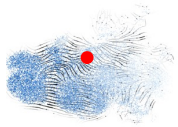
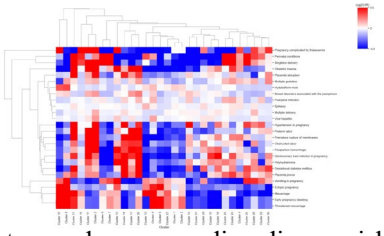
a

Analyze the following EHR data, all collected from one pregnant woman at different visits, and see if she's facing any risk during pregnancy and delivery. Then estimate the fetal biological age, analyze the health status of her offspring after birth, if facing risk in any disease, provide some recommendations.

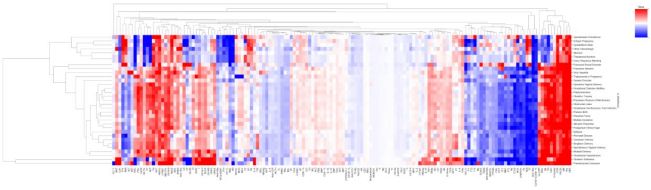
b



Input location (cluster 14) on the trajectory map



Clusters and corresponding disease risks



Disease indicator feature vs input feature
(for disease prediction)

c

**Gestational clock estimation**: 25 weeks
**Fetal clock estimation**: 25 weeks
**Cluster assignment**: cluster 14
**Current disease**: None
**High risk disease**: hypertension in pregnancy
**Hypertension in pregnancy probability**: 10%
**Offspring high risk disease**: None

d

**Hypertension in pregnancy**

Diagnostic criteria

Building on the diagnostic and initial management framework, recent multicenter randomized trials have clarified the impact of antihypertensive therapy for mild chronic hypertension in pregnancy. The CHAP trial demonstrated that active treatment to maintain blood pressure below 140/90 mm Hg significantly reduced the composite risk of preeclampsia with severe features, medically indicated preterm birth, placental abruption, and fetal or neonatal death, without increasing the risk of small-for-gestational-age infants or other adverse neonatal outcomes. This supports a more proactive approach to blood pressure control in pregnancy, with a therapeutic goal of 110–135/85 mm Hg, as endorsed by the American Diabetes Association and other international societies.

Early intervention strategies

extend beyond pharmacotherapy. Low-dose aspirin (81 mg daily, initiated between 12 and 28 weeks gestation, optimally before 16 weeks) is recommended for individuals at moderate or high risk for preeclampsia, with robust evidence for reduction in preeclampsia, preterm birth, and perinatal mortality. Additionally, validated home blood pressure monitoring and telemedicine platforms are emerging as adjuncts for earlier detection and management of hypertensive disorders in pregnancy, potentially improving outcomes through timely intervention.

Treatment options

For acute management of severe hypertension (SBP ≥160 mm Hg or DBP ≥110 mm Hg), rapid-acting agents such as intravenous labetalol, intravenous hydralazine, or immediate-release oral nifedipine are recommended, with comparable efficacy and safety profiles according to network meta-analyses and Cochrane reviews. The choice among these agents should be guided by provider familiarity and institutional protocols. Magnesium sulfate remains the standard for seizure prophylaxis in preeclampsia and eclampsia.

