

(Team 1) ePartogram-GPT: Reinventing the ePartogram through Language Modeling

Ayeeshi Poosarla, Anna Vargas, Samhith Bhrugubanda, Sidharth Raghavan

Motivation:

The partogram is a crucial decision-support tool used during childbirth to help Skilled Birth Attendants (SBAs) monitor labor progression and identify moments for early interventions. The partogram tracks key labor parameters such as cervical dilation, fetal heart rate, contractions per 10 minutes, and other maternal and fetal vital signs. This tool helps SBAs detect deviations from normal labor progression and alerts them of warning signs such as prolonged labor or fetal distress. These features are critical for ensuring maternal and neonatal safety.

Despite these new tools, improving global maternal health remains a significant challenge. While the maternal mortality ratio declined by 34% from 2000 to 2020 [1], approximately 800 women die daily from preventable pregnancy and childbirth causes [2]. Unfortunately, 95% of these deaths occur in low and middle-income countries (LMICs) [3, 4]. This disparity shines a light on the inequality in reproductive healthcare access globally.

Addressing this need aligns with Sustainable Development Goal (SDG) 3, which aims to reduce maternal and neonatal mortality, and SDG 9, which pushes for innovation in healthcare infrastructure. The current challenges in LMIC's labor care systems are compounded by the pressing need for effective and scalable solutions to monitor labor progression. SBAs are burdened with heavy patient loads, which often lead to complications being detected too late, resulting in a high occurrence of Cesarean sections. To mitigate these factors, Skilled Birth Attendants (SBAs) in LMICs need a tool that **provides automated decision support** and **streamlined data input** to improve labor management.

Approach:

To develop the ePartogram-GPT, we relied on the WHO Labor Care Guide to create a system that addresses the needs of the SBAs. Our system assesses patient conditions, records observations, checks clinical thresholds, plans appropriate interventions, and suggests preliminary diagnoses based on SBA's written or spoken query.

ePartogram-GPT utilizes the Zephyr-7b-beta model to process SBA inputs due to its ability to handle complex medical queries, provide context-aware responses, and be stored on mobile devices. While more complex models such as GPT-4 and Llama have been validated commercially, they cannot be stored on mobile devices and require a connection to the internet.

To optimize ePartogram-GPT's responses, prompt engineering and Retrieval-Augmented Generation (RAG) was used to tailor outputs to specific patient summaries. Skilled birth attendant queries are processed through natural language processing and a digital WHO Labor

Care Guide is automatically filled. Alerts and follow-up questions are reciprocated by the ePartogram through thresholds and an interpretation of the data.

Aim 1: The development of a language model that can interpret clinical cases and provide decision-support

ePartogram-GPT was adapted from the Zephyr-7b-beta model, and further fine-tuned through Retrieval Augmented Generation (RAG) and prompt engineering. The latest maternal health guidelines and training materials were acquired from the World Health Organization, UNICEF, USAID, and Jhpiego. Within this database included the WHO Labor Care Guide, WHO Managing Complications in Pregnancy and Childbirth: A guide for midwives and doctors, and clinical rules for the eZazi mobile app (see Supplementary Table 1 for full list of resources). Clinical rules were further simplified and rewritten for clarity by the co-authors. Retrieval Augmented Generation was used to split the database into 1000 character chunks, store the chunks in a Chroma-vector database, and to create a retriever for finding relevant document chunks based on the search query. By manipulating the RAG database, and testing scenarios, it was found that certain training documents, specifically those that over-described empathetic care or provided instructions on how to use various maternal tools, made the AI assistant worse at diagnosis prediction and decision-support. As such, we identified and excluded those resources. Prompt engineering via Langchain was then used to define the system as an AI assistant specializing in analyzing and interpreting medical documents (additional instructions were provided). The user input was formatted to separate the context of the patient from the query, enabling more precise and consistent responses.

To establish a quantitative basis for assessing ePartogram-GPT's effectiveness, we employed the Term Frequency-Inverse Document Frequency (TF-IDF) natural language processing model. This approach extracted relevant terms from both the correct answer and the solution's response. These terms typically encompassed action plans, assessments, and case-specific jargon. Similarity scores were calculated using cosine similarity, providing a numerical measure of the alignment between the correct answer and ePartogram-GPT's output. Responses from ePartogram-GPT were independently generated from ten clinical cases: there were five unique scenarios with diagnosis and plan of care queries. We found that the percentage of similarity across the ten cases was 13.6%. The similarity scores reflect our model's inability to align with the training answer key. While the cases used for validation were taken from various *Jhpiego* training materials, our maternal health database did not include their respective learning resources, due to the size of the database. Nevertheless, we noted that many of the most important words that the model determined were not relevant to its meaning (see Supplementary Table 3). Furthermore, qualitative analysis of the responses found that the model was better at diagnoses than plans, and often over communicated rather than distilling the most important information. Future attempts will train the ePartogram with a curation of *Jhpiego* materials, include prompt engineering to avoid statements that overgeneralize care,

and validate responses on 50-100 case studies and scenarios, generated by clinicians and adapted from partograms.

Aim 2: Streamlined data input, alerts, and reminders via Natural Language Processing

In addition to interpreting clinical cases and providing diagnoses, the ePartogram-GPT synthesizes user queries and extracts labor-related metrics that are present on the partogram and the WHO Labor Care Guide. Pattern matching tools such as RegEx and spaCy - an industrial strength open-source library - were used to conduct natural language processing on the user input. Metrics were identified and converted to their respective unit or symbol on the WHO Labor Care Guide. Thresholds and monitoring frequencies were extracted from the WHO Labor Care Guide Manual and eZazi warning system. The datetime of each chat is recorded and appended to both a dataframe and a csv file for chat history retention. The functionality of this system was tested on time-based scenarios presented in the WHO Labor Care Guide Manual. Due to GPU RAM limit issues, the full functionality of this feature integrated with ePartogram-GPT was not validated. To mitigate this issue, we plan to optimize memory management through offloading models to disk space.

For the full pipeline, excluding training materials, please see the ePartogram-GPT [GitHub](#).

Future Plans:

We plan to continue validation and fine-tuning of ePartogram-GPT for deployment in LMIC healthcare settings. First, we will incorporate *Jhpiego* Learning Resource Packages to optimize the model's performance in planning care, and fine-tune prompt engineering to limit answer lengths to the most relevant information. We will also incorporate new methods of RAG that further simplify token processing, reducing compute power and increasing efficacy [5]. Through discussions with skilled birth attendants (SBA) and clinicians, we hope to generate test cases and validate model outputs to ensure it meets the nuanced demands of maternal healthcare. Additionally, a mobile or tablet application will be developed to assess the feasibility of using ePartogram-GPT in real-world scenarios. This will also serve as a platform for field testing. Further adjustments will also be made to account for local dialects and language variations to enhance the tool's accessibility and usability. ePartogram-GPT must be incorporated with the eZazi app or a similar system. This integration will create a unified tool for maternal healthcare, leveraging advanced AI capabilities to complement existing workflows. Subsequently, field-testing of eZazi with the enhanced ePartogram-GPT will be conducted to evaluate its practical impact. During this phase, cultural sensitivity and adherence to local regulations will be prioritized, with the model being refined based on observations and feedback from healthcare providers and community members. These steps aim to ensure a culturally appropriate, effective solution that supports maternal and neonatal care.

References

- [1] Khalil A, Samara A, O'Brien P, Coutinho CM, Quintana SM, Ladhani SN. A call to action: the global failure to effectively tackle maternal mortality rates. *The Lancet Global Health*. 2023;11(8):e1165-e1167. doi:10.1016/S2214-109X(23)00247-4
- [2] Maternal mortality rates and statistics. UNICEF DATA. Accessed October 16, 2024. <https://data.unicef.org/topic/maternal-health/maternal-mortality/>
- [3] Maternal mortality. Accessed October 16, 2024. <https://www.who.int/news-room/fact-sheets/detail/maternal-mortality>
- [4] Kurjak A, Stanojević M, Dudenhausen J. Why maternal mortality in the world remains a tragedy in low-income countries and shame for high-income ones: will sustainable development goals (SDG) help? *Journal of Perinatal Medicine*. 2023;51(2):170-181. doi:10.1515/jpm-2022-0061
- [5] Gupta, S., Ranjan, R., & Singh, S. N. (2024). A Comprehensive Survey of Retrieval-Augmented Generation (RAG): Evolution, Current Landscape and Future Directions. *arXiv preprint arXiv:2410.12837*.

