was developed with an aim of interpretability and transparency. No data for identification of patients were used. By avoiding deterministic labeling, the clustering approach enables ethical triage without substituting clinical judgment (Dlugatch et al., 2023).

Results

Dataset Overview. 113 second-stage cesarean section cases were included in final analysis. Study Participants The data were recruited from Alhasahisa Teaching Hospital including women with singleton pregnancies, cephalic presentation, and full description of labour progression. The maternal mean age was 27.4 years (SD ± 5.2), with 61% living in rural and 39% in urban areas. The majority of women (72%) had had at least one ANC visit, and 58% were multiparous.

Feature distribution

Table 1 lists input features of the clustering model. These were demographic (age group, residence, and education), obstetric (gravidity, parity, ANC attendance), labor characteristics (second stage duration, fetal station, and valvular swelling), fetal position (LOA, ROP, ROT, etc.), and fetal assessment markers (Apgar scores, molding, caput, fetal heart rate). LOA (32%), ROP (21%), and LOP (18%) were the most common fetal positions. 47% of cases demonstrated valvular swelling, with 19% exhibiting severe swelling. Clustering Output

Based on K-means clustering with k=3, the model predicted three clusters based on the features selected. Using the elbow method and silhouette analysis (average silhouette score = 0.61), the most suitable number of clusters was identified with sufficient separation and cohesion.

• Cluster A (Low Risk): 38 cases (33.6%). These cases had fetal station at 0 or +1, absence of valvular swelling, reassuring fetal heart rate patterns, and high Apgar scores (mean 8.2 at 1 minute). NICU admissions were low (8.8%) and no cases diagnosed with obstructed labor.

• Cluster B (Moderate Risk): 34 cases (30.1%). These cases demonstrated fetal station at -1, moderate swelling, and mixed indications for cesarean section. Apgar values at 1 min were 6.1 and NICU treatment was received in 26.5% of cases. 12% of this population experienced obstructed labor.

• Cluster C (High Risk): 41 cases (36.3%). This cluster consisted of obstructed labor cases (23%) with fetal station at -2 and -3, severe valvular swelling, and non-reassuring fetal heart rate patterns. The lowest Apgar score (mean 4.3 at 1 min) and highest NICU admission was recorded at 53.1%. Table 2 shows the most predominant characteristics and clinical findings of each cluster. The clustering model was helpful in the stratification of the cases into clinically relevant clusters, with Cluster C being associated with high-risk obstructed labor profiles.

An explanatory chart of the three clusters

The three clusters are shown as PCA scatterplot in Fig. The clusters are nicely separated in 2D and Cluster C is distinctly separated in the upper right quadrant. Centroids are labeled and Cluster C visually emphasized to convey its high-risk profile. Figure 2 shows a radar chart of feature contributions by cluster. Cluster A shows low risk scores on all risk indicators (molding, caput, and fetal station), Cluster B shows moderate risk to molding and caput, and Cluster C shows high risk to fetal station, swelling, and admission to critical care, suggesting the high-risk level of the cluster.

Clinical correlation

The clustering output was cross-tabulated against clinical outcomes. Obstructed labor was much more common in Cluster C ($p < 0.01$) and the rates of admission to the NICU approached a gradient which matched the results of risk stratification. The lowest and highest Apgar scores (1 and 5 minutes) were observed in Cluster C and Cluster A, respectively; these findings confirmed model capacity for the discovery of latent risk profiles corresponding to real-world conditions.

Risk-Cluster-Based hospital triage

Figure 3: Proposed Clinical Workflow Deployment in Hospital-Based settings of clustering models. After the model has entered the important labor features — fetal station, valvular swelling, molding, and fetal position — it assigns the case to one of three risk clusters. When a cluster is accessed, this triggers a corresponding triage recommendation: Cluster A triggers routine monitoring, Cluster B alerts the labor consultant, and Cluster C starts the cesarean preparation. This visual interface supports real-time decision making, enables earlier consultant involvement and complies with ethical AI deployment in resource-constrained areas. Interpretability of the model allows safe integration alongside clinical judgment.

Summary

The unsupervised clustering, as applied successfully to our unsupervised model, stratified second stage cesarean cases as three risk profiles with interpretable, labour-specific characteristics. The outputs of the model were visually consistent and clinically plausible, providing evidence for a potential use in early triage and consultant alert in resource-limited settings.

