© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

AI-Assisted Cardiotocography Interpretation for Early Detection of Fetal Distress: Implications for Obstetric Care in Saudi Arabia

AMINA MAJID

Abstract- Cardiotocography (CTG) remains one of the most frequently used methods for assessing fetal wellbeing during labor, particularly in pregnancies requiring continuous intrapartum surveillance. Despite its widespread use, conventional visual CTG interpretation is limited by interobserver and interobserver variability, inconsistent application of clinical guidelines, signal artefacts, and the relatively low positive predictive value of abnormal patterns for actual fetal hypoxia. Artificial intelligence (AI), machine learning (ML), and deep-learning techniques are increasingly being investigated as tools for improving the consistency and predictive value of CTG interpretation. This review evaluates contemporary evidence on AI-assisted CTG analysis for the early recognition of fetal compromise and examines its potential application within Saudi Arabian maternity services. A structured literature review was undertaken using evidence published primarily between 2020 and 2025. Current research demonstrates that machine-learning and deep-learning models can identify CTG abnormalities, predict fetal acidemia, and provide near-real-time assessment of fetal-heart-rate responses. Multicenter deep-learning studies have demonstrated promising discrimination for severe neonatal acidemia, while AI-based models may reduce subjectivity by continuously evaluating fetal-heart-rate and uterine-contraction patterns. However, clinical implementation remains limited by inconsistent outcome definitions, retrospective datasets, signal-quality problems, inadequate external validation, algorithmic bias, class imbalance, and limited explainability. AI should therefore augment rather than replace obstetric judgment. For Saudi Arabia, integrating AI-assisted CTG with electronic maternal records, central fetal-monitoring platforms, and structured escalation protocols could support high-volume maternity hospitals and align with the digital-health objectives of Saudi Vision 2030. A clinician-led human-AI model is proposed to strengthen fetal surveillance while maintaining patient safety, transparency, and professional accountability.

Keywords: Cardiotocography; Artificial Intelligence; Fetal Hypoxia; Fetal Distress; Fetal Monitoring; Machine Learning; Deep Learning; Clinical Decision Support; Saudi Arabia; Vision 2030.

I. INTRODUCTION

Intrapartum fetal monitoring is intended to identify fetuses that are becoming unable to tolerate the physiological demands of labor. Uterine contractions intermittently reduce uteroplacental blood flow and oxygen delivery. Most healthy fetuses compensate adequately; however, fetuses with reduced placental or physiological reserve may progressively develop hypoxemia, metabolic acidemia, neurological injury, or other adverse neonatal outcomes. Cardiotocography records fetal heart rate together with uterine activity and is widely used to identify changes that may indicate evolving fetal compromise. Clinical CTG interpretation typically considers baseline fetal heart rate, variability, accelerations, decelerations, contraction frequency, and changes in these features over time. Current clinical guidance stresses that CTG should not be interpreted in isolation. Maternal observations, gestational age, fetal risk factors, medications, meconium, labor progression, oxytocin administration, infection, and other clinical variables must also be considered. NICE guidance specifically recommends assessing the whole clinical picture and examining temporal changes rather than responding to an isolated CTG feature.

A major limitation of conventional CTG interpretation is subjectivity. Different clinicians may classify the same tracing differently, while the same observer may reach different conclusions when reassessing a trace. A systematic review of international CTG guidelines also identified substantial differences in how major guidance

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

systems classify variability, decelerations, accelerations, and other fetal-heart-rate features.

This variability creates two patient-safety problems. Failure to recognize true fetal deterioration can delay intervention and increase the risk of metabolic acidemia, hypoxic-ischemic injury, neonatal intensive-care admission, and stillbirth. Conversely, overinterpretation of abnormal CTG patterns can contribute to unnecessary operative vaginal delivery or caesarean birth.