Appendix:

Supplementary Table 1:

| Retrieval Augmented Generation Directory |
|--|
| Labor Care Guide - WHO |
| Labor Care Guide Manual - WHO |
| Clinical Rules Table (provided by Dr. Harshadkumar Sanghvi, Jhpiego) |
| Managing Complications in Pregnancy and Childbirth: A guide for midwives and doctors - WHO |
| Antenatal Care - Life-threatening Complications - Jhpiego/USAID |
| Childbirth, Postpartum, and Newborn Care - Life-threatening Complications - Jhpiego/USAID |
| Guidance for Monitoring Labour Progression - ePartogram Development Team |
| Guidance for Monitoring the Baby's Well-Being - ePartogram Development Team |

Supplementary Table 2:

Similarity scores were calculated using cosine similarity between TF-IDF representations of the chat output and Jhpiego reference. Additionally, we independently compared the outputs for the diagnosis and plan of care, meaning we cleared the chat history, reentered the case summary, and then asked for either the diagnosis or plan of care.

| Test Case and Context | Similarity Score - Diagnosis | Qualitative Check - Diagnosis | Similarity Score - Plan of Care | Qualitative Check - Point of Care |
|------------------------------------|-------------------------------------|--|--|--|
| 1 - Common Newborn Problems | 0.20 | More detailed than the answer key, included the same concepts. | 0.17 | More detailed than the answer key, included the same concepts. |
| 2 - Common Newborn Problems | 0.09 | More detailed than the answer key, included the same concepts. | 0.09 | Very different support, not specific enough to case. |
| 3 - Convulsions or Spasms | 0.14 | Output is much more detailed than the answer key, providing a more specific diagnosis compared to the answer key. Output included correct main concepts. | 0.07 | Output does not include the intervals for IV fluid administration and is very broad compared to the answer key. Both focus on the same concepts. |
| 4 - Breathing Difficulty | 0.14 | Output is much more detailed than the answer key, providing a more specific diagnosis compared to the answer key. | 0.21 | Output contains correct main ideas but is missing specific numbers for drug interventions. |

| | | | | |
|-------------------------------|------|--|------|---|
| | | Output included correct main concepts. | | |
| 5 - Maternal Infection | 0.18 | Both contained the same main ideas, but output was more persistent on continuous monitoring. | 0.07 | Output had no correlation with the answer key and focused on other interventions that the answer key did not include. |

Supplementary Table 3:

| | |
|--|---|
| Case 1 Important Words (Diagnosis): | '60' '70' 'ability' 'adequate' 'airways' 'analysis' 'asphyxia' 'assess' 'assessment' 'baby' 'based' 'birth' 'blood' 'breathe' 'breathing' 'breaths' 'care' 'careful' 'catheter' 'causes' 'chest' 'clearance' 'compromising' 'considered' 'consistent' 'cultures' 'delivered' 'determining' 'did' 'difficulty' 'disorders' 'distress' 'drainage' 'efforts' 'ensure' 'evidenced' 'facilitate' 'feeding' 'findings' 'following' 'gas' 'guide' 'help' 'hypoxia' 'identify' 'indrawing' 'intensive' 'investigations' 'levels' 'lungs' 'management' 'mentioned' 'metabolic' 'mild' 'minute' 'moderate' 'monitoring' 'nasal' 'nasogastric' 'necessary' 'neonatal' 'nutrition' 'oxygen' 'percussion' 'persists' 'physiotherapy' 'poor' 'postural' 'prevent' 'prongs' 'prudent' 'rate' 'rays' 'receive' 'regular' 'remove' 'require' 'required' 'requiredresuscitation' 'respiratory' 'saturation' 'secretions' 'sessions' 'severity' 'sign' 'specialized' 'specific' 'steps' 'strategies' 'suck' 'support' 'swallow' 'synchronously' 'taken' 'therapy' 'toxicity' 'transfer' 'tube' 'underlying' 'unit' |
| Case 1 Important Words (Plan): | '24' '60' 'according' 'address' 'addressed' 'adequate' 'aids' 'airway' 'apnoea' 'appropriate' 'assess' 'assessment' 'baby' 'based' 'begins' 'body' 'box' 'breastfeeding' 'breastmilk' 'care' 'causes' 'chest' 'closely' 'decreased' 'detailed' 'difficulties' 'donor' 'effort' 'encouraged' 'established' 'evaluation' 'feeding' 'flow' 'following' 'given' 'grunting' 'health' 'healthcare' 'help' 'hours' 'indrawing' 'intake' 'iv' 'levels' 'maintain' 'manage' 'mechanical' 'minute' 'monitored' 'nasal' 'nipple' 'observationsrecorded' 'oral' 'overall' 'oxygen' 'pattern' 'physical' 'plan' 'positioning' 'positive' 'possible' 'potential' 'presence' 'pressure' 'professional' 'promptly' 'prongs' 'provide' 'provided' 'questions' 'range' 'rate' 'rds' 'reassurance' 'recommend' 'respiratory' 'sessions' 'severe' 'shields' 'shorter' 'signs' |

| | |
|--|--|
| | 'status' 'stay' 'stimulators' 'support' 'synchronization' 'syndrome' 'tachypnea' 'test' 'therapy' 'thorough' 'toask' 'transient' 'ttn' 'undergoes' 'underlying' 'use' 'ventilation' 'vital' 'weight' |
| Case 2 Important Words (Diagnosis): | 'ability' 'additional' 'antibiotics' 'appropriate' 'aspirate' 'baby' 'blood' 'body' 'category' 'cause' 'causing' 'certainly' 'check' 'chest' 'close' 'comprehensive' 'decreased' 'dehydration' 'diagnostic' 'digestive' 'discomfort' 'disorders' 'drowsiness' 'ensure' 'evaluated' 'evaluation' 'fed' 'feeding' 'fever' 'fight' 'findings' 'fluids' 'gastroesophageal' 'gerd' 'growth' 'habits' 'having' 'health' 'healthcare' 'hypoglycemia' 'important' 'inability' 'include' 'infection' 'infections' 'insulin' 'intolerance' 'issues' 'lactase' 'lactose' 'lethargy' 'levels' 'loses' 'low' 'manage' 'medication' 'metabolic' 'monitoring' 'nasal' 'necessary' 'nutrients' 'observation' 'occurs' 'optimal' 'options' 'performed' 'plan' 'poor' 'possible' 'problems' 'professional' 'ray' 'receiving' 'recommended' 'reduced' 'reflux' 'regulate' 'rehydration' 'requires' 'respiratory' 'rule' 'sepsis' 'signs' 'sugar' 'supplements' 'takes' 'test' 'tests' 'thirst' 'tolerance' 'treatment' 'underlying' 'urine' 'weakness' 'working' |
| Case 2 Important Words (Plan): | '12' '120' '24' 'ability' 'acetaminophen' 'additional' 'address' 'addressed' 'antibiotics' 'antiviral' 'appropriate' 'ask' 'assess' 'assessed' 'assessment' 'baby' 'based' 'baths' 'body' 'breast' 'breastfeeding' 'breastmilk' 'care' 'closely' 'decreased' 'encouraged' 'ensure' 'evaluation' 'expressed' 'feeding' 'fever' 'findings' 'fluids' 'following' 'formula' 'gain' 'health' 'healthcare' 'hours' 'infection' 'infections' 'initial' 'intake' 'iv' 'lactose' 'manage' 'management' 'measures' 'monitor' 'necessary' 'nipple' 'nutrients' 'nutritional' 'observation' 'optimal' 'oral' 'output' 'period' 'physical' 'plan' 'possible' 'potential' 'professional' 'promptly' 'provide' 'provided' 'questions' 'reassessed' 'reassurance' 'receiving' 'recommend' 'rehydrated' 'rehydration' 'required' 'respiratory' 'results' 'sepsis' 'shields' 'signs' 'solution' 'source' 'sponge' 'status' 'stay' 'stimulators' 'supplemental' 'supplements' 'support' 'temporary' 'test' 'thorough' 'time' 'tolerated' 'tract' 'treated' 'treatment' 'tube' 'underlying' 'urinary' 'weight' |
| Case 3 Important Words (Diagnosis): | '25' 'additionally' 'address' 'addressing' 'administering' 'anticonvulsant' 'baby' 'based' 'blood' 'checking' 'classic' 'condition' 'consistent' 'context' 'convulsions' 'crucial' 'diagnosis' 'dl' 'experiencing' 'feeding' 'findings' 'glucose' 'hypoglycemia' 'includes' 'indicators' 'initial' 'level' 'levels' 'low' 'management' 'medication' 'mentioned' 'mg' 'mmol' 'necessary' 'provided' 'reason' 'recent' 'recurrence' 'result' 'seizure' 'seizures' 'sign' 'sugar' 'symptoms' 'treating' 'underlying' |
| Case 3 Important Words (Plan): | '10' '12' '24' '25' '30' '45' '80' 'addressing' 'adequate' 'aims' 'alternative' 'anticonvulsant' 'appropriate' 'ask' 'baby' 'blood' 'body' 'bolus' 'breast' 'breastfeed' 'care' 'carefully' 'caregivers' 'conditions' 'consecutive' 'consulting' 'continued' 'convulsions' 'covered' 'cup' 'daily' 'day' 'discussed' 'dl' 'educated' 'encouraged' |

| | |
|--|---|
| | <p>'ensure' 'ensuring' 'given' 'glucose' 'hours' 'hypoglycemia' 'infusion'</p> <p>'iv' 'kg' 'level' 'levels' 'line' 'management' 'measured' 'measurement'</p> <p>'measurements' 'medication' 'method' 'mg' 'minutes' 'ml' 'mmol' 'monitoring' 'mrs' 'necessary' 'needed' 'neonatologist' 'neurological'</p> <p>'newborn' 'nutrition' 'observed' 'ongoing' 'parents' 'pediatrician' 'percent' 'plan' 'provided' 'providing' 'questions' 'readings' 'reassurance' 'repeated' 'required' 'requires' 'safety' 'seizures' 'service' 'signs' 'slowly' 'spoon' 'status' 'stay' 'stays' 'subtle' 'sugar' 'summary' 'support' 'symptoms' 'transferring' 'treat' 'treating'</p> <p>'tube' 'warm' 'weight'</p> |
| Case 4 Important Words (Diagnosis): | <p>'80' 'airway' 'aspiration' 'babies' 'baby' 'based' 'birth' 'born' 'breathing' 'cannula' 'cardiovascular' 'care' 'chest' 'closely' 'common'</p> <p>'consistent' 'context' 'continuous' 'corrective' 'depending' 'diagnosis'</p> <p>'difficulty' 'distress' 'ensure' 'factors' 'fetal' 'findings' 'flow' 'healthcare' 'hfnc' 'high' 'hours' 'incomplete' 'indrawing' 'infection'</p> <p>'initiate' 'levels' 'low' 'lungs' 'maternal' 'measures' 'mechanical' 'meconium' 'moderate' 'monitor' 'nasal' 'ncpap' 'noted' 'occur' 'optimal'</p> <p>'outcomes' 'oxygen' 'positive' 'pressure' 'preterm' 'problem' 'production' 'provide' 'provided' 'providers' 'rapidly' 'rate' 'rds' 'replacement' 'required' 'requiring' 'respiratory' 'result' 'severity' 'specialized' 'status' 'support' 'surfactant' 'symptoms' 'syndrome' 'term' 'therapy' 'underdeveloped' 'various' 'ventilation'</p> |
| Case 4 Important Words (Plan): | <p>'12' 'baby' 'based' 'begins' 'blood' 'breathing' 'care' 'causes' 'cdh' 'chest' 'clinical' 'collaboration' 'congenital' 'continuously' 'diaphragmatic' 'distress' 'encouraged' 'ensure' 'flow' 'given' 'glucose'</p> <p>'grunting' 'healthcare' 'hernia' 'hours' 'including' 'indicate' 'indrawing' 'iv' 'make' 'management' 'mask' 'measured' 'mechanical'</p> <p>'metabolic' 'mg' 'minute' 'ml' 'mmol' 'moderate' 'monitor' 'monitored'</p> <p>'monitoring' 'mrs' 'nasal' 'necessary' 'neonatologist' 'neurological' 'normal' 'nutrition' 'observations' 'ongoing' 'overall' 'oxygen' 'parents' 'pediatrician' 'persists' 'picture' 'plan' 'pneumonia' 'positive' 'presence' 'pressure' 'prongs' 'provide' 'provided' 'providers' 'questions' 'range' 'rapid' 'rate' 'rds' 'reassess' 'reassurance' 'receives' 'recorded' 'regular' 'required' 'requires' 'respiratory' 'response' 'retractions' 'rsv' 'signs' 'specialist' 'specialized' 'status' 'stay' 'supplemental' 'support' 'supportive' 'symptoms' 'syncytial' 'syndrome' 'treated' 'underlying' 'ventilation'</p> <p>'virus' 'volume' 'weight'</p> |
| Case 5 Important Words (Diagnosis): | <p>'18' '24' 'accordingly' 'appropriate' 'baby' 'based' 'birth' 'care' 'circumstance' 'classified' 'close' 'closely' 'complications' 'confirmed'</p> |

| | |
|---------------------------------------|---|
| | <p>'context' 'delivery' 'detected' 'diagnosis' 'early' 'evaluation' 'fact'</p> <p>'far' 'following' 'given' 'hours' 'increases' 'infection' 'infections'</p> <p>'initial' 'membranes' 'monitored' 'monitoring' 'mrs' 'needs'</p> <p>'newborn'</p> <p>'newborns' 'noted' 'observed' 'outlined' 'possible' 'potential'</p> <p>'predisposition' 'prevent' 'principles' 'problems' 'prolonged'</p> <p>'provide'</p> <p>'quickly' 'recommended' 'requires' 'risk' 'routine' 'rupture'</p> <p>'ruptured'</p> <p>'section' 'sepsis' 'service' 'sick' 'signs' 'stated' 'testing' 'transfer'</p> <p>'treated' 'treatment'</p> |
| Case 5 Important Words (Plan): | <p>'36' '37' '48' '5^oc' 'airway' 'approach' 'appropriate' 'babies' 'baby'</p> <p>'balance' 'care' 'carefully' 'continuous' 'control' 'culture'</p> <p>'deficiencies' 'degree' 'delays' 'depending' 'developmental'</p> <p>'developmentally' 'discussed' 'distress' 'early' 'ears' 'electrolyte'</p> <p>'encouraged' 'enteral' 'equipment' 'eyes' 'findings' 'fluid' 'fluids'</p> <p>'gestation' 'hand' 'healthcare' 'hospital' 'hours' 'hygiene'</p> <p>'immature'</p> <p>'immune' 'implement' 'include' 'including' 'increased' 'incubators'</p> <p>'infection' 'infections' 'initiate' 'instability' 'intensive'</p> <p>'intestines' 'involvement' 'involves' 'isolation' 'kangaroo'</p> <p>'laboratory'</p> <p>'listened' 'loss' 'lungs' 'maintain' 'makes' 'management' 'managing'</p> <p>'measures' 'mechanical' 'monitor' 'mrs' 'multidisciplinary' 'nasal'</p> <p>'ncpap' 'need' 'needed' 'neonatal' 'neonatologists' 'nicu' 'normal'</p> <p>'nursery' 'nurses' 'nutrition' 'nutritional' 'nutritionists' 'observed'</p> <p>'optimize' 'parenteral' 'preterm' 'provide' 'providers' 'respiratory'</p> <p>'stimulation' 'support' 'temperature' 'therapists' 'therapy'</p> <p>'thermoregulation' 'underdeveloped' 'unit' 'urinary' 'urine' 'use'</p> |