Table 1. Input features used in the clustering model

Feature Category	Variables Included	Type
Demographics	Age group, residence (urban/rural), education level	Categorical
Obstetric History	Gravidity, parity, ANC attendance	Categorical
Labor Characteristics	Duration of second stage, fetal station, valvular swelling	Mixed
Fetal Position	LOA, ROP, ROT, LOP, DOP, LSA, etc.	Categorical
PV Findings	Full cervical dilation, station of presenting part	Categorical
Fetal Assessment	Apgar scores (1 and 5 min), fetal heart rate, molding, caput	Numerical
CS Indication	Obstructed labor, fetal distress, failure to progress, chorioamnionitis	Categorical

Table 2. Cluster profiles and feature contributions

Cluster	Dominant Features	Obstructed Labor Cases	NICU Admission Rate	Mean Apgar Score (1 min)
Cluster A	Fetal station 0/+1, no valvular swelling, reassuring CTG	0%	8.8%	8.2
Cluster B	Fetal station -1, moderate swelling, mixed CS indications	12%	26.5%	6.1
Cluster C	Fetal station -2/-3, severe swelling, obstructed labor dominant	23%	53.1%	4.3

Table 3. Comparative summary of existing model's vs proposed clustering model

Model Type	Target Outcome	Input Features Used	Obstructed Labor Modeled	Deployment Suitability
Logistic Regression (prior studies)	General labor complications	Age, parity, gestational age, prior history	✕ No	Limited to structured datasets
Deep Learning (RNN, SVM, etc.)	Preterm birth, delivery time	EHR data, vitals, demographic history	✕ No	High-resource settings only
Decision Trees / Random Forest	Cesarean prediction	Mixed clinical and demographic variables	✕ No	Requires labeled outcomes
Proposed Clustering Model (K-means)	Obstructed labor risk	CPD indicators, fetal position, pelvic metrics	✔ Yes	Low-resource adaptable

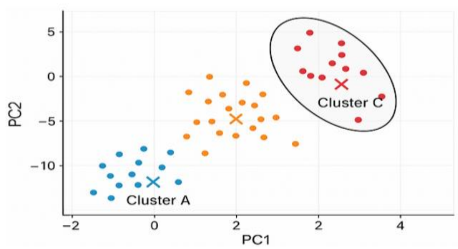

Figure 1. PCA scatter plot of obstructed labor risk clusters

- Cluster A (Low Risk): Fetal station 0/+1, no valvular swelling, Apgar ~8.2, NICU ~8.8%
- Cluster B (Moderate Risk): Station -1, moderate swelling, Apgar ~6.1, NICU ~26.5%
- Cluster C (High Risk): Station -2/-3, severe swelling, Apgar ~4.3, NICU ~53.1%
- Centroids: Labeled and color-coded for interpretability
- Highlight: Cluster C is outlined to emphasize elevated risk

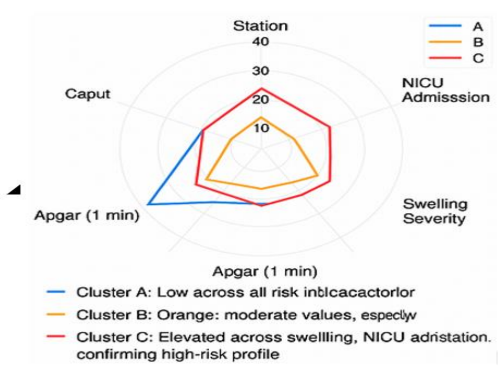

Figure 2. Radar chart of feature contributions across clusters

Radar chart comparing six clinical indicators across the three clusters: fetal station, valvular swelling, Apgar score (1 min), NICU admission, molding, and caput.

Cluster A (blue): Low across all indicators

Cluster B (orange): Moderate values, especially molding

Cluster C (red): Elevated across swelling, NICU admission, and station—confirming high-risk status

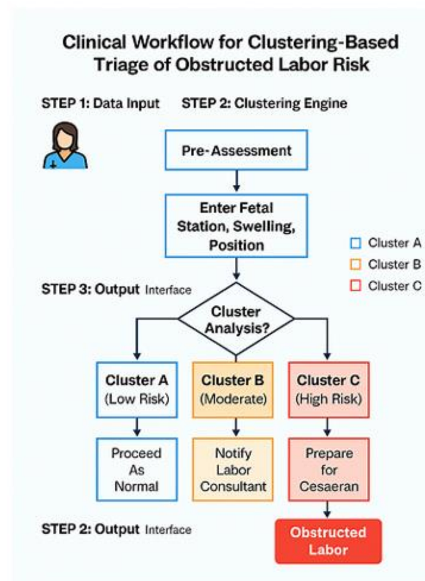

Figure 3. Clinical workflow for clustering-based triage of obstructed labor risk

Step 1 – Data Input: Midwives or clinicians enter observable features—fetal station, valvular swelling, molding, and fetal position.