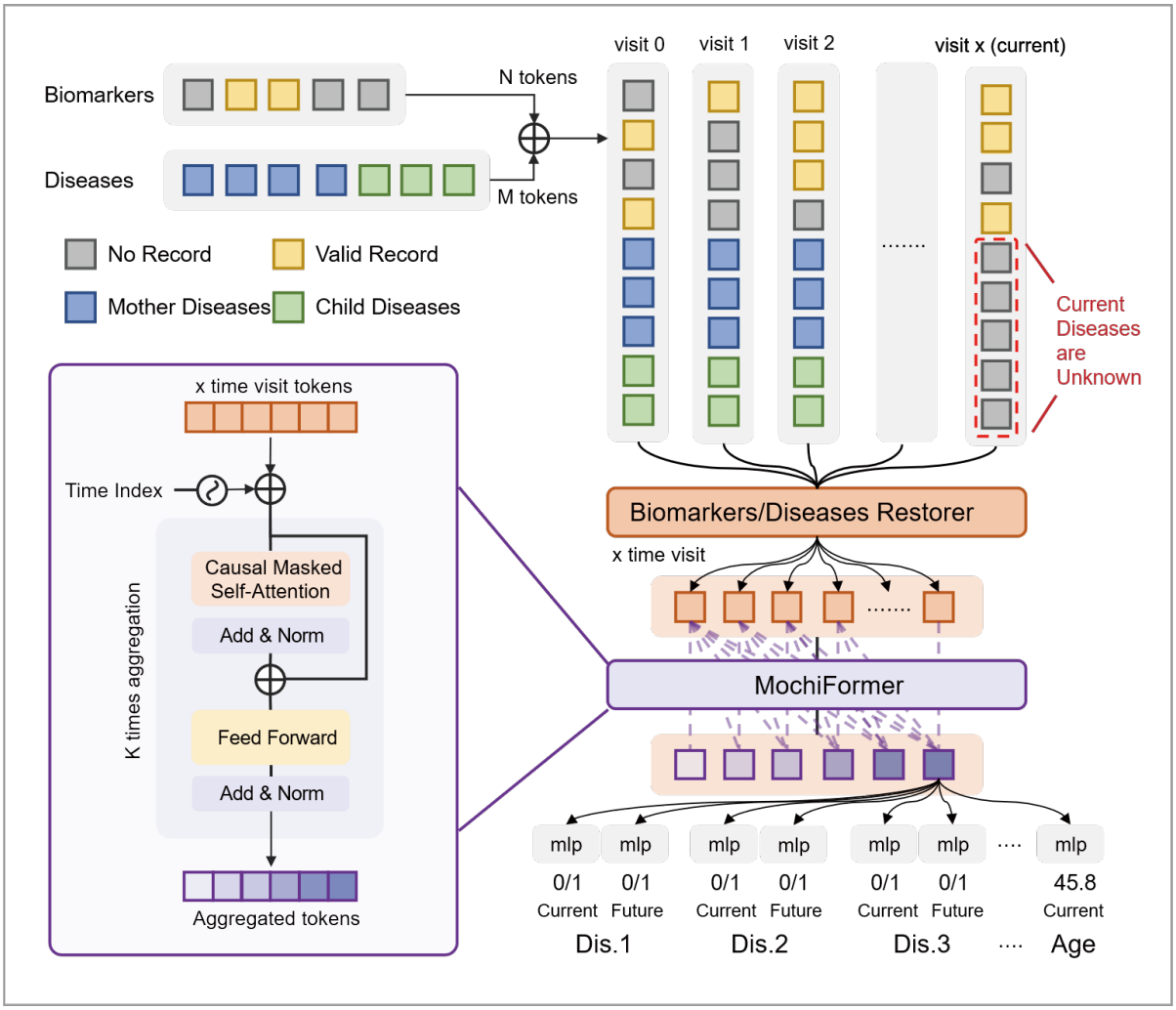
References

1. https://www.ahajournals.org/doi/10.1161/CIR.0000000000001356?url_ver=Z39.88
2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed
2. https://www.nejm.org/doi/full/10.1056/NEJMoa2201295?utm_source=openevidence
3. https://diabetesjournals.org/care/article/48/Supplement_1/S207/157549/10-Cardiovascular-Disease-and-Risk-Management
4. https://www.nejm.org/doi/full/10.1056/NEJMra2109523?utm_source=openevidence
5. https://pmc.ncbi.nlm.nih.gov/articles/PMC9031058/
6. https://www.ahajournals.org/doi/full/10.1161/CIR.0000000000000772?rfr_dat=cr_pub++0pubmed&url_ver=Z39.88-
2003&rfr_id=ori%3Arid%3Acrossref.org

**Supplementary Figure 1. Example of MoChiAgent clinical analysis and evidence-grounded recommendation workflow.**

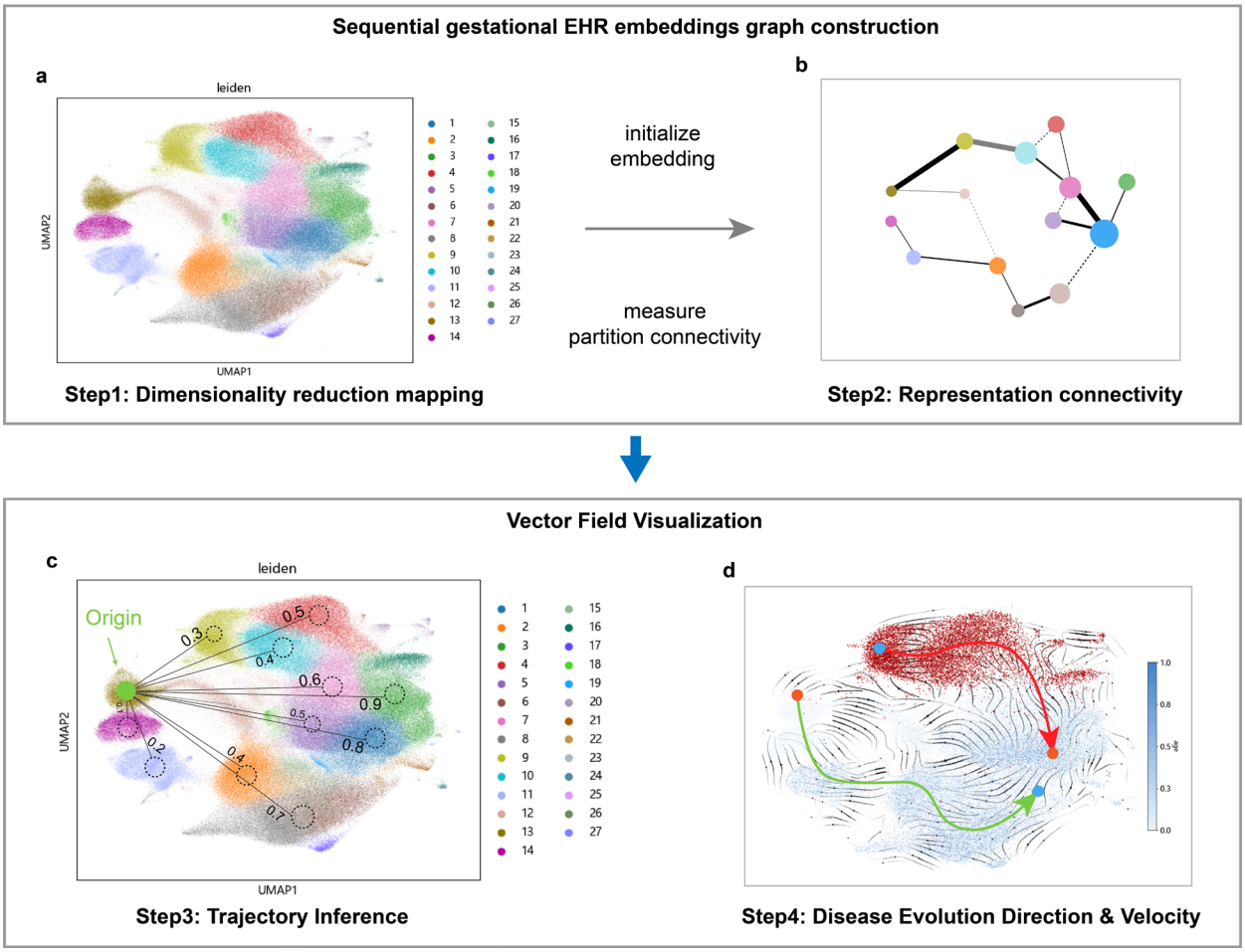
**a.** Example user query requesting analysis of longitudinal EHR data from a pregnant patient across multiple visits, including risk assessment during pregnancy and delivery, fetal biological age estimation, postnatal health evaluation of the offspring, and generation of clinical recommendations. **b.** Model interpretation and contextualization pipeline. The patient's visit representation is mapped onto the gestational trajectory embedding space to determine its location within the learned health-state manifold. Cluster-level disease risk patterns and feature–disease associations are visualized to contextualize the prediction. **c.** Example structured

model output summarizing clinical inference, including estimated gestational and fetal biological clocks, cluster assignment, current disease status, predicted high-risk conditions, and corresponding risk probabilities. **d.** Evidence-grounded clinical guidance generated by the system for the predicted high-risk condition (hypertension in pregnancy), including diagnostic criteria, early intervention strategies, treatment options, and supporting literature references retrieved from trusted medical sources.



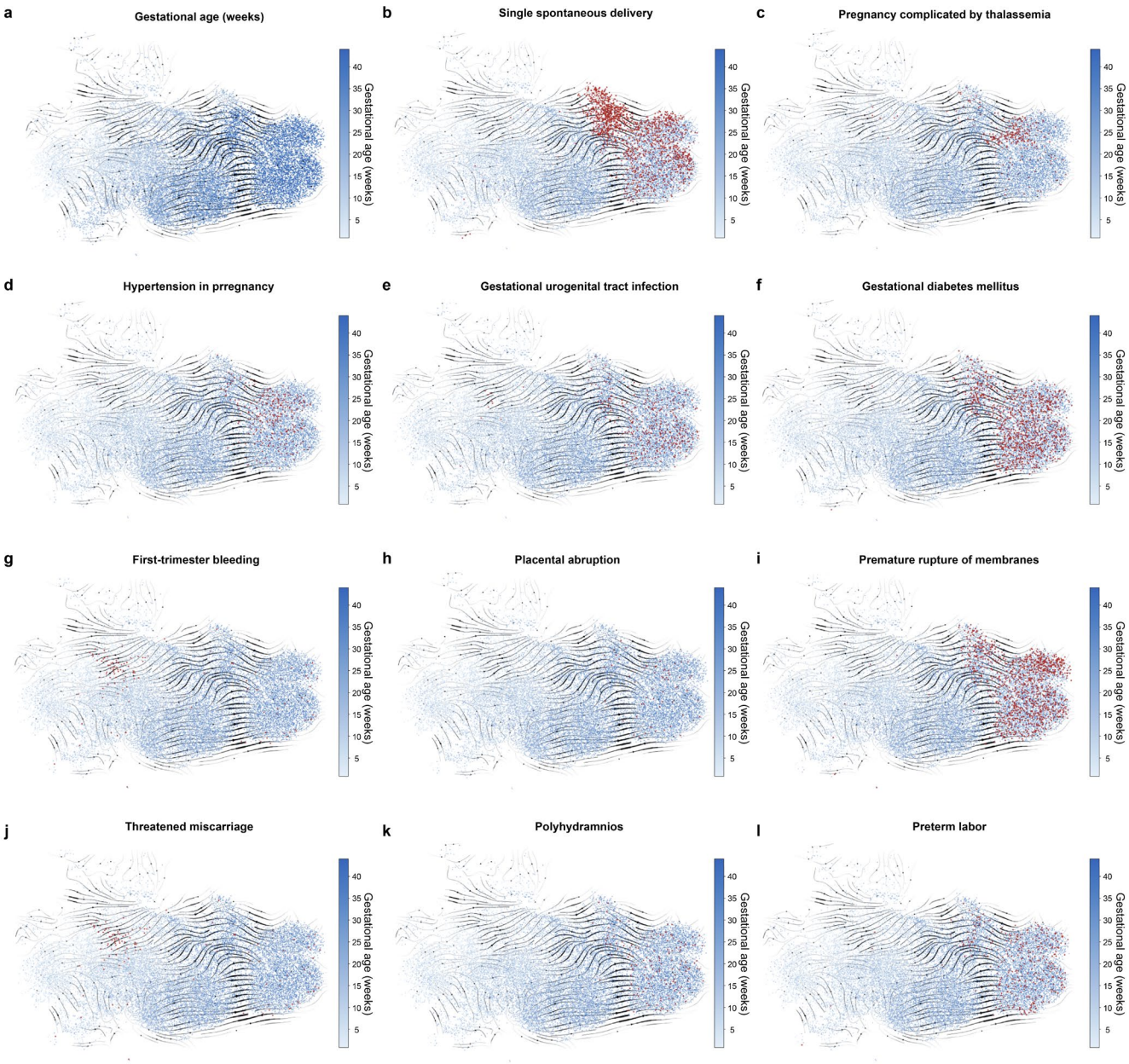

**Supplementary Figure 2. Architecture of the MoChiFormer model for longitudinal EHR analysis.**

Longitudinal electronic health record data, including biomarker measurements and disease diagnoses for mothers and children, are encoded as visit-level tokens across time. Valid and missing records are represented as structured tokens, and disease tokens distinguish maternal and child conditions. A biomarker–disease restoration module first reconstructs incomplete observations and handles missing entries, including unknown current disease states. Temporal visit tokens are then processed through the MoChiFormer architecture, which employs causal masked self-attention to model longitudinal dependencies across visits and aggregate time-aware representations. The resulting patient embeddings are passed to task-specific multilayer perceptron (MLP) heads to perform multiple downstream predictions, including current and future disease risks and biological age estimation.



