

Utility of Multimodal Longitudinal Imaging Data for Dynamic Prediction of Cardiovascular and Renal Disease: The CARDIA Study

Hieu Nguyen, PhD, Johns Hopkins University, Baltimore, MD, USA, hnguye78@jhmi.edu

Henrique D. Vasconcelos, MD, MSc, Johns Hopkins Hospital, Baltimore, MD, USA, hvascon1@jhu.edu

Kimberley Keck, Johns Hopkins Hospital, Baltimore, MD, USA, kkeck2@jhmi.edu

John Carr, MD, Vanderbