



OPEN

## Incorporating longitudinal history of risk factors into atherosclerotic cardiovascular disease risk prediction using deep learning

Jingzhi Yu, Xiaoyun Yang, Yu Deng, Amy E. Krefman, Lindsay R. Pool, Lihui Zhao, Xinlei Mi, Hongyan Ning, John Wilkins, Donald M. Lloyd-Jones, Lucia C. Petito & Norrina B. Allen 

It is increasingly clear that longitudinal risk factor levels and trajectories are related to risk for atherosclerotic cardiovascular disease (ASCVD) above and beyond single measures. Currently used in clinical care, the Pooled Cohort Equations (PCE) are based on regression methods that predict ASCVD risk based on cross-sectional risk factor levels. Deep learning (DL) models have been developed to incorporate longitudinal data for risk prediction but its benefit for ASCVD risk prediction relative to the traditional Pooled Cohort Equations (PCE) remain unknown. Our study included 15,565 participants from four cardiovascular disease cohorts free of baseline ASCVD who were followed for adjudicated ASCVD. Ten-year ASCVD risk was calculated in the training set using our benchmark, the PCE, and a longitudinal DL model, *Dynamic-DeepHit*. Predictors included those incorporated in the PCE: sex, race, age, total cholesterol, high density lipid cholesterol, systolic and diastolic blood pressure, diabetes, hypertension treatment and smoking. The discrimination and calibration performance of the two models were evaluated in an overall hold-out testing dataset. Of the 15,565 participants in our dataset, 2170 (13.9%) developed ASCVD. The performance of the longitudinal DL model that incorporated 8 years of longitudinal risk factor data improved upon that of the PCE [AUROC: 0.815 (CI 0.782–0.844) vs 0.792 (CI 0.760–0.825)] and the net reclassification index was 0.385. The brier score for the DL model was 0.0514 compared with 0.0542 in the PCE. Incorporating longitudinal risk factors in ASCVD risk prediction using DL can improve model discrimination and calibration.

The Pooled Cohort Equations (PCE) were developed by the American College of Cardiology (ACC) and American Heart Association (AHA) in 2013 and updated in 2018 using data from 9 longitudinal cohort studies as a tool for clinicians to predict 10-year risk of atherosclerotic cardiovascular disease (ASCVD)<sup>1,2</sup>. The PCE are a set of race- and sex-specific Cox proportional hazards models, that include widely-accepted clinical and behavioral risk factors for ASCVD, including age, sex, race, systolic (SBP) and diastolic blood pressure (DBP), total cholesterol, high density lipid-protein (HDL) cholesterol, smoking status, and type 2 diabetes. In clinical practice, risk predictions from the PCE are a key criterion to determine eligibility for moderate to high intensity statins and hypertension treatments<sup>1,3</sup>. However, numerous studies have found the performance of the PCE varies across demographic groups<sup>4–6</sup>; c-statistics from these studies ranged from 0.55 to 0.77 (average: 0.70) in men and 0.61 to 0.82 (average: 0.74) in women<sup>7,8</sup>. Additionally, current clinical guidelines provide more ambivalent and complex treatment recommendations for those who fall in the borderline (5% to 7.5%) and intermediate risk groups (7.5% to 20%)<sup>9</sup>. A more accurate and robust risk prediction algorithm can help physicians better assess an individual's risk, allowing them to make more appropriate treatment decisions.

A growing number of studies have demonstrated that long-term risk factor levels are associated with an individual's risk for the development of ASCVD. For instance, incident CVD risk was shown to be dependent on cumulative exposure to LDL-C<sup>10</sup>. In a separate study, incident CVD and survival were also found to be associated with 10-year cumulative SBP<sup>11</sup>. Hence, long-term risk factor patterns may be predictive of ASCVD risk above and beyond cross-sectional levels<sup>12</sup>. In a prior study, after including 5-year and 10-year cumulative blood pressure measurements in the PCE, researchers found a moderate improvement in the net reclassification index<sup>13</sup>. Additionally, full integration of multiple longitudinal trajectories of clinical factors into ASCVD prediction is

Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.  
✉ email: [Norrina-allen@northwestern.edu](mailto:Norrina-allen@northwestern.edu)

now feasible in clinical practice given advances in computing and electronic medical record (EMR) systems that allow clinicians to access longitudinal risk factor data for their patients.

In recent years, deep learning methods have been applied to many clinical predictive and classification problems to much success<sup>14–16</sup>. Compared with traditional statistical methods, deep learning methods are often superior at processing and creating representations of complex data, such as radiology images and unstructured physician notes<sup>17,18</sup>, without the need of prior feature engineering or selection<sup>15,19</sup>. Hence, deep learning can more thoroughly extract and leverage the rich features stored in longitudinal data such as longitudinal blood pressure measurements recorded in the electronic health records (EHR) for predictive tasks.

In this study, we incorporated cross-sectional and longitudinal clinical and behavioral risk factor levels into a state-of-the-art deep learning architecture to create a new prediction model for 10-year risk of incident ASCVD in a pooled cohort of 4 US-based, diverse longitudinal cohorts. We evaluated our model's predictive performance in comparison to that of the PCE in the overall population and in key population subgroups to better understand the importance of longitudinal data for ASCVD risk prediction. Moreover, we determined the importance of each clinical variable used in the prediction model. Lastly, we performed additional evaluations of the model performance in the borderline and intermediate risk groups to better understand our model's potential impact on clinical decision making.

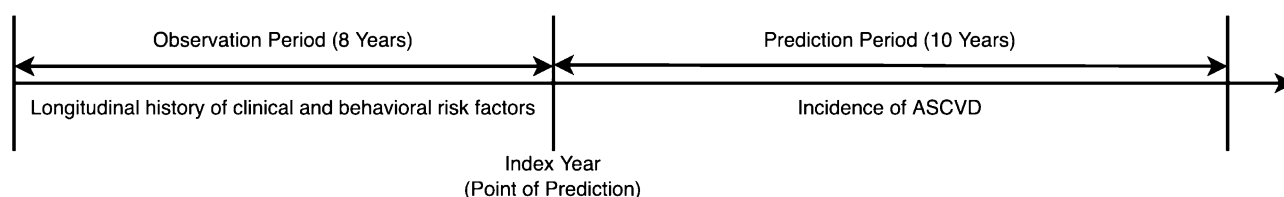
## Methods

### Study population

The four longitudinal cohorts used in this study contributed data to the Cardiovascular Lifetime Risk Pooling Project (LRPP): the Framingham Heart Study, Framingham Offspring Study, Coronary Artery Risk Development in Young Adults (CARDIA) Study, and Atherosclerosis Risk in Communities (ARIC) Study<sup>20</sup>. These cohorts were selected for their number of participants, duration of follow-up, number of participant visits, and consistency of measurement of CVH risk factors.