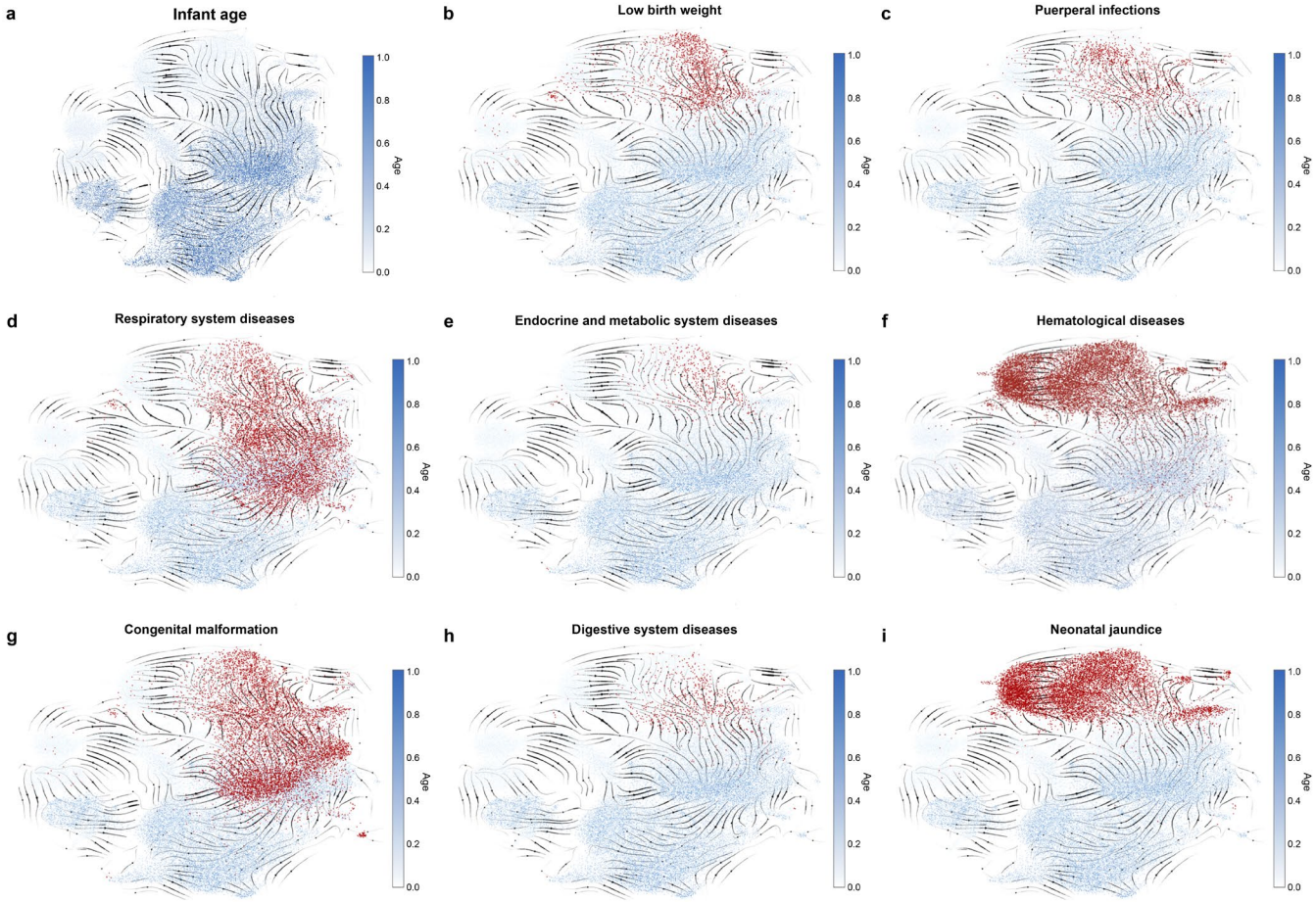

**Supplementary Figure 3. Construction and visualization of sequential gestational EHR embedding trajectories.**

**a,** Dimensionality reduction of longitudinal gestational EHR embeddings using UMAP, followed by Leiden clustering to identify distinct maternal health states. **b,** Construction of a representation graph based on embedding connectivity, where nodes represent clusters and edges denote transition relationships between states. **c,** Trajectory inference from a defined origin state, estimating transition probabilities among clusters to reconstruct potential developmental or disease progression paths. **d,** Vector field visualization of disease evolution in the embedding space, illustrating the direction and velocity of state transitions across the gestational health landscape.



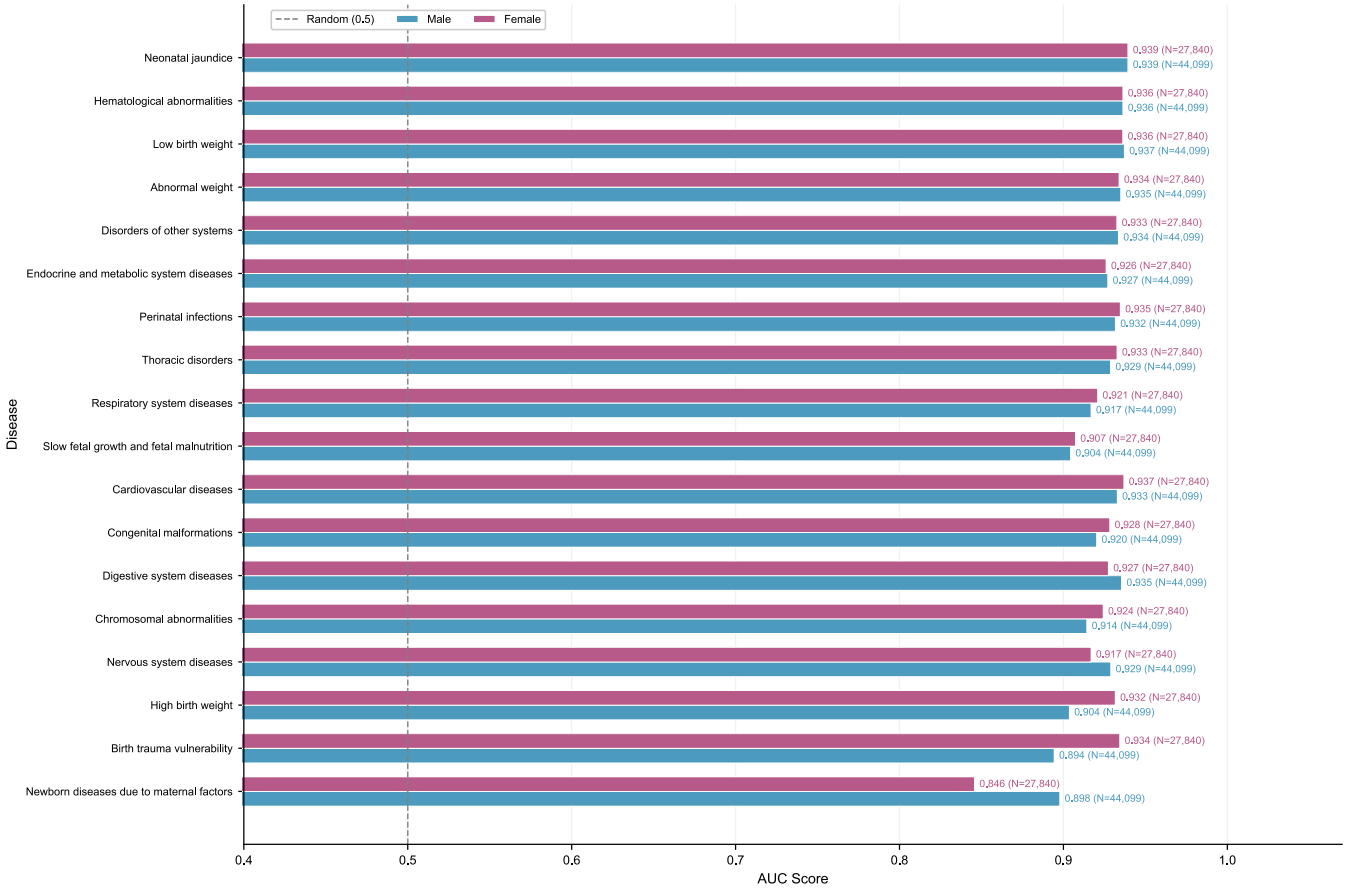

**Supplementary Figure 4. Vector field visualization of gestational trajectories and pregnancy-related conditions in the EHR embedding space.**

**a,** Distribution of gestational age (in days) across the embedding manifold, illustrating the temporal progression of pregnancy. **b-l,** Overlay of representative pregnancy outcomes and complications on the same trajectory landscape, including single spontaneous delivery (b), pregnancy complicated by thalassemia (c), hypertension in pregnancy (d), gestational urogenital tract infection (e), gestational diabetes mellitus (f), first-trimester bleeding (g), placental abruption (h), premature rupture of membranes (i), threatened miscarriage (j), polyhydramnios (k), and preterm labor (l). Each point represents an individual visit embedded in the longitudinal EHR representation space, with streamlines indicating inferred developmental trajectories. Red points denote visits associated with the corresponding condition, while blue points represent the background distribution of all visits, revealing disease-specific regions and progression patterns along gestational trajectories.



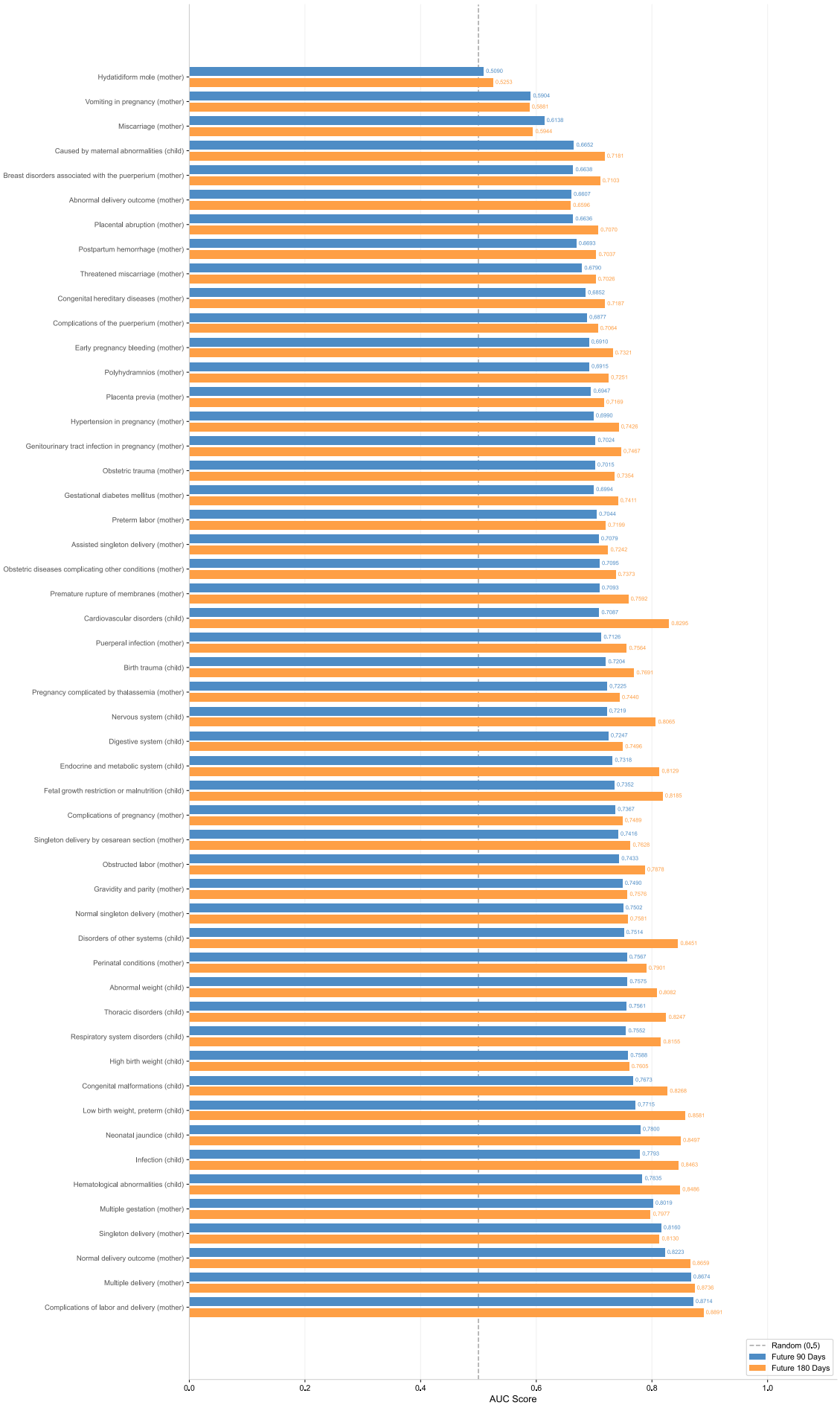

**Supplementary Figure 5. Infant developmental trajectories and disease distributions in the EHR embedding space.**

**a,** Distribution of infant age across the embedding manifold, illustrating developmental progression in early life. **b-i,** Overlay of representative neonatal and pediatric conditions on the same trajectory landscape, including low birth weight (b), puerperal infections (c), respiratory system diseases (d), endocrine and metabolic system diseases (e), hematological diseases (f), congenital malformations (g), digestive system diseases (h), and neonatal jaundice (i). Each point represents an individual visit embedded in the longitudinal EHR representation space, with streamlines indicating inferred developmental trajectories. Red points denote visits associated with the corresponding condition, while blue points represent the background distribution of all visits, highlighting disease-specific regions and patterns along infant developmental trajectories.



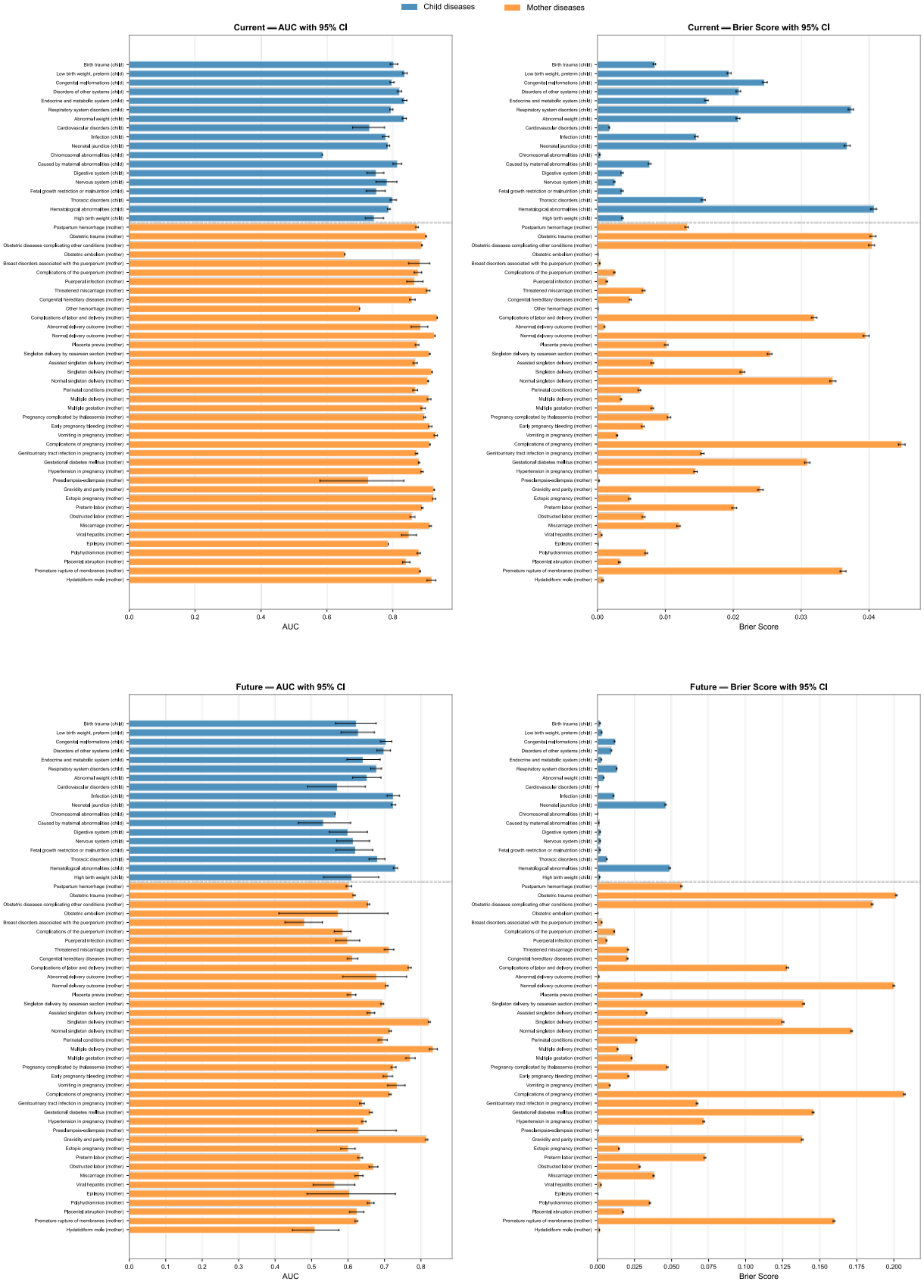

**Supplementary Figure 6. Sex-specific performance of infant disease prediction.**

Area under the receiver operating characteristic curve (AUC) for predicting major neonatal and pediatric disease categories using the MoChiFormer model, stratified by sex. Blue bars represent performance in male infants and pink bars represent performance in female infants, with the dashed line indicating the random baseline (AUC = 0.5). Across a wide range of conditions, including neonatal jaundice, hematological abnormalities, respiratory diseases, congenital malformations, and metabolic disorders, the model achieves consistently high predictive accuracy for both sexes. Sample sizes for each group are indicated alongside the reported AUC values.