Artificial intelligence offers a potential solution by enabling continuous and standardized analysis of large quantities of CTG data. Recent AI systems can extract fetal-heart-rate patterns automatically, quantify changes over time, and estimate the probability of fetal compromise. Deep-learning methods can additionally learn complex relationships directly from raw CTG signals without depending exclusively on predefined clinical features.

Saudi Arabia represents an important setting for the development of such technologies. Vision 2030 and the Health Sector Transformation Program emphasize digital healthcare, integrated systems, preventive care, advanced analytics, and artificial intelligence. The 2024 Health Sector Transformation Report also highlights the establishment of an Operations and Artificial Intelligence Center to strengthen healthcare through AI and advanced data analytics.

This review therefore examines the current evidence for AI-assisted CTG interpretation, its ability to identify fetal compromise, its limitations, and its potential integration into obstetric practice in Saudi Arabia.

II. LITERATURE REVIEW

2.1 Cardiotocography and Fetal Compromise

CTG simultaneously records fetal heart rate and uterine contractions. Fetal-heart-rate behavior reflects interactions between the autonomic nervous system, cardiovascular regulation, fetal oxygenation, uteroplacental circulation, medications, sleep cycles, and the mechanical effects of labor.

A reassuring fetal-heart-rate pattern generally indicates intact autonomic regulation and adequate physiological reserve. Abnormal patterns can include persistent tachycardia or bradycardia, reduced variability, repetitive decelerations, prolonged deceleration, or progressive deterioration in the relationship between fetal-heart-rate recovery and uterine contractions.

Although the phrase fetal distress remains widely used, “suspected fetal compromise” or “suspected fetal hypoxia” is more precise because abnormal CTG does not automatically prove oxygen deprivation. Confirmation of clinically significant compromise frequently relies retrospectively on umbilical-cord arterial pH, base excess, neonatal condition, Apgar score, or neurological outcomes.

This distinction is important for artificial intelligence because an algorithm learns according to the outcome definition provided by researchers.

2.2 Limitations of Conventional CTG Interpretation

Visual interpretation generally follows structured clinical guidance. Three-tier approaches commonly classify traces as normal, suspicious/intermediate, or pathological/abnormal.

However, guideline-based interpretation remains challenging.
Ben M’Barek, Jauvion and Ceccaldi (2023) reported that visual CTG assessment is associated with substantial interobserver variability and noted that computerized methods have been investigated for decades in response to this problem. Modern machine-learning techniques have renewed interest because they are capable of processing much larger clinical datasets and more complex temporal relationships.

A systematic review of clinical practice guidelines similarly identified meaningful differences between major international CTG frameworks, including disagreement about the significance of individual fetal-heart-rate characteristics.

Human factors further influence interpretation. A clinician managing several laboring patients

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

simultaneously may face fatigue, interruptions, emergency workload, incomplete information, and pressure to make rapid decisions. Junior and senior clinicians may also interpret identical patterns differently.

These limitations make CTG particularly suitable for decision-support technology, provided the technology does not remove the broader clinical context.

2.3 Computerized CTG Analysis

Computerized CTG analysis predates modern artificial intelligence. Earlier systems usually extracted predefined fetal-heart-rate characteristics and applied fixed clinical rules.

Modern systems employ statistical learning and pattern-recognition techniques.

Typical CTG variables include:

- Fetal-heart-rate baseline
- Short-term variability
- Long-term variability
- Accelerations
- Early decelerations
- Variable decelerations
- Late decelerations
- Prolonged decelerations
- Uterine-contraction frequency
- Duration of abnormal patterns
- Frequency-domain features
- Temporal relationship between contractions and fetal responses

The machine-learning model then determines relationships between these variables and an outcome such as normal fetal status, neonatal acidemia, low Apgar score, or another indicator of compromise.

Mendis et al. (2023) reviewed computerized CTG technologies and described both conventional machine-learning and deep-learning approaches. They emphasized that automated interpretation could potentially improve fetal-compromise detection but identified data quality, preprocessing, dataset limitations, and translation into clinical practice as major barriers.