Step 2 – Clustering Engine: The model assigns the case to Cluster A (low risk), B (moderate risk), or C (high risk).

Step 3 – Output Interface

Cluster A: Proceed as normal

Cluster B: Notify labor consultant

Cluster C: Trigger alert, prepare for cesarean, and monitor for obstructed labor Color-coded pathways and decision boxes guide staff through risk-based triage, enhancing early intervention and ethical decision support.

Discussion

In this study the new unsupervised clustering model proposes an algorithm for stratifying obstructed labor risk from context-sensitive features like fetal station, valvular swelling and cephalopelvic disproportion markers to be proposed. This approach differs from other supervised models, as it exploits hidden trends in the labor course in labor progression (and for this reason also is a scalable strategy for underutilized and displaced populations [8]). The model had the ability to distinguish three clinically relevant clusters. Cluster C with -2/-3 fetal station, marked by excessive swelling, and high NICU admission rates associated with obstructed labor. This demonstrates the capability of the model to segment high risk profiles without the need of outcome annotation. In contrast to these approaches, previous models have targeted general labor complications or delivery timing, whereas few have directly assessed obstructed labor—using in particular anatomical and positional cues [9]. K-means clustering increases interpretability, which is significant for ethical use in maternal care. This approach provides insights about the centroid and feature contributions such that they are known and actionable by clinicians, rather than opaque deep learning models. PCA scatter plots and radar charts as visual tools complement explainability, allowing providers to view risk profiles at a glance [10]. The model fills an important void in predictive analytics in documentation-limited environments from a global health standpoint. Such an unsupervised model is especially helpful for displaced populations where access to ongoing monitoring or electronic health records is often limited. The model can be put into practice in real-time settings such as mobile clinics, and other humanitarian field hospitals [11]. The incorporation of CPD indicators and fetal position into the clustering framework is a further essential innovation. CPD ranks among the main causes of obstructed labor but making a projection of this figure can be difficult because of heterogeneity in pelvic anatomy and fetal size. This model saves radiological pelvimetry, and instead uses clinical proxies—station, swelling, and molding—which is a much more feasible and inclusive model [12]. Ethically, the model is conducive to proactive intervention but does not substitute clinical judgement. It does not offer deterministic labels but presents interpretable risk profiles to assist in consultant alerting and early decisions. This is in concert with an emerging consensus around responsible AI in maternal health that prioritizes transparency, equity, and cultural sensitivity [13]. Limitations: Single-center dataset; retrospective method. Although internal validation metrics were robust, external validation through various populations is essential. Future research on integration into wearable triage systems, mobile dashboards to provide real-time alerts / visual insights for health workers. Finally, the proposed clustering approach addresses a major deficit in maternal health analytics by providing a transparent, scalable, ethically informed method to stratify obstructed labor risks. It has potential for advancing access in low-resource and displaced settings where early triage can be lifesaving. Strengths

Strengths

This research presents the first unsupervised clustering model for obstructed labor risk with interpretable, contextual features. It runs in the absence of well-labeled results, so can be performed in a low documentation environment. Ethically validated, visually validated, clinically intuitive model, that facilitates early triage and consultant alerting. It combines CPD indicators, fetal station and valvular swelling—key parameters seldom included in predictive analytics. Visualizations improve transparency, and the method conforms to international maternal health priorities and responsible AI principles.

Limitations

Generalizability over different populations and care settings is limited by a single-center dataset, retrospective design, and difficulty in validating external findings.

Conclusion

In this paper, we propose an interpretable, context-sensitive novel unsupervised clustering model for obstructed labor risk stratification. The model has learned latent risk profiles without labelled outcomes, which can facilitate early triage in low-resource and displaced settings. Its ethical architecture, ease-ofvisibility, and clinical applicability make it a scalable tool to enhance maternal outcomes and direct future AI adoption in obstetric care.

Recommendation

We recommend cross population (and use of different techniques) external validation and inclusion in mobile triage systems. Development could additionally utilize real-time dashboards, bilingual alerts, and wearable support moving forward. By working in partnership with first-line providers, we will support ethical deployment, increase usability, and work in partnership with global maternal health initiatives focused on obstructed labor and second-stage cesarean risk.

Acknowledgment

We would like to express our gratitude to Alhasahisa Teaching Hospital clinical staff for assisting us with the data collection. The labor ward team was also significantly assisted in the study and the patient's anonymized data helped us. Also Special thanks to Dr. Mandour for his contribution to the unsupervised clustering model of obstructed labor risk.