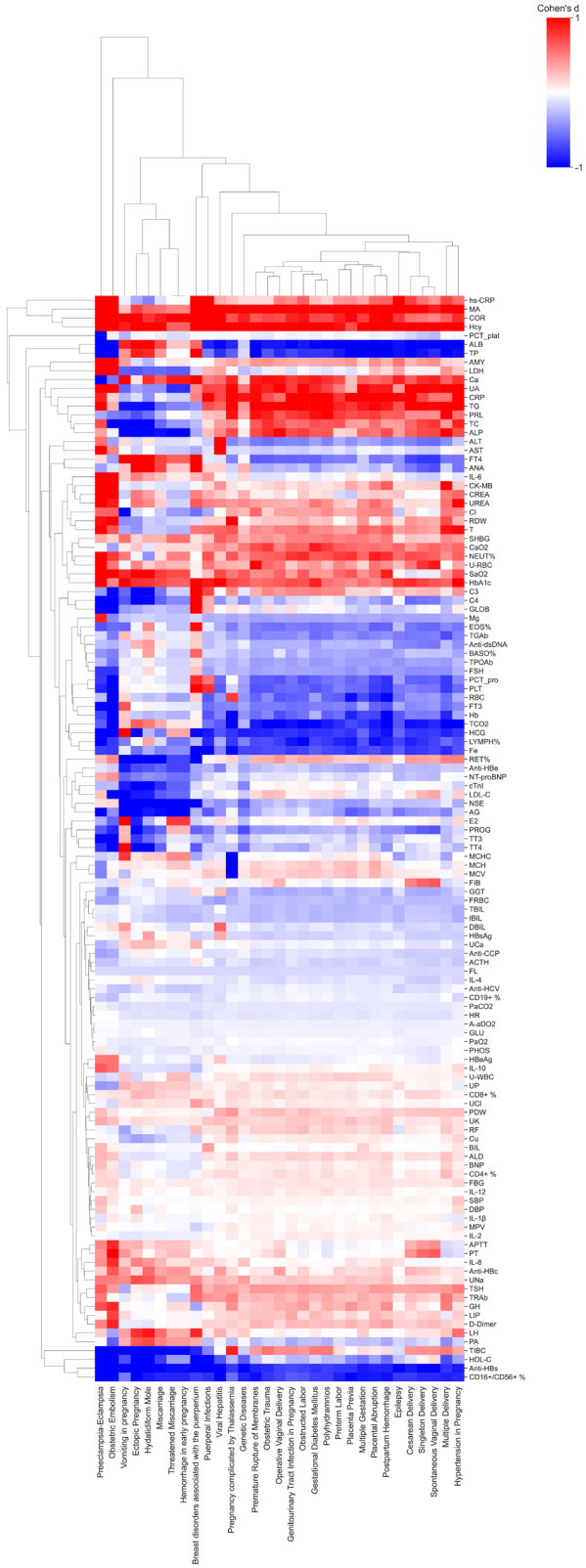

**Supplementary Figure 7. Predictive performance of MoChiFormer for future maternal and infant health outcomes at varying time horizons.**

Comparison of the area under the receiver operating characteristic curve (AUC) for predicting the onset of specific maternal and infant conditions over two distinct time windows: future 90 days (blue bars) and future 180 days (orange bars). The y-axis lists a diverse range of clinical outcomes and complications, labeled by the affected subject (mother or child). The dashed vertical line indicates the random baseline (AUC = 0.5), and specific AUC values are annotated at the end of each bar. MoChiFormer demonstrates robust long-term predictive capability, achieving high accuracy across both short-term and medium-term forecasting horizons for major conditions such as complications of labor, multiple delivery, and various pediatric system disorders.




**Supplementary Figure 8. Overall performance of MoChiFormer for current diagnosis and future prediction of maternal and infant conditions.**

**a,b,** Performance for current disease diagnosis tasks, evaluated using the mean area under the receiver operating characteristic curve (AUC) (a) and Brier score (b). Blue bars represent infant diseases (e.g., birth trauma, low birth weight), and orange bars represent maternal diseases (e.g., postpartum hemorrhage, placenta previa). Error bars indicate standard errors. The model demonstrates high diagnostic accuracy (AUCs consistently exceeding 0.8 or 0.9) and excellent calibration (low Brier scores) across a broad spectrum of obstetric and neonatal conditions. **c, d,** Performance for future disease prediction tasks, evaluated using mean AUC (c) and Brier score (d). Similar to the diagnostic tasks, blue bars denote infant outcomes and orange bars denote maternal outcomes. While future prediction presents a more challenging task than current diagnosis, MoChiFormer maintains robust predictive power and reliable risk probability estimation for prospective maternal complications and infant developmental disorders.

