

Structured Bayesian modelling advances longitudinal EHR research

Xin Chen,¹ Zhonghua Li,¹ Wenyun Yang,¹ Tiantian Liu,¹ and Fangrong Yan^{1,*}

¹Department of Biostatistics, China Pharmaceutical University, Nanjing 211198, China

*Correspondence: f.r.yan@163.com

Received: July 29, 2026; Accepted: September 12, 2026; Published Online: September 14, 2026; <https://doi.org/10.59717/j.xinn-nutri.2026.100036>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Chen X., Li Z., Yang W., et al. (2026). Structured Bayesian modelling advances longitudinal EHR research. *The Innovation Nutrition* 1:100036.

Longitudinal electronic health records (EHRs) are routinely collected clinical data, including diagnoses, procedures, medications, laboratory results and other events recorded during healthcare delivery. Compared with traditional cohort studies, EHRs offer greater scale and richer real-world clinical detail, but their use for research is complicated by incomplete and inconsistent documentation shaped by patterns of healthcare utilization. Recent artificial intelligence (AI) approaches, particularly Transformer-based deep learning models such as Med-BERT and Delphi-2M, have addressed part of this challenge by adapting language-model architectures to structured EHR sequences.¹⁻² These models are flexible and powerful, but they do not directly provide transparent disease phenotypes or explicit links between longitudinal morbidity and genetic discovery.

Urbut and colleagues address this gap with ALADYNOULLI, a structured Bayesian generative framework that jointly models longitudinal diagnoses, age and genetic predisposition.³ Patients are represented by time-varying loadings on a small set of latent disease signatures, and each signature encodes age-dependent risks for related conditions. As disease risk is modelled as an additive mixture of signature-specific probabilities, the framework can represent multiple simultaneous disease processes and persistent chronic conditions more naturally than treating diseases independently.

Across UK Biobank, Mass General Brigham and All of Us, comprising more than 683,000 participants and 348 diseases, ALADYNOULLI recovered 21 signatures with considerable cross-cohort preservation despite differences in population composition, healthcare delivery and coding practice. Several signatures had plausible biological content, including familial hypercholesterolaemia within an ischaemic cardiovascular signature and clonal haematopoiesis within an inflammatory and critical-care signature. Signature-based association analyses also identified genetic loci missed by conventional single-trait analyses, while the explicit likelihood enabled inverse-probability weighting for participation bias.

The central advance of ALADYNOULLI is therefore not simply a new risk predictor, but a model-based framework for defining reproducible longitudinal disease phenotypes. This comment focuses on three issues raised by that advance: what structured Bayesian trajectories gain and lose relative to Transformer-based EHR models, whether the Bayesian specification is fully used for uncertainty quantification, and what would be needed to move from prognostic trajectories toward actionable digital twins.

STRUCTURED PHENOTYPES VERSUS FLEXIBLE PREDICTION

ALADYNOULLI and Transformer-based EHR models are two complementary answers to different modelling questions (Figure 1). Transformer models are strong for event prediction, as self-attention can capture nonlinear, long-range and context-dependent associations and can, in principle, combine diagnoses with medications, laboratory measurements, clinical text and other modalities. Delphi-2M, for example, illustrates how a generative sequence model can estimate future event rates and generate plausible disease trajectories. These properties make Transformer-based approaches natural candidates for multimodal representation learning, forecasting and synthetic trajectory generation. The main limitation is that flexible representations can be difficult to interpret as stable disease phenotypes, and may partly reflect coding practice or healthcare-use patterns rather than disease biology.

ALADYNOULLI makes a more constrained design choice, but that constraint is precisely where its scientific value lies. It represents patients through a limited number of smooth, age-dependent disease signatures, rather than through a high-dimensional neural state. This structure makes the resulting phenotypes easier to inspect biologically, compare across cohorts and use in likelihood-based bias adjustment. Genetic covariates are also explicitly linked to patient-specific signature loadings, which helps

connect longitudinal morbidity with genetic discovery. Its main limitation is the corresponding loss of representational flexibility, as a fixed additive signature model may miss high-order interactions, abrupt transitions or uncommon disease trajectories that Transformer-based EHR models can represent.