2.4 Machine Learning in CTG Interpretation

Machine-learning algorithms frequently used for CTG include:

- Logistic regression
- Decision trees
- Random forests
- Support-vector machines
- Gradient boosting
- Artificial neural networks
- Ensemble classifiers

Francis et al. (2024) reviewed 36 studies applying signal processing and machine learning to CTG. Their review concluded that ML showed significant potential for detecting fetal hypoxia but identified substantial methodological weaknesses. Many studies repeatedly used the same open-access database, and fetal metabolic acidemia was defined using inconsistent umbilical-pH thresholds.

This is important because a model that achieves high accuracy on a small public dataset may perform differently in actual maternity units.

Clinical populations vary according to:

- Maternal characteristics
- Gestational age
- Obstetric complications
- Monitoring equipment
- Clinical protocols
- Medications
- Labor-management strategies

External validation is consequently essential before implementation.

2.5 Deep Learning and Fetal Acidemia Prediction

Deep-learning models reduce dependence on manual feature selection.

Convolutional neural networks can automatically identify complex patterns within fetal-heart-rate signals. Recurrent neural networks and long short-term memory models are designed for sequential information and can assess how CTG characteristics evolve over time.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

Recent research demonstrates increasingly strong performance.

McCoy et al. developed and externally validated deep-learning models using a large electronic fetal-monitoring dataset. More than 124,000 monitoring files were initially available, and 10,182 traces appropriately linked with umbilical-cord gas measurements formed the primary analytical dataset.
The best-performing model achieved an area under the receiver-operating-characteristic curve of approximately 0.85 for severe acidemia defined as umbilical arterial pH below 7.05. When severe acidemia was combined with significant base-excess abnormality, performance reached approximately 0.89.

These results are clinically important because they demonstrate that fetal-heart-rate data contain patterns capable of predicting biochemical compromise beyond conventional visual categorization.

Nevertheless, even strong AUC values do not guarantee clinical benefit. Whether AI actually reduces neurological injury or unnecessary caesarean delivery requires prospective assessment.

2.6 DeepCTG Models

Another significant development is the DeepCTG research program.
DeepCTG 1.0 used clinically interpretable fetal-heart-rate characteristics to estimate fetal hypoxia. Instead of producing an entirely opaque result, the approach incorporated recognizable CTG parameters, demonstrating how computational prediction and clinical interpretability can coexist.

DeepCTG 2.0 subsequently used a multicenter dataset involving 27,662 cases from five hospitals. The convolutional neural-network model was evaluated across different centers rather than only within one dataset.

For severe neonatal acidemia, cross-center AUC values ranged from approximately 0.74 to 0.83. The investigators concluded that additional clinical factors could further strengthen prediction because maternal and fetal risk factors are not fully represented within CTG signals alone.
This finding strongly supports multimodal decision support rather than CTG-only AI.

2.7 Near-Real-Time Artificial Intelligence

An important requirement for AI-assisted fetal monitoring is real-time functionality.
Melaet et al. (2024) developed an AI system that assessed whether the fetal heart rate responded normally to uterine activity. Instead of simply categorizing an entire trace retrospectively, the model learned how the individual fetus had previously responded and compared expected responses with emerging behaviour.

The system achieved an AUC of approximately 0.96 when distinguishing normal from pathological events compared with majority expert annotations.

This individualized approach is particularly interesting because fetal reserve differs between pregnancies. An algorithm capable of learning a fetus's previous response to contractions may identify deterioration earlier than fixed thresholds.
However, agreement with clinicians does not necessarily prove improvement in neonatal outcomes. Clinical trials remain necessary.

III. RESEARCH METHODOLOGY

3.1 Research Design

This study adopted a structured narrative-review methodology with systematic literature-search elements.

A narrative synthesis was considered appropriate because AI-assisted CTG research contains substantial methodological heterogeneity. Studies differ in datasets, CTG acquisition, signal preprocessing, algorithms, fetal-compromise definitions, outcome measures, and validation methods.