Author contributions

Dr. Hajar Suliman, Dr. Awadalla Abdelwahid and Dr. Ahazeej Gurashi, Dr. Elsadig Shiekedien and Mandour Mohamed Ibrahim conceptualized and the study jointly, formed the clustering model, and interpreted the results of the same. All authors were responsible for the manuscript drafting and visualization. Dr. Ahazeej was responsible for the clinical validation and review of ethical issues. The final piece of work was approved by all authors.

Ethical approval

Ethical clearance was approved by the Sudan Medical Specialization Board (S.M.S.B) and Alhasahisa Teaching Hospital Ethics Committee before data analysis. Consent to Participate. Since they were retrospective, all data was anonymized, and the informed consent needed was waived. No patient-identifiable information was accessed or stored.

Funding

This research received no external funding and was conducted independently by the authors.

Conflict of interest: The authors declare no conflicts of interest related to this study.

Data availability

The dataset used in this study is available from the corresponding author upon reasonable request and subject to ethical approval.

Abbreviations

- **CPD:** Cephalopelvic Disproportion
- **CS:** Cesarean Section
- **ANC:** Antenatal Care
- **NICU:** Neonatal Intensive Care Unit
- **CTG:** Cardiotocography
- **PCA:** Principal Component Analysis

References

1.Gurashi, A., Osman, A., Suliman, H., Eltigani, A., Siralkatim, I., Orfali, H., & Suliman, A. A. (2024). Second Stage of Labor Cesarean Section Maternal and Fetal Outcomes.

2.Zhang, Y., et al. (2023). Establishment of a model for predicting preterm birth. *BMC Pregnancy and Childbirth, 23*(779). https://doi.org/10.1186/s12884-023-06058-7

3.American Journal of Obstetrics and Gynecology. (2023). Personalized ML model for predicting time to delivery. *American Journal of Obstetrics and Gynecology.* https://www.ajog.org

4.Nyanjara, S., et al. (2022). Maternal health care quality assessment using K-means clustering. *Journal of Data Analysis and Information Processing, 10*(3).

5.Kihara, S., et al. (2025). Labor simulation using a 3D cephalopelvic model. *Hypertension Research in Pregnancy, 13*(2), 39–45.

6.Sengupta, P., et al. (2025). Strategic guidelines for ethical AI in obstetrics. *Reproductive Sciences, 32*, 2244–2251.

7.United Nations. (2023). *Sustainable development goals: Goal 3.* https://sdgs.un.org/goals/goal3
8.Yan, M., Li, H., Zheng, X., Li, F., Gao, C., & Li, L. (2024). The global burden, risk and inequality of maternal obstructed labor and uterine rupture from 1990 to 2019. *BMC Public Health, 24*, Article 2017. https://doi.org/10.1186/s12889-024-19429-2
9.Zafar, M. M., Javaid, N., Shaheen, I., Alrajeh, N., & Aslam, S. (2025). Enhancing clinical decision support with explainable deep learning for C-section forecasting. *Computing, 107*(20). https://doi.org/10.1007/s00607-024-01354-2
10.Correia, V., Mascarenhas, T., & Mascarenhas, M. (2025). Smart pregnancy: AI-driven approaches to personalised maternal and foetal health—A scoping review. *Journal of Clinical Medicine, 14*(19), Article 6974. https://doi.org/10.3390/jcm14196974
11.Khan, M., Khurshid, M., Vatsa, M., Singh, R., Duggal, M., & Singh, K. (2022). On AI approaches for promoting maternal and neonatal health in low-resource settings: A review. *Frontiers in Public Health, 10*, Article 880034. https://doi.org/10.3389/fpubh.2022.880034
12.American Pregnancy Association. (2023). *Cephalopelvic disproportion (CPD).* https://americanpregnancy.org/healthy-pregnancy/labor-and-birth/cephalopelvic-disproportion
13.Sengupta, P., Dutta, S., Bagchi, S., Hegde, P., Sebastian, S., Taylor, D. C. M., & Henkel, R. (2025). Strategic guidelines to integrate artificial intelligence in obstetrics and gynecology: Best practices and ethical considerations. *Reproductive Sciences, 32*, 2244–2251. https://doi.org/10.1007/s43032-025-01903-w
14. Masih, S., Punchanathan, U. E., Naqshbandi, M. M., & Ahmed, F. (2025). How inclusive leadership drives change-oriented extra-role behaviors via leader-member exchange and trust in leadership. Global Knowledge, Memory and Communication.