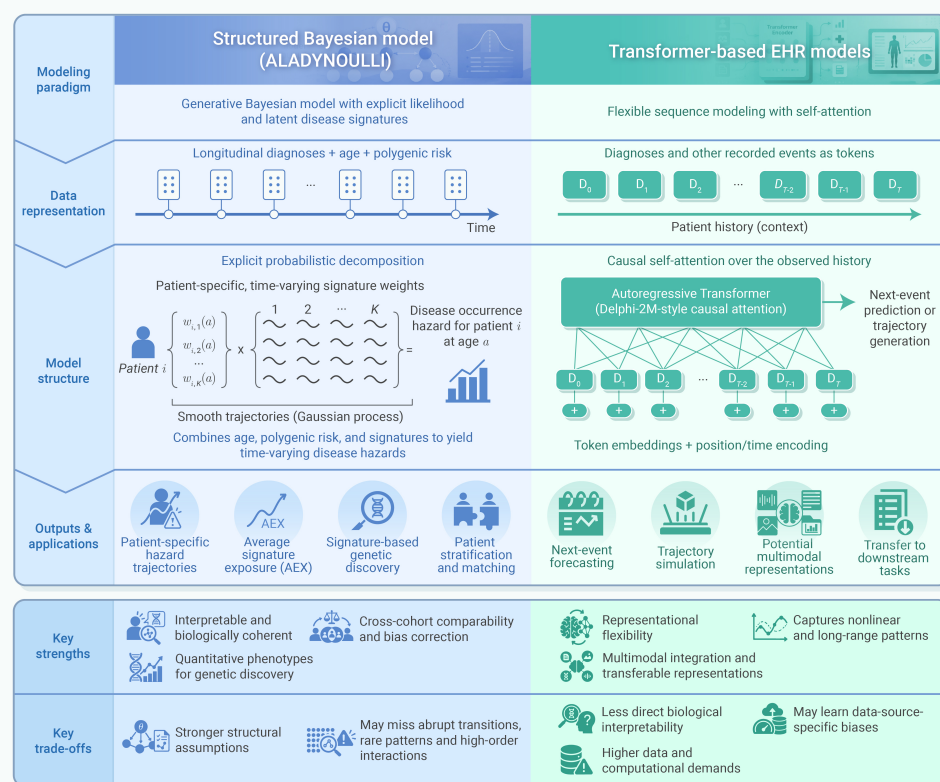
The consequences of these design choices are evident in the outputs reported by the two studies. In ALADYNOULLI, hypercholesterolaemia, coronary atherosclerosis and myocardial infarction contribute to an interpretable ischaemic cardiovascular signature. Patient-specific lifetime exposure to this signature was then used as a quantitative phenotype for genetic association analysis, identifying 56 lead loci in UK Biobank. Delphi-2M, by contrast, was evaluated primarily through future event rates and generated disease trajectories. Its disease embeddings and explainability analyses revealed temporal relationships between diagnoses, but the published study did not define a phenotype or use its representations for genetic locus discovery. These findings illustrate why the choice of model should be guided by the research question. ALADYNOULLI is well suited to interpretable latent phenotyping, genetic discovery and disease-level risk estimation, whereas Transformer-based models are better suited to flexible event prediction, multimodal modelling and trajectory simulation. The two approaches should be treated as complementary rather than as competing alternatives.

Computational requirements present another trade-off. Both ALADYNOULLI and Delphi-2M were trained on approximately 400,000 UK Biobank participants. The ALADYNOULLI implementation reports 8–10 min of training per 10,000-participant batch on a multicore CPU, with four or more cores recommended, and approximately 5–7 hours for the UK Biobank training when run sequentially. The 2.2-million-parameter Delphi-2M model took approximately one hour to train on an NVIDIA V100 GPU. Direct comparison of computational efficiency is difficult because the two implementations were evaluated on different hardware. However, ALADYNOULLI's batchwise implementation can run on standard multicore CPU hardware without GPU acceleration, which may lower the infrastructure required for training. On the other hand, the computational cost of standard self-attention grows quadratically with sequence length. Extending Transformer models to longer histories and multimodal EHR data may therefore require much greater computational resources.

BAYESIAN OPPORTUNITY FOR QUANTIFYING UNCERTAINTY

ALADYNOULLI's Bayesian specification is valuable not only for regularizing a generative model, but also for providing a basis for quantifying uncertainty. In the current implementation, inference relies on the maximum a posteriori (MAP), which is computationally sensible at Biobank scale and supports robust point estimates. A natural extension would be to use the Bayesian specification more explicitly to quantify uncertainty around latent trajectories, disease associations and derived signature exposures.