3.2 Data Sources

Relevant evidence was identified using:

- PubMed
- Scopus

- Web of Science
- ScienceDirect
- SpringerLink
- IEEE-related research
- Clinical guideline databases
- Saudi Vision 2030 publications

The principal literature period was 2020–2025, while earlier established clinical guidance was considered when necessary for explaining standard fetal-monitoring principles.

3.3 Search Strategy

Keywords included:

"cardiotocography AND artificial intelligence"
"CTG AND machine learning"
"fetal hypoxia AND artificial intelligence"
"fetal acidemia AND deep learning"
"computerized cardiotocography"
"CTG interpretation variability"
"fetal monitoring AND clinical decision support"
"artificial intelligence AND obstetrics"
"Saudi Arabia AND digital healthcare"

3.4 Inclusion Criteria

Eligible evidence included:

1. Peer-reviewed clinical research.
2. Systematic or scoping reviews.
3. Machine-learning CTG studies.
4. Deep-learning fetal-monitoring studies.
5. Studies predicting hypoxia or neonatal acidemia.
6. Computerized CTG analyses.
7. Relevant professional clinical guidance.
8. Saudi digital-health policy documents.

3.5 Exclusion Criteria

The review excluded:

- Animal-only studies.
- Studies unrelated to fetal monitoring.
- Purely technical algorithms without meaningful obstetric application.
- Duplicate publications.
- Non-English publications.
- Studies with insufficient methodological information.

3.6 Data Extraction

Information was extracted regarding:

- Study population
- Dataset size
- CTG variables
- Signal preprocessing
- AI architecture
- Reference outcome
- Validation approach
- Accuracy measures
- Clinical interpretation
- Limitations

3.7 Data Analysis

Evidence was organized into the following themes:

1. Conventional CTG limitations.
2. Machine-learning interpretation.
3. Deep-learning prediction.
4. Fetal acidemia detection.
5. Explainable AI.
6. Human-AI collaboration.
7. Saudi healthcare implementation.

IV. RESULTS AND DISCUSSION

4.1 Overall Evidence

The literature demonstrates substantial progress in computerized fetal monitoring.

Aeberhard et al. reviewed 40 studies evaluating AI or machine-learning systems for CTG interpretation and concluded that numerous approaches were promising. However, none had achieved widespread clinical acceptance. Larger, better curated and more consistently labelled datasets were identified as essential for development of clinically reliable systems.

Francis et al. reached a similar conclusion. Machine-learning methods can detect fetal hypoxia, but inconsistent benchmarks and repeated use of small datasets limit reproducibility and generalizability.

Therefore, the current evidence supports AI-assisted rather than AI-controlled fetal monitoring.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

Table 1. Conventional versus AI-Assisted CTG Interpretation

Parameter	Conventional CTG	AI-Assisted CTG
Interpretation	Visual clinician assessment	Algorithm + clinician
Consistency	Observer dependent	Potentially standardized
Continuous analysis	Workload dependent	Automated
Complex pattern detection	Experience dependent	Data driven
Fatigue effect	Possible	None for algorithm
Clinical context	Strong	Limited unless integrated
Explainability	Generally high	Model dependent
Real-time trend analysis	Variable	Potentially continuous
Final decision	Clinician	Clinician should retain authority

4.2 Potential for Earlier Detection

The strongest potential advantage of AI is continuous trend analysis.

Human clinicians generally review CTG over defined periods and recognize deterioration based on visual assessment. AI can simultaneously quantify:

- Baseline changes
- Variability trends
- Deceleration frequency
- Deceleration duration
- Recovery patterns
- Contraction frequency
- Cumulative abnormality
- Changes over prolonged labour

Subtle deterioration may therefore be recognized before a trace becomes obviously pathological.

Near-real-time work by Melaet et al. supports this concept, while the large deep-learning work of McCoy et al. demonstrates that CTG contains clinically meaningful information capable of predicting cord-blood acidemia.

4.3 Clinical Performance Must Be Interpreted Carefully

Accuracy alone is insufficient.

Severe fetal hypoxia is relatively uncommon, creating class imbalance. A model can achieve high overall accuracy by correctly identifying normal fetuses while still missing an unacceptable proportion of pathological cases.

Clinically relevant metrics therefore include:

- Sensitivity
- Specificity
- Positive predictive value
- Negative predictive value
- AUC
- Calibration
- False-negative rate
- False-positive rate

The appropriate threshold also depends on the intended clinical use.

A screening system should prioritize sensitivity because missed compromise can be catastrophic. However, excessive sensitivity may produce large numbers of false alarms and potentially increase unnecessary operative delivery.

The ideal AI system must therefore balance early detection with avoidance of over-intervention.

4.4 Explainable AI

One of the greatest barriers to AI adoption is the black-box problem.

An alert that simply states "high fetal risk" provides limited clinical value.

A preferable system would explain:

"Risk increased because recurrent late decelerations developed, baseline increased progressively, and variability declined during the previous thirty minutes."

Explainability provides several benefits:

- Supports clinical trust.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

- Allows rapid verification.
- Facilitates teaching.
- Improves audit.
- Helps identify erroneous alerts.
- Supports professional accountability.

Ben M'Barek and colleagues have emphasized that computerized CTG systems should provide interpretable information alongside predicted fetal-hypoxia risk.

4.5 Signal Quality and Artefacts
Reliable AI depends upon reliable input.

CTG recordings frequently contain:
- Missing data
- Transducer displacement
- Maternal-heart-rate contamination
- Poor signal acquisition
- Sudden nonphysiological changes
- Equipment-related artefacts

Mendis et al. describe preprocessing as an essential stage of automated CTG analysis. Signal cleaning, artefact detection, interpolation, normalization, and appropriate selection of CTG segments directly affect model performance.

An effective hospital system should therefore perform signal-quality assessment before fetal-risk assessment.

V. APPLICATION TO OBSTETRIC CARE IN SAUDI ARABIA

Saudi maternity hospitals could benefit from AI-assisted CTG in several ways.
Large maternity facilities frequently monitor numerous labouring women simultaneously. A centralized AI-supported system could continuously evaluate all active CTGs and prioritize traces requiring review.

Figure 1. AI-Assisted CTG Clinical Workflow

CTG Signal Acquisition
↓
Automated Signal-Quality Assessment
↓
Fetal Heart Rate + Uterine Activity Analysis
↓
AI Pattern Recognition
↓
Dynamic Fetal-Risk Classification
↓
Low Risk → Routine Monitoring
Intermediate Risk → Obstetric Review
High Risk → Immediate Senior Review
↓
Maternal + Fetal + Labour Assessment
↓
Clinical Management
↓
Maternal and Neonatal Outcome

The system should never automatically order emergency delivery.

It should trigger professional review.

5.1 Integration with Electronic Health Records
CTG-only prediction cannot fully represent clinical risk.

The same CTG may have different significance in a healthy term fetus and a fetus with severe growth restriction or placental insufficiency.

Future Saudi systems could integrate:
- CTG
- Gestational age
- Maternal vital signs
- Previous obstetric history
- Fetal growth
- Meconium
- Oxytocin exposure
- Maternal fever
- Labour progress
- Medications
- Pregnancy complications

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

This would create a more clinically realistic decision-support platform.

Figure 2. Human-AI Collaborative Fetal Monitoring Model

CTG + Maternal Data + Obstetric History + Labour Progress + Fetal Risk Factors
↓
AI Clinical Decision-Support Engine
↓
Transparent Risk Alert
↓
Obstetrician / Senior Midwife Review
↓
Multidisciplinary Clinical Decision
↓
Continue Monitoring / Correct Reversible Cause / Further Assessment / Expedite Birth
↓
Neonatal Outcome

5.2 Saudi Vision 2030 Alignment
The proposed model fits Saudi healthcare transformation priorities.

Vision 2030 describes a movement from fragmented and reactive healthcare toward integrated, accessible, preventive, and digitally enabled services.

The Health Sector Transformation Report 2024 specifically identifies artificial intelligence and advanced analytics as areas of development within Saudi healthcare.

AI-assisted CTG could support these objectives through:
- Standardized maternity monitoring.
- Digital clinical decision support.
- Improved patient safety.
- Earlier escalation.
- Centralized monitoring.
- Remote specialist support.
- Structured clinical audit.
- Development of national obstetric datasets.

Table 2. AI Approaches and Applications in CTG

AI Technique	Potential CTG Application	Strength	Major Limitation
Logistic Regression	Acidemia prediction	Highly interpretable	Limited complex modelling
Random Forest	Fetal-status classification	Handles nonlinear variables	Overfitting risk
Support Vector Machine	Normal/pathological classification	Effective with engineered data	Requires careful preprocessing
Gradient Boosting	Risk prediction	Strong tabular performance	Calibration required
CNN	Raw CTG analysis	Automatic feature learning	Requires large datasets
LSTM	Temporal FHR analysis	Captures sequential changes	Computational complexity
Transformer	Long-range CTG patterns	Powerful temporal modelling	Data intensive
Hybrid AI + Rules	Clinical decision support	Easier clinical integration	Depends on rule quality

VI. CLINICAL RISKS AND IMPLEMENTATION CHALLENGES

AI introduces new risks as well as benefits.

6.1 Automation Bias
Clinicians may assume the computer is correct even when clinical findings suggest otherwise.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722768

Clinical protocols should explicitly state that AI recommendations are advisory.

6.2 False Alerts
Excessive alerts could increase:
- Unnecessary intervention.
- Staff workload.
- Alarm fatigue.
- Operative delivery.

Algorithms therefore require appropriate calibration.

6.3 Algorithmic Bias
Models trained in one population may perform differently in another.

Saudi implementation should therefore include validation involving local maternal populations rather than relying exclusively on European or North American datasets.

6.4 Data Privacy
CTG recordings linked to maternal clinical information represent sensitive healthcare data.

Systems require:
- Secure storage.
- Encryption.
- Controlled access.
- De-identification for research.
- Audit trails.
- Compliance with Saudi regulations.

6.5 Clinical Liability
Responsibility must remain clearly defined.
If AI provides an incorrect recommendation, healthcare institutions need transparent protocols determining how alerts should be reviewed, documented, and overridden.

VII. RECOMMENDATIONS

Saudi tertiary maternity hospitals should initially implement AI-assisted CTG through controlled pilot programs rather than immediate nationwide deployment.

First, every proposed algorithm should undergo independent validation using Saudi obstetric populations.

Second, AI should function as a second reader and continuous surveillance tool, while obstetricians retain final clinical authority.

Third, fetal-risk alerts should be explainable. Clinicians should be able to see which CTG characteristics produced the warning.

Fourth, AI should eventually integrate electronic maternal information rather than analyze CTG alone.

Fifth, hospitals should define escalation pathways:

AI Warning → Responsible Clinician → Senior Obstetric Review → Defined Clinical Action

Sixth, maternity teams should receive structured training in both CTG interpretation and AI-supported fetal monitoring.

Seventh, routine quality audits should compare:
- AI prediction.
- Human interpretation.
- Cord arterial pH.
- Apgar scores.
- NICU admission.
- Hypoxic-ischaemic encephalopathy.
- Operative delivery.
- Missed fetal compromise.
- False alerts.

Finally, Saudi research centers should establish multicenter fetal-monitoring databases containing high-quality CTG signals and standardized neonatal outcomes.

VIII. FUTURE DIRECTIONS

Future research should move beyond simple classification of normal, suspicious, and pathological traces.

The next generation of systems should use dynamic risk prediction, estimating how the probability of fetal compromise changes minute by minute.