As the examination schedules differed across cohorts, the number of exams within timeframes varied. To include the largest number of exams across the different studies while balancing the size of the timeframe for the study, we used 8 years of longitudinal data as the timeframe for CVD risk factor ascertainment (observation period). For consistency with the PCEs, outcomes were then measured over a 10-year follow-up period. Thus, to maximize the number of exams included in our study, we included data beginning at the following index exams (i.e. the exam at which risk factor follow-up began) for the included studies (Fig. 1): year 15 for the Framingham Heart Study, year 10 for the Framingham Offspring Study, year 18 for the CARDIA study, and year 1 for the ARIC study. The exact start and end years of each cohort as well as their mean and interquartile range of the number of exams in each cohort are shown in Table 1.

Eligible participants were over 40 and under the age of 75 years at the point of prediction (i.e. the end of the 8 year observation period), had no record of self-report or diagnosed ASCVD at the index exam or during the 8 year observation period, and had at least one measurement of SBP, DBP, total cholesterol and HDL cholesterol. The LRPP is approved by the Northwestern IRB and this study utilized de-identified data from each of the included cohorts in LRPP. Written informed consent was obtained for all participants and analysis were performed in accordance with relevant guidelines.



**Figure 1.** Timeframe of the study. The longitudinal history of clinical and behavior risk factors within the 8 years prior to the index year were included in our prediction model. Our prediction model predicts the outcome within 10 years after the index year.

Cohort	Official start year	Start year of observation period	End year of observation period	End of follow-up	Average number of exams	Interquartile range
ARIC	1987	1987	1995	2005	3	3–3
CARDIA	1985	2003	2011	2021	2	2–2
FHS	1948	1963	1971	1981	2	1–3
FOF	1971	1981	1989	1999	2	2–2

**Table 1.** The official start year, start year of the observation period (after adjustment), end year of the 8 year follow-up period, average number of exams within the 8 year follow-up period as well as the interquartile range of the number of exams by each cohort.

### Outcome: ASCVD incidence

The outcome in our study was ASCVD incidence, defined as the incidence of coronary heart disease, ischemic stroke, or CVD-related death, over a 10-year period that began at the end of the observation period (Fig. 1)<sup>11,20</sup>. Coronary heart disease and ischemic stroke were adjudicated by review of medical records by study investigators<sup>20</sup>. Participants without any recorded event at the end of the study, or who died of other causes during the follow-up period were considered right censored.

### Features: CVD risk factors

CVD risk factors included in the original PCE include systolic BP, diastolic BP, total cholesterol, and HDL cholesterol, and were measured 1–4 times during the 8-year observation period. Blood pressure was measured using standard methods by clinic staff in the various cohorts<sup>20,21</sup>. Fasting HDL-C, total cholesterol measurements and blood glucose were collected via blood serum<sup>20,21</sup>. Diagnosis of diabetes and treatment for hypertension, predictors also included in the PCE, were self-reported at the index visit<sup>20,21</sup>. Age, sex, race, ethnicity, smoking status (current/former smoker vs. never smoker), and alcohol consumption were self-reported at the index visit<sup>20,21</sup>.

### Statistical analysis

The deep learning model used in this study is *Dynamic-DeepHit*, which enabled the incorporation of longitudinal risk factor data in a dynamic fashion to estimate 10-year risk of incident ASCVD<sup>22</sup>. The *Dynamic-DeepHit* model has been demonstrated to have substantial improvements over traditional predictive methods, including the Cox Proportional Hazards Model, in predicting cystic fibrosis outcomes<sup>22</sup>.

The *Dynamic-DeepHit* model consists of two neural networks: (1) a recurrent neural network (RNN) that processes the longitudinal measurements and predicts future measurements of time-varying covariates, and (2) a fully connected neural network that estimates the probability of the specific event at a given time. RNNs are commonly used for machine learning problems involving temporal or sequential data and can capture long-term dependencies in the data. The *Dynamic-DeepHit* model also utilizes an attention mechanism that identifies important longitudinal measurements when making risk predictions, which improves predictive performance. The second neural network takes as input the learned representations that are output from the first neural network along with the last recorded set of behavioral and clinical covariates (e.g. the most recent CVD risk factor measurements at the end of the 8-year observation period). The output layer of the second neural network converts the learned relationships between the risk factors and outcome into the 10-year risk of incident CVD.

To explore the reasons for any improvements in the predictive power we also implemented a cross-sectional *DeepHit* model. This allowed us to disentangle whether the improvements were due to the incorporation of the longitudinal data or simply to the complexity of the neural network modeling methods. The *DeepHit* model was fitted on only the last set of measurements for each participant within the 8-year observation period. We also fit the traditional PCE model, to understand its performance in this sample.

Data pre-processing included randomly splitting the dataset into 3 parts, called training, tuning, and testing, at a 3:1:1 ratio. The *Dynamic-DeepHit* and cross-sectional *DeepHit* models were trained in the training dataset and corresponding hyperparameters were tuned in the tuning dataset. The training data for the PCE included both the training and tuning datasets. The testing dataset, not used in model development, was used for validation. The participants were the same in each of the respective datasets for each model.

We assessed model discrimination and calibration of all 3 models. We calculated and compared the Area Under the Receiver Operator Curve (AUROC) for all models to evaluate model discrimination, the ability of the model to discriminate those who have a higher risk of having an event from those at lower risk. Brier scores were used to evaluate the calibration of the model; lower scores indicate better calibration, the extent of the estimated risk correspond to observed event rates<sup>23</sup>.

The trained *Dynamic-DeepHit* model was evaluated in the following population groups: Black males, Black females, other (White, Hispanic, Asian) males, other females, under 60 years old and 60 or over years old. These demographic groups were chosen to mirror the same classifications used for the sex- and race-specific PCE. As in the overall analysis, the AUROCs were compared between corresponding population subgroups.

To understand the importance of each predictor in the *Dynamic-DeepHit* model, we took a leave-one-out approach. We removed one predictor at a time from the *Dynamic-DeepHit* model and retrained and retested the model. The change in the testing dataset AUROC was calculated for each feature removed: the greater the change in AUROC, the greater the importance of the predictor. To also understand the role of longitudinal clinical risk factors better in the *Dynamic-DeepHit* model, we examined the average trajectories of SBP, DBP, total cholesterol and HDL for the individuals whose predicted risk increased and those whose risk decreased in the *Dynamic-DeepHit* model. Trajectories were created via generalized estimating equations (GEE) to account for correlation between repeated measurements for individuals. The trajectories were visualized across exam times with the 95% confidence bands.

Current blood pressure and cholesterol control guidelines use risk thresholds based on the PCE to inform clinical care. Physicians are advised to prescribe medium intensity statins if an individual's ASCVD risk is over 7.5%. However, differentiation of individuals between the borderline and intermediate PCE risk groups could be improved. We calculated the net reclassification index (NRI) between the PCE and the *Dynamic-DeepHit* model, to understand how the *Dynamic-DeepHit* model changed individuals' risk classification. We then conducted additional analysis to better understand the performance of the *Dynamic-DeepHit* model in borderline and intermediate groups, and how clinical behavior would be affected if the risk derived from the *Dynamic-DeepHit* model was used instead of risk from the PCE.

All statistical analysis was performed using Python version 3.8 and R 4.0.2. A 5% type-I error rate was used when calculating all confidence intervals.

## Results

### Baseline characteristics

Baseline demographics and measurements of CVD risk factors included in the PCE are described in Table 2. Pooled cohort participants included in this study were 55% female, 27% non-Hispanic Black and 50 years old on average. We found participants who developed ASCVD in prediction period had significantly higher levels of ASCVD risk factors compared with the participants who did not develop ASCVD. Baseline demographics and clinical characteristics of the participants by cohort is presented in Supplemental Table S1.

### Performance of models

Table 3 shows the discrimination of the three models in the training and testing datasets. The AUCs for the PCE and the longitudinal *Dynamic-DeepHit* model were 0.792 (CI 0.760–0.825) and 0.815 (CI 0.782–0.844), respectively. The *Dynamic-DeepHit* model shows slight improvement in discrimination upon the PCE model. The cross-sectional deep learning model achieved an AUC of 0.807 (CI 0.778–0.838) (Supplemental Table S3). The continuous net reclassification index (NRI) for the *Dynamic-DeepHit* model compared with the PCE was 0.385. The Brier Score for the PCE model was 0.054, 0.052 for the cross-sectional deep learning model and 0.051 for the longitudinal deep learning model, showing meaningful improvement in model calibration.

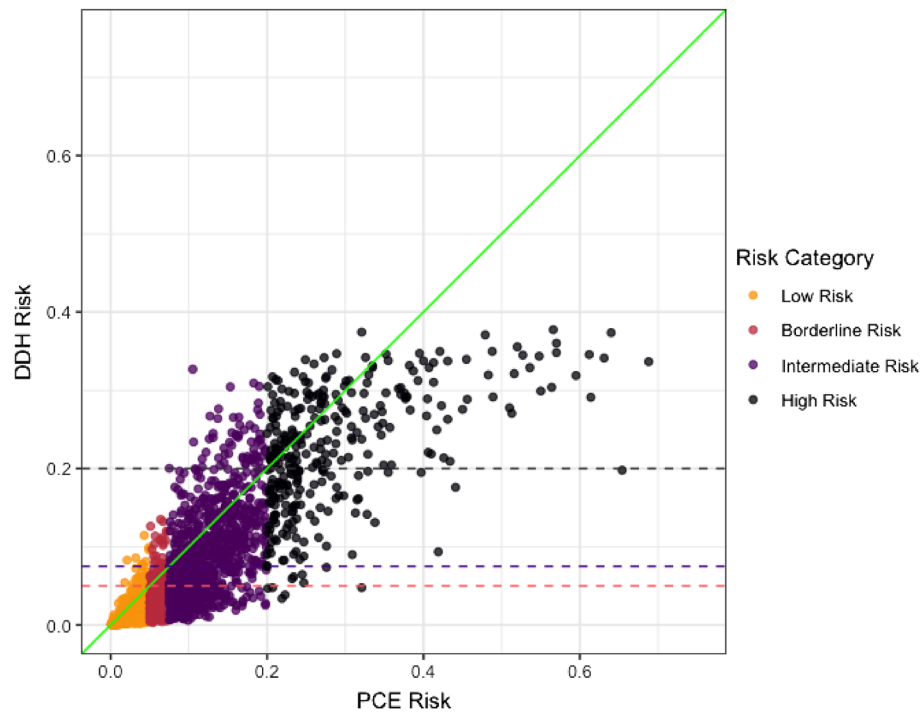
The predicted risks derived from the *Dynamic-DeepHit* model were found to be generally lower than the risks derived from the PCE (Fig. 2). In Fig. 3, the calibration of the *Dynamic-DeepHit* model is compared with the calibration of the PCE by comparing the predicted risk and observed risk within each decile of predicted risk. The PCE is shown to over-predict 10-year ASCVD risk, especially within the top 40% of predicted risk, which corresponds to the 7.5% risk threshold used in clinical guidelines. Comparatively, the calibration of the *Dynamic-DeepHit* model is consistently better along the entire spectrum of risk.