Such uncertainty quantification could be especially useful when EHR evidence is sparse or ambiguous. For patients with short or fragmented records, posterior uncertainty around signature loadings could distinguish confident phenotypic assignment from weak evidence. For rare diseases or small subgroups, uncertainty around disease-signature associations could prevent overinterpretation of unstable patterns. In genetic analyses, propagating uncertainty in lifetime signature exposure could help separate robust genetic signals from uncertainty introduced by phenotype construction. And in clinical prediction, credible intervals may be more informative than point estimates when decisions depend on whether a patient is clearly above or near an action threshold. As full Markov Chain Monte Carlo (MCMC) may be unrealistic at Biobank scale, future work could examine whether scalable variational, Laplace or ensemble-based approximations can provide useful uncertainty quantification.⁴



FROM PROGNOSTIC TRAJECTORIES TO ACTIONABLE DIGITAL TWINS

ALADYNOULLI represents disease risk as longitudinal, patient-specific trajectories that can be updated as clinical histories accrue. This feature makes it relevant to medical digital twins, which also require patient-specific models that evolve over time and can be used to simulate future outcomes under alternative interventions. However, ALADYNOULLI's trajectories remain prognostic rather than interventional. The model learns from observational disease histories, but it does not identify how the same patient would evolve under different treatment strategies. Turning this type of model into part of an actionable digital twin would require treatment-aware longitudinal data, explicit causal identification assumptions, counterfactual trajectories under well-defined interventions and prospective validation of individualized treatment response predictions.⁵ The key challenge is that EHRs record care as delivered, so treatment choices reflect indication, disease severity and health-care access, while post-treatment outcomes may be incompletely captured. These features make it difficult to infer treatment effects from observed data. Future work will need to connect trajectory modelling with explicit causal designs if digital twins are to move from forecasting disease risk to guiding clinical action.

CONCLUSION

In summary, ALADYNOULLI shows the value of imposing interpretable structure on longitudinal EHR data, while Transformer models offer complementary flexibility for complex sequence modelling. Further use of the Bayesian framework could make uncertainty in patient trajectories and derived phenotypes more explicit. Looking ahead, combining structured phenotypes with flexible representations, scalable uncertainty quantification and causal modelling may support more reliable disease forecasting and

Figure 1. Complementary paradigms for longitudinal EHR modelling ALADYNOULLI uses explicit probabilistic decomposition, smooth latent disease signatures and patient-specific loadings to support interpretable disease-level prediction, genetic discovery and bias adjustment. Transformer-based EHR models use flexible sequence representations and self-attention to support event-level prediction, multimodal learning and trajectory generation. The two paradigms address different but complementary objectives, and future hybrid models could combine flexible encoders with structured latent phenotypes and calibrated uncertainty.

provide a stronger basis for clinically actionable digital twins.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used ChatGPT in order to polish the language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

REFERENCES

- Rasmy L., Xiang Y., Xie Z., et al. (2021). Med-BERT: Pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *npj Digit. Med.* **4**:86. DOI:10.1038/s41746-021-00455-y
- Shmatko A., Jung A.W., Gaurav K., et al. (2025). Learning the natural history of human disease with generative transformers. *Nature* **647**:248–256. DOI:10.1038/s41586-025-09529-3
- Urbat S.M., Ding Y., Nakao T., et al. (2026). A Bayesian framework for longitudinal EHR and genetic discovery. *Nature* **656**:924–935. DOI:10.1038/s41586-026-10780-5
- Blei D.M., Kucukelbir A. and McAuliffe J.D. (2017). Variational inference: A review for statisticians. *J. Am. Stat. Assoc.* **112**:859–877. DOI:10.1080/01621459.2017.1285773
- Corral-Acero J., Margara F., Marciniak M., et al. (2020). The 'digital twin' to enable the vision of precision cardiology. *Eur. Heart J.* **41**:4556–4564. DOI:10.1093/eurheartj/ehaa159

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 82273735; 82473737; 82404380; 82674786). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization: Xin Chen and Fangrong Yan; Writing--Original Draft: Xin Chen and Zhonghua Li; Writing--Review & Editing: Tianian Liu, Wenyun Yang and Fangrong Yan. All authors contributed to and approved the manuscript.

DECLARATION OF INTERESTS

Fangrong Yan is an Editorial Board member of The Innovation Nutrition and was blinded from reviewing or making final decisions on the manuscript. Peer review was handled independently of this member and their research group. The other authors declare no conflicts of interest.