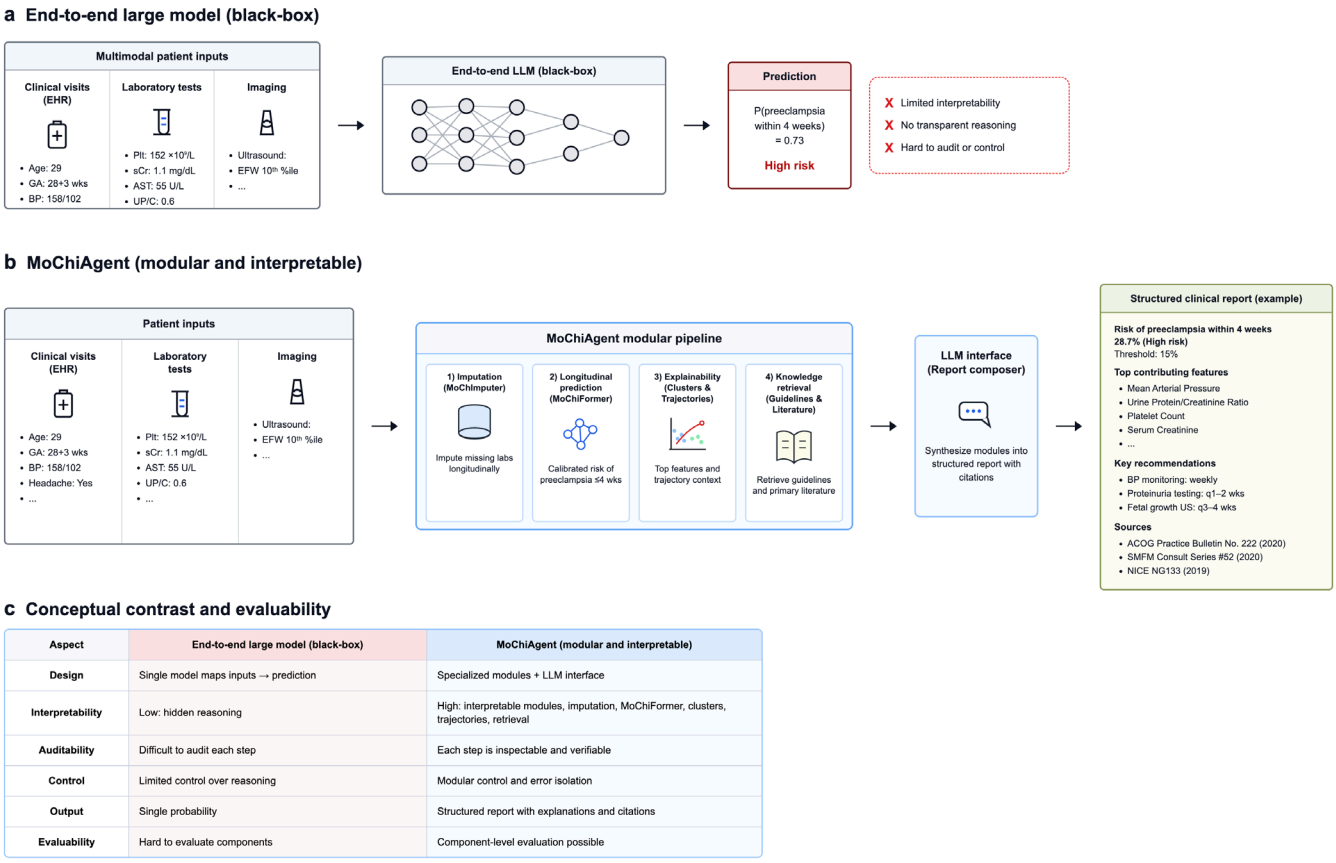



**Supplementary Figure 9. Associations between clinical biomarkers and pregnancy-related conditions.**

Hierarchical clustering heatmap showing effect sizes (Cohen's d) of clinical biomarkers across pregnancy-related diseases and obstetric outcomes. Rows represent laboratory results and clinical measurements, while columns correspond to pregnancy complications and delivery outcomes. Color intensity indicates the magnitude and direction of the association, with red denoting higher values in the disease group and blue denoting lower values relative to controls. Both results and conditions are hierarchically clustered to reveal correlated groups across related pregnancy complications.


**Supplementary Figure 10. Performance of MoChiFormer stratified by maternal age groups.**

**a,** Diagnostic performance evaluated using the area under the receiver operating characteristic curve (AUC) for representative maternal and infant conditions, stratified by maternal age (in years). Columns represent distinct age groups (18–20, 21–30, 31–40, 41+). The blue bars indicate the AUC for MoChiFormer, and the dashed vertical line represents the random baseline (AUC = 0.5). The model demonstrates robust predictive accuracy across the primary childbearing age groups (21–40) for both maternal complications (top section) and infant outcomes (bottom section). Missing bars in the older age groups (41+) reflect data sparsity for pregnancies in these demographics. **b**, Calibration performance evaluated using the Brier score for the same set of conditions and age strata. Lower Brier scores indicate better calibration between predicted probabilities and observed outcomes. The consistently low scores across the populated age groups indicate that MoChiFormer maintains reliable probabilistic risk estimation across different maternal ages.




**c Conceptual contrast and evaluability**

| Aspect | End-to-end large model (black-box) | MoChiAgent (modular and interpretable) |
|---|---|---|
| **Design** | Single model maps inputs → prediction | Specialized modules + LLM interface |
| **Interpretability** | Low: hidden reasoning | High: interpretable modules, imputation, MoChiFormer, clusters, trajectories, retrieval |
| **Auditability** | Difficult to audit each step | Each step is inspectable and verifiable |
| **Control** | Limited control over reasoning | Modular control and error isolation |
| **Output** | Single probability | Structured report with explanations and citations |
| **Evaluability** | Hard to evaluate components | Component-level evaluation possible |

**Supplementary Figure 11. Conceptual schematic of the MoChiAgent system architecture relative to end-to-end large models.**

**a,** Conventional end-to-end large models map multimodal patient inputs (e.g., clinical visits, laboratory tests, and medical images) directly to predictions through a single black-box architecture, with limited interpretability and limited control over the reasoning underlying each prediction. **b,** The proposed MoChiAgent architecture decomposes the analysis into specialized, individually inspectable modules, which include imputation and longitudinal prediction by MoChiFormer, cluster- and trajectory-based interpretability, and a literature retrieval module. An LLM-based interface composes these modules into a structured, citation-backed clinical report. **c,** Conceptual contrast between the two designs: the modular MoChiAgent architecture isolates the empirically validated predictive engine from the interface layer, enabling transparent inspection of each step and providing a natural target for future quantitative benchmarking against end-to-end LLM baselines. The schematic illustrates the intended advantage in interpretability and component-level evaluability.