Variables	Overall	Non-ASCVD	ASCVD
	(N = 15,565)	(N = 13,395)	(N = 2170)
Age (Mean, SD)	50.2 (7.2)	49.6 (7.0)	53.7 (6.9)
Sex			
Male	7028 (45.2%)	5798 (43.3%)	1230 (56.7%)
Female	8537 (54.8%)	7597 (56.7%)	940 (43.3%)
Race			
Other	11,366 (73.0%)	9750 (72.8%)	1616 (74.5%)
Black	4199 (27.0%)	3645 (27.2%)	554 (25.5%)
Systolic blood pressure, mmHg	121 (18.7)	120 (17.6)	131 (21.9)
Diastolic blood pressure, mmHg	74.8 (11.6)	74.1 (11.2)	78.9 (12.9)
Total cholesterol, mg/dL	207 (42.9)	205 (41.6)	223 (47.3)
High density lipid, mg/dL	51.6 (16.5)	52.4 (16.5)	46.9 (15.6)
Smoker (vs. never smoker)	4623 (29.7%)	3752 (28.0%)	871 (40.1%)
Diabetes	1410 (9.1%)	966 (7.2%)	444 (20.5%)
Treated for hypertension	3726 (23.9%)	2946 (22.0%)	780 (35.9%)

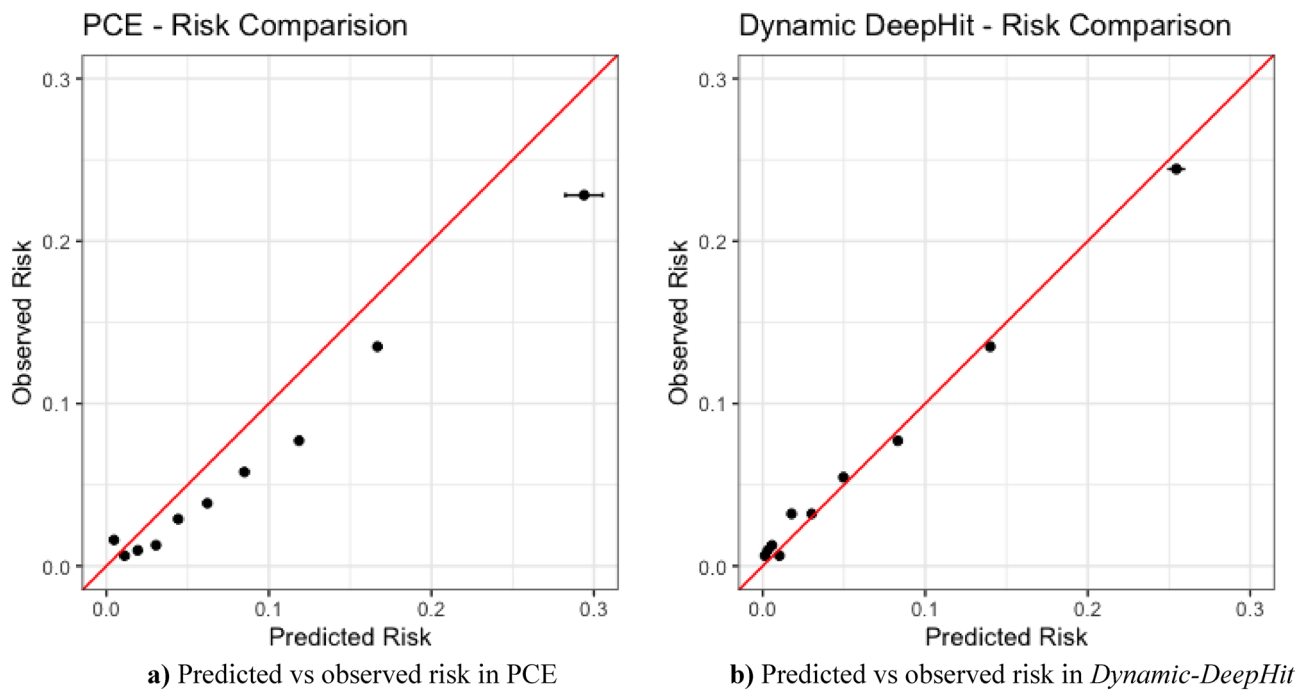
**Table 2.** Baseline demographics of LRPP analytic dataset. Risk factors in the ASCVD group were significantly higher than those in the Non-ASCVD group.

Categories	PCE training AUROC	<i>Dynamic-DeepHit</i> training AUROC	PCE testing AUROC	<i>Dynamic-DeepHit</i> testing AUROC
Full model	0.801 (0.786–0.816)	0.817 (0.791–0.843)	0.792 (0.760–0.825)	0.815 (0.782–0.844)
Race and sex				
Other males	N/A	N/A	0.779 (0.732–0.826)	0.801 (0.753–0.848)
Other females	N/A	N/A	0.780 (0.712–0.848)	0.801 (0.737–0.864)
Black males	N/A	N/A	0.826 (0.756–0.897)	0.820 (0.751–0.888)
Black females	N/A	N/A	0.801 (0.726–0.877)	0.821 (0.751–0.890)
Age				
< 60 years	N/A	N/A	0.781 (0.721–0.842)	0.803 (0.747–0.858)
≥ 60 years	N/A	N/A	0.667 (0.615–0.719)	0.698 (0.646–0.749)
PCE risk categories				
Borderline and intermediate	N/A	N/A	0.652 (0.594–0.709)	0.688 (0.634–0.742)

**Table 3.** Discrimination performance of models and in population subgroups.



**Figure 2.** Predicted risk (actual value) comparison between two models. Risk category derived from PCE risk category thresholds. horizontal dashed lines show the numeric thresholds for the risk categories used by the PCE mapped on to the risk predicted by the *Dynamic-DeepHit* model.



**Figure 3.** Observed events compared with average predicted risk within each predicted risk decile. Confidence intervals are also shown via error bars. (a) Predicted vs observed risk in PCE. (b) Predicted vs observed risk in *Dynamic-DeepHit*.

### Model performance in population subgroups

The discriminative performance of the *Dynamic-DeepHit* and PCE models within different population groups is shown in Table 3. The *Dynamic-DeepHit* model performed relatively better than the PCE in the other-males group (0.801, CI 0.753–0.848 vs 0.779, CI 0.732–0.826), other-females group (0.801, CI 0.737–0.764 vs 0.780, CI 0.712–0.848), and the Black-females group (0.821, CI 0.751–0.890 vs 0.801, CI 0.726–0.877). However, it underperformed in the Black-males group (0.820, CI 0.751–0.888 vs 0.826, CI 0.756–0.897). In the under-60-years-old group, the *Dynamic-DeepHit* model had an AUROC of 0.803 (CI 0.747–0.858) compared with the PCE's AUROC of 0.781 (CI 0.721–0.842) and in the over-60-years-old group, the *Dynamic-DeepHit* model had an AUROC of 0.698 (CI 0.646–0.749) compared with the PCE's AUROC of 0.667 (0.615–0.719). The *Dynamic-DeepHit* model outperforms the PCE in three of the four demographic groups outlined by the PCE.

### Feature importance

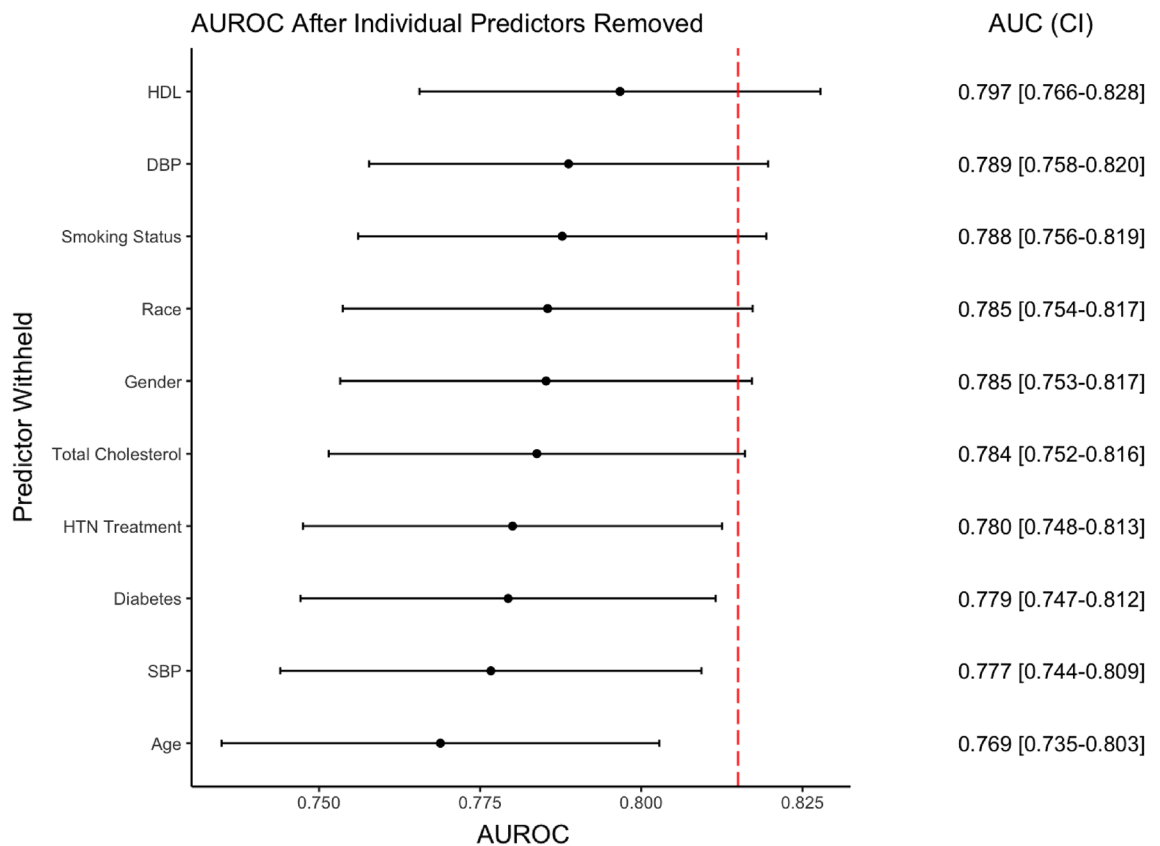
The results of the leave-one-out feature importance analysis are shown in Fig. 4. After removing age from the model, the greatest decrease in AUROC was observed (0.769, CI 0.735–0.803); thus, age is considered the most important variable in the model. Following age, longitudinal SBP was the second most important predictor, with the AUROC reduced to 0.777 (CI 0.744–0.809). Diabetes diagnosis and hypertension treatment were the most important categorical predictors, with AUROCs reduced to 0.779 (CI 0.747–0.812) and 0.780 (CI 0.748–0.813) when these predictors were removed respectively.

Figure 5 shows the longitudinal trajectories of clinical risk factors, including SBP, DBP, total cholesterol and HDL among the individuals whose risk increased and those whose risk decreased after switching to the *Dynamic-DeepHit* model for ASCVD risk prediction. Between the two groups, the average terminal measurements of SBP and total cholesterol were similar, the historical measurements of those risk factors were higher among those whose predicted risk increased in *Dynamic-DeepHit* model.