Multimodal models could incorporate:
CTG + Maternal Vital Signs + Laboratory Results + Fetal Growth + Medications + Oxytocin + Labour Progress + Clinical Risk Factors

Explainable AI should become a central development requirement.

Federated learning could also allow several Saudi hospitals to contribute to algorithm development without transferring identifiable patient-level data into one centralized research database.

Another promising direction is personalized fetal modelling. Instead of comparing every fetus with only population averages, algorithms could learn the individual fetus's previous response to contractions and identify when that response begins to deteriorate.
Most importantly, research must determine whether AI improves real clinical outcomes.

Future prospective trials should assess whether AI-assisted CTG reduces:

- Severe neonatal acidemia.
- Hypoxic-ischaemic encephalopathy.
- Delayed intervention.
- Avoidable emergency delivery.
- Unnecessary operative birth.
- Neonatal intensive-care admission.

Technical accuracy should no longer be considered sufficient evidence of success.

IX. CONCLUSION

Artificial intelligence has considerable potential to strengthen cardiotocography interpretation by providing standardized, continuous, and data-driven assessment of fetal-heart-rate and uterine-contraction patterns.

Recent machine-learning and deep-learning studies demonstrate promising ability to detect abnormal CTG patterns and predict neonatal acidemia. Large-scale and multicenter studies represent important improvements over earlier research based on small public datasets. Nevertheless, major barriers remain, including outcome-definition inconsistency, signal artefacts, class imbalance, algorithmic bias, limited explainability, insufficient external validation, and uncertain impact on clinical outcomes.

AI should therefore not replace the obstetrician.
The most appropriate model is human-AI collaboration, where computational systems provide continuous surveillance and transparent risk warnings while clinicians interpret those warnings using maternal condition, fetal characteristics, labour progression, examination findings, and professional judgment.

Saudi Arabia is particularly well positioned to evaluate this approach because its healthcare transformation strategy increasingly emphasizes artificial intelligence, integrated digital systems, advanced analytics, healthcare quality, and patient safety.

Through locally validated algorithms, secure digital infrastructure, multidisciplinary training, standardized escalation protocols, and prospective clinical evaluation, AI-assisted cardiotocography could become an important component of modern Saudi maternity care and contribute to safer maternal and neonatal outcomes under Saudi Vision 2030.

REFERENCES

[1] Aeberhard, J. L., Radan, A. P., Delgado-Gonzalo, R., Strahm, K. M., Sigurthorsdottir, H. B., Schneider, S., & Surbek, D. (2023). Artificial intelligence and machine learning in cardiotocography: A scoping review. European Journal of Obstetrics & Gynecology and Reproductive Biology, 281, 54–62.

[2] Ben M'Barek, I., Jauvion, G., & Ceccaldi, P. F. (2023). Computerized cardiotocography analysis during labor: A state-of-the-art review. Acta Obstetricia et Gynecologica Scandinavica, 102(2), 130–137.

[3] Francis, F., Luz, S., Wu, H., Stock, S. J., & Townsend, R. (2024). Machine learning on

cardiotocography data to classify fetal outcomes: A scoping review. Computers in Biology and Medicine, 172, 108220.

[4] Mendis, L., Palaniswami, M., Brownfoot, F., & Keenan, E. (2023). Computerised cardiotocography analysis for the automated detection of fetal compromise during labour: A review. Bioengineering, 10(9), 1007.

[5] McCoy, J. A., Levine, L. D., Wan, G., Chivers, C., Teel, J., & La Cava, W. G. (2025). Intrapartum electronic fetal heart rate monitoring to predict acidemia at birth with the use of deep learning. American Journal of Obstetrics and Gynecology, 232(1), 116.e1–116.e9.

[6] Ben M'Barek, I., Jauvion, G., Merrer, J., Koskas, M., Sibony, O., Ceccaldi, P. F., Le Pennec, E., & Stirnemann, J. (2025). DeepCTG 2.0: Development and validation of a deep learning model to detect neonatal acidemia from cardiotocography during labor. Computers in Biology and Medicine, 184, 109448.

[7] Ben M'Barek, I., Jauvion, G., Vitrou, J., Holmström, E., Koskas, M., & Ceccaldi, P. F. (2023). DeepCTG 1.0: An interpretable model to detect fetal hypoxia from cardiotocography data during labor and delivery. Frontiers in Pediatrics, 11, 1190441.

[8] Melaet, R., de Vries, I. R., Kok, R. D., Oei, S. G., Huijben, I. A. M., van Sloun, R. J. G., van Laar, J. O. E. H., & Vullings, R. (2024). Artificial intelligence based cardiotocogram assessment during labor. European Journal of Obstetrics & Gynecology and Reproductive Biology, 295, 75–85.

[9] Brown, J., et al. (2023). Clinical practice guidelines for intrapartum cardiotocography interpretation: A systematic review. Australian and New Zealand Journal of Obstetrics and Gynaecology.

[10] Zullo, F., Di Mascio, D., Raghuraman, N., Wagner, S., Brunelli, R., Giancotti, A., et al. (2023). Three-tiered fetal heart rate interpretation system and adverse neonatal and maternal outcomes: A systematic review and meta-analysis. American Journal of Obstetrics and Gynecology, 229(4), 377–387.

[11] NICE. (2022). Fetal Monitoring in Labour: NICE Guideline NG229. National Institute for Health and Care Excellence.

[12] Tsiligkeridou, S., Bolou, A., Xanthos, T., & Gourounti, K. (2023). Perinatal and neonatal outcomes using cardiotocography versus STAN and cardiotocography: A systematic review. Maedica, 18(4), 684–694.

[13] Pardasani, R., Vitullo, R., Harris, S., Yapici, H., & Beard, J. (2025). Development of a novel artificial intelligence algorithm for interpreting fetal heart rate and uterine activity data in cardiotocography.

[14] Saudi Vision 2030. (2024). Health Sector Transformation Report 2024. Kingdom of Saudi Arabia.

[15] Saudi Vision 2030. (2024). Vision 2030 Annual Report 2024. Kingdom of Saudi Arabia.

[16] Francis, F., Luz, S., Wu, H., Stock, S. J., & Townsend, R. (2024). Machine-learning approaches for fetal hypoxia detection and the challenges of clinical implementation. Computers in Biology and Medicine.

[17] Ben M'Barek, I., Jauvion, G., & Ceccaldi, P. F. (2023). Machine learning and deep learning applications in computerized intrapartum cardiotocography. Acta Obstetricia et Gynecologica Scandinavica.

[18] Mendis, L., Palaniswami, M., Brownfoot, F., & Keenan, E. (2023). Artificial intelligence approaches for automated fetal-compromise detection. Bioengineering.

[19] McCoy, J. A., et al. (2025). Deep-learning prediction of fetal acidemia using intrapartum electronic fetal monitoring. American Journal of Obstetrics and Gynecology.

[20] Aeberhard, J. L., et al. (2023). Artificial-intelligence-supported interpretation of fetal heart-rate monitoring. European Journal of Obstetrics & Gynecology and Reproductive Biology.

[21] Melaet, R., et al. (2024). Near-real-time AI assessment of fetal-heart-rate response during

labour. European Journal of Obstetrics & Gynecology and Reproductive Biology.

[22] National Institute for Health and Care Excellence. (2023). Intrapartum Care: NICE Guideline NG235. NICE.

[23] Saudi Health Sector Transformation Program. (2024). Artificial Intelligence and Advanced Analytics in Healthcare Transformation. Kingdom of Saudi Arabia.

[24] Saudi Vision 2030. (2024). Digital Healthcare, Integrated Services and Preventive Health Transformation. Kingdom of Saudi Arabia.

[25] Ministry of Health, Kingdom of Saudi Arabia. (2024). Digital Health and Artificial Intelligence Initiatives in the Saudi Healthcare System.