### Borderline risk stratification

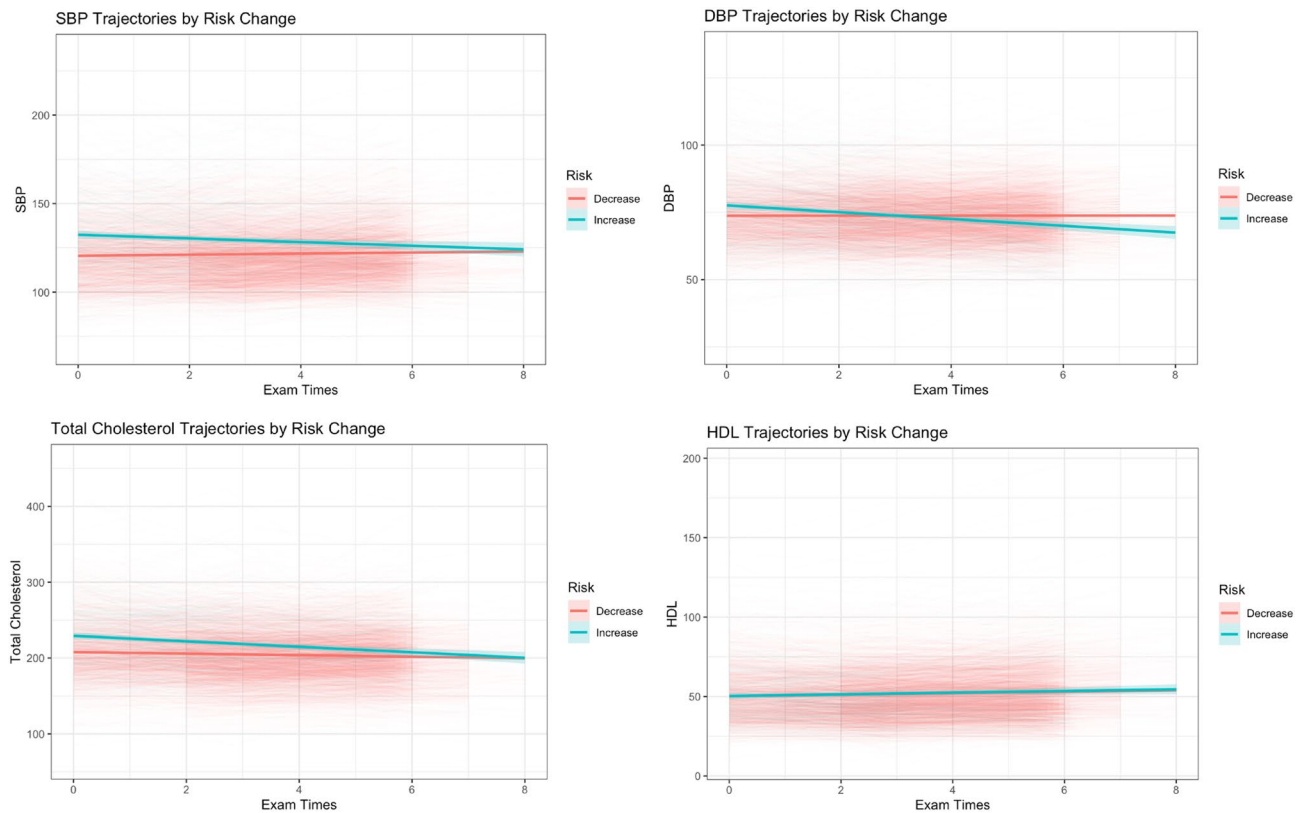
Among the individuals in the borderline and intermediate risk groups determined by the risk derived from the PCE, the AUC from the *Dynamic-DeepHit* model was higher than that from the PCE: 0.688 (CI 0.634–0.742) versus 0.652 (CI 0.594–0.709). The NRI for the *Dynamic-DeepHit* model between the borderline and intermediate group was 0.322. The Brier score was 0.069 for the PCE compared with 0.067 for the *Dynamic-DeepHit* model, again showing some improvement in the model calibration.

Given the 7.5% risk threshold for moderate-intensity statin prescription, we examined the individuals whose risk crossed the threshold in both directions to understand the *Dynamic-DeepHit* model's potential impact on



**Figure 4.** Feature Importance. AUROC was calculated via retraining and retesting the model while removing predictors individually. Red dashed line indicates the AUROC of the full *Dynamic-DeepHit* model.





**Figure 5.** Longitudinal histories of clinical risk factors compared between individuals whose risk increased and decreased in the *Dynamic-DeepHit* model from the PCE. Generalized linear models were used to smooth the longitudinal clinical risk factor trajectories. Color bands show the 95% confidence interval.

clinical decision making. In our testing dataset, among those who would be prescribed statins under the PCE risk ( $N = 1213$ ), 33% ( $N = 405$ ) would not be prescribed statins under the new risk provided by the *Dynamic-DeepHit* model, and 95% ( $N = 386$ ) of those individuals would not develop ASCVD. Among those who were not prescribed statins using the PCE ( $N = 1900$ ), 2% ( $N = 34$ ) would be recommended to prescribe statins under the *Dynamic-DeepHit* model. However, of those individuals, only 3% ( $N = 1$ ) developed ASCVD within 10 years.

## Discussion

### Principal findings

In this study, we have demonstrated that by incorporating longitudinal data of the same clinical and behavioral predictors as in the PCE using a state-of-the-art and validated deep learning model we can improve the calibration of predicting 10-year ASCVD risk. We leveraged data from 4 diverse cohorts for model training and testing and found that the longitudinal deep learning model outperformed the PCE both in the overall cohort and in specific subpopulations. We have demonstrated that the longitudinal deep learning model has clinical value through improved discrimination and greater calibration for those with borderline risk of ASCVD, thus providing physicians more reliable estimates of risk for clinical decision making.

### Deep learning in ASCVD risk prediction

Longitudinal trends of clinical factors such as blood pressure and cholesterol have long been established to be of clinical importance<sup>13</sup>. While this is not the first study to incorporate longitudinal data for predicting ASCVD, to our knowledge, it is the first study that uses a deep learning approach. Prior studies used methods such as including aggregate summary statistics of the longitudinal clinical data in the PCE or landmark models that could update data at fixed time intervals<sup>13,24,25</sup>. This foundational work led to minor improvements in model discrimination; however, we were able to achieve better performance because we utilized a deep learning method. A key advantage of deep learning models is their ability to recognize complex patterns by utilizing multiple layers of artificial neural networks, which are composed of inter-connected nodes. This advantage manifests in two ways in the *Dynamic-DeepHit* model. First, the improvement in the discrimination of the cross-sectional *DeepHit* model over the PCE demonstrates that given the same cross-sectional data, neural networks can make better predictions of ASCVD than the PCE. Second, the RNN can create robust representations of longitudinal clinical data, preserving critical information for ASCVD risk prediction.

## Clinical implications

Through evaluating the *Dynamic-DeepHit* model in various population demographic groups, we found that the model improves risk prediction in Black females compared with the PCE<sup>26–31</sup>. This indicates that incorporating longitudinal data may allow physicians to make more accurate treatment decisions and reducing health outcome disparities in these high-risk groups.

Among the individuals categorized as borderline- and intermediate-risk by the PCE, the *Dynamic-DeepHit* model improved discrimination and was better calibrated. One-third of the individuals in the intermediate PCE risk groups had overestimated 10-year ASCVD risk, which indicates the potential for over-prescribing. In these individuals, the *Dynamic-DeepHit* model slightly under-estimates risk, that it is better at ruling out people who will not have ASCVD events, while not as good as identifying those who will get ASCVD. As current clinical guidance requires further risk analysis for the individuals in these risk groups, guideline-concordant treatment is less optimal. By providing a better calibrated risk assessment, clinicians may be less concerned with over-prescribing and feel more confident in prescribing guideline-concordant treatment given the predicted risks from the *Dynamic-DeepHit* model.

The feature importance analysis shows that longitudinal measurements of clinical variables have meaningful influence on the performance on the *Dynamic-DeepHit* model. In the *Dynamic-DeepHit* model, longitudinal SBP was the most important modifiable predictor, while total cholesterol was found to be relatively important as well. Similar to prior research<sup>21</sup>, age was found to be an important predictor in the *Dynamic-DeepHit* model. In addition, diabetic status, sex, and smoking status were also found to influence the AUROC of the model. In the observed 8-year trajectories of SBP, DBP, and total cholesterol, for the individuals whose risk changed (Fig. 5), at the population level, the aggregate terminal measurements were similar. If prediction occurred only using those terminal measurements, a similar risk profile between those with increased risk and decreased risk would be assumed. However, the *Dynamic-DeepHit* model picked up separation in the historical values of those clinical factors, which contributed to the model identifying the differences in risk profiles of the two groups of individuals. Combined with the results of the feature importance analysis, this evidence further supports that longitudinal histories of clinical predictors can provide additional insight in evaluating ASCVD risk profiles.

With the proliferation of EHRs, longitudinal data is readily accessible. In addition, with the advent of cloud and edge computing, it is possible to deliver intensive computing capabilities to the EHR for supporting sophisticated machine learning or deep learning models for clinical risk prediction. This study shows, with further validation, deep learning models can be a powerful tool to aid clinicians to leverage the silos of currently untapped historical patient data in the EHR to improve patient cardiovascular outcomes. New methods of interpreting these models will also add confidence in adoption among physicians.

## Limitations

There are several limitations in our study. First, the cohorts used in this study may not reflect the clinical conditions of present-day patients, who are more likely to be on CVD treatments, such as statins. Therefore, given the limited information we had on statin usage, we did not exclude any participants who may have been on statin treatment. Second, data was recorded more sparsely in the cohort studies, whereas clinical measurements are often more frequent in clinical practice<sup>32</sup>. The quality of the data stored in the EHR could be also compromising, due to varying clinical contexts of when the data was collected. While these data problems exist, deep learning methods are still one of the best tools to overcome such issues<sup>33,34</sup>. On the other hand, EHRs often do not contain up to 8 years of longitudinal data on patients. As this study is a proof of concept, further work is needed to explore efficacy and utility of incorporating longitudinal risk factors into ASCVD risk prediction within EHRs.

## Data availability

Data used in this manuscript is not publicly available due to prior legal agreements. However, readers may reach out to the corresponding author to receive access to the pooled data source.

Received: 2 October 2023; Accepted: 8 January 2024

Published online: 31 January 2024

## References

1. Grundy, S. M. *et al.* 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: A report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J. Am. Coll. Cardiol.* **73**(24), e285–e350. <https://doi.org/10.1016/j.jacc.2018.11.003> (2019).
2. Stone, N. J. *et al.* 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J. Am. Coll. Cardiol.* **63**(25 Pt B), 2889–934. <https://doi.org/10.1016/j.jacc.2013.11.002> (2014).
3. Whelton, P. K. *et al.* 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: A report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Hypertension* **71**(6), e13–e115. <https://doi.org/10.1161/HYP.000000000000065> (2018).
4. Ridker, P. M. & Cook, N. R. Statins: New American guidelines for prevention of cardiovascular disease. *The Lancet* **382**(9907), 1762–1765. [https://doi.org/10.1016/S0140-6736\(13\)62388-0](https://doi.org/10.1016/S0140-6736(13)62388-0) (2013).
5. DeFilippis, A. P., Young, R. & Blaha, M. J. Calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann. Intern. Med.* **163**(1), 68–69. <https://doi.org/10.7326/L15-5105-2> (2015).
6. Wallisch, C. *et al.* Re-estimation improved the performance of two Framingham cardiovascular risk equations and the Pooled Cohort equations: A nationwide registry analysis. *Sci. Rep.* **10**(1), 8140. <https://doi.org/10.1038/s41598-020-64629-6> (2020).
7. Damen, J. A. *et al.* Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: A systematic review and meta-analysis. *BMC Med.* **17**(1), 109. <https://doi.org/10.1186/s12916-019-1340-7> (2019).



8. Lloyd-Jones, D. M. *et al.* Use of risk assessment tools to guide decision-making in the primary prevention of atherosclerotic cardiovascular disease: A special report from the American Heart Association and American College of Cardiology. *J. Am. Coll. Cardiol.* **73**(24), 3153–3167. <https://doi.org/10.1016/j.jacc.2018.11.005> (2019).
9. Wong, N. D. Cardiovascular risk assessment: The foundation of preventive cardiology. *Am. J. Prev. Cardiol.* **1**, 100008. <https://doi.org/10.1016/j.ajpc.2020.100008> (2020).
10. Domanski, M. J. *et al.* Time course of LDL cholesterol exposure and cardiovascular disease event risk. *J. Am. Coll. Cardiol.* **76**(13), 1507–1516. <https://doi.org/10.1016/j.jacc.2020.07.059> (2020).
11. Reges, O. *et al.* Association of cumulative systolic blood pressure with long-term risk of cardiovascular disease and healthy longevity: Findings from the lifetime risk pooling project cohorts. *Hypertension* **77**(2), 347–356. <https://doi.org/10.1161/hypertensionaha.120.15650> (2021).
12. Rospleszcz, S. *et al.* Temporal trends in cardiovascular risk factors and performance of the Framingham Risk Score and the Pooled Cohort Equations. *J. Epidemiol. Community Health* **73**(1), 19–25. <https://doi.org/10.1136/jech-2018-211102> (2019).
13. Pool, L. R., Ning, H., Wilkins, J., Lloyd-Jones, D. M. & Allen, N. B. Use of long-term cumulative blood pressure in cardiovascular risk prediction models. *JAMA Cardiol.* **3**(11), 1096–1100. <https://doi.org/10.1001/jamacardio.2018.2763> (2018).
14. Lewis, M. *et al.* Comparison of deep learning with traditional models to predict preventable acute care use and spending among heart failure patients. *Sci. Rep.* **11**(1), 1164. <https://doi.org/10.1038/s41598-020-80856-3> (2021).
15. Shickel, B., Tighe, P. J., Bihorac, A. & Rashidi, P. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE J. Biomed. Health Inform.* **22**(5), 1589–1604. <https://doi.org/10.1109/JBHI.2017.2767063> (2018).
16. Si, Y. *et al.* Deep representation learning of patient data from Electronic Health Records (EHR): A systematic review. *J. Biomed. Inform.* **115**, 103671–103671. <https://doi.org/10.1016/j.jbi.2020.103671> (2021).
17. Zhao, Y. *et al.* BERTSurv: BERT-Based Survival Models for Predicting Outcomes of Trauma Patients. [arXiv:2103.10928](https://arxiv.org/abs/2103.10928). Accessed March 01, 2021. <https://ui.adsabs.harvard.edu/abs/2021arXiv210310928Z> (2021).
18. Zeng, Z., Deng, Y., Li, X., Naumann, T. & Luo, Y. Natural language processing for EHR-based computational phenotyping. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **16**(1), 139–153. <https://doi.org/10.1109/TCBB.2018.2849968> (2019).
19. Wang, F., Casalino, L. P. & Khullar, D. Deep learning in medicine: Promise, progress, and challenges. *JAMA Intern. Med.* **179**(3), 293–294. <https://doi.org/10.1001/jamainternmed.2018.7117> (2019).
20. Wilkins, J. T. *et al.* Data resource profile: The cardiovascular disease lifetime risk pooling project. *Int. J. Epidemiol.* **44**(5), 1557–1564. <https://doi.org/10.1093/ije/dyv150> (2015).
21. Berry, J. D. *et al.* Lifetime risks of cardiovascular disease. *N. Engl. J. Med.* **366**(4), 321–329. <https://doi.org/10.1056/NEJMoa1012848> (2012).
22. Lee, C., Yoon, J. & Schaar, M. V. Dynamic-DeepHit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Trans. Biomed. Eng.* **67**(1), 122–133. <https://doi.org/10.1109/TBME.2019.2909027> (2020).
23. Van Calster, B. & Vickers, A. J. Calibration of risk prediction models: Impact on decision-analytic performance. *Med. Decis. Mak.* **35**(2), 162–169. <https://doi.org/10.1177/0272989x14547233> (2015).
24. Sayadi, M., Zare, N., Attar, A. & Ayatollahi, S. M. T. Improved landmark dynamic prediction model to assess cardiovascular disease risk in on-treatment blood pressure patients: A simulation study and post hoc analysis on SPRINT data. *Biomed. Res. Int.* **2020**, 2905167. <https://doi.org/10.1155/2020/2905167> (2020).
25. Paige, E. *et al.* Landmark models for optimizing the use of repeated measurements of risk factors in electronic health records to predict future disease risk. *Am. J. Epidemiol.* **187**(7), 1530–1538. <https://doi.org/10.1093/aje/kwy018> (2018).
26. Roger, V. L. *et al.* Heart disease and stroke statistics—2011 update: A report from the American Heart Association. *Circulation* **123**(4), e18–e209. <https://doi.org/10.1161/CIR.0b013e3182009701> (2011).
27. Ferdinand, K. C. *et al.* Disparities in hypertension and cardiovascular disease in blacks: The critical role of medication adherence. *J. Clin. Hypertens. (Greenwich)* **19**(10), 1015–1024. <https://doi.org/10.1111/jch.13089> (2017).
28. Jolly, S., Vittinghoff, E., Chattopadhyay, A. & Bibbins-Domingo, K. Higher cardiovascular disease prevalence and mortality among younger blacks compared to whites. *Am. J. Med.* **123**(9), 811–818. <https://doi.org/10.1016/j.amjmed.2010.04.020> (2010).
29. Mizuno, K. *et al.* Usefulness of pravastatin in primary prevention of cardiovascular events in women: Analysis of the Management of Elevated Cholesterol in the Primary Prevention Group of Adult Japanese (MEGA study). *Circulation (New York, NY)* **117**(4), 494–502. <https://doi.org/10.1161/CIRCULATIONAHA.106.671826> (2008).
30. Mosca, L., Barrett-Connor, E. & Wenger, N. K. Sex/gender differences in cardiovascular disease prevention what a difference a decade makes. *Circulation (New York, NY)* **124**(19), 2145–2154. <https://doi.org/10.1161/CIRCULATIONAHA.110.968792> (2011).
31. Ridker, P. M. *et al.* Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N. Engl. J. Med.* **359**(21), 2195–2207. <https://doi.org/10.1056/NEJMoa0807646> (2008).
32. Cohen, D. J. *et al.* Primary care practices' abilities and challenges in using electronic health record data for quality improvement. *Health Affairs Web Exclus.* **37**(4), 635–643. <https://doi.org/10.1377/hlthaff.2017.1254> (2018).
33. Rajkomar, A. *et al.* Scalable and accurate deep learning with electronic health records. *NPJ Digit. Med.* **1**, 18. <https://doi.org/10.1038/s41746-018-0029-1> (2018).
34. Xu, D., Hu, P. J., Huang, T. S., Fang, X. & Hsu, C. C. A deep learning-based, unsupervised method to impute missing values in electronic health records for improved patient management. *J. Biomed. Inform.* **111**, 103576. <https://doi.org/10.1016/j.jbi.2020.103576> (2020).

## Author contributions

J.Y., X.Y. contributed to the analysis of the project and writing of the manuscript. Y.D., A.K., L.P., L.Z., X.M., H.N., J.W., D.L.J. contributed to writing and editing of the manuscript. L.P., N.A. contributed to writing and editing of the manuscript, as well as the formulation of research project.

## Funding

The Lifetime Risk Pooling Project was supported in its inception by the National Institutes of Health/National Heart, Lung, and Blood Institute (R21 HL085375).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-51685-5>.

**Correspondence** and requests for materials should be addressed to N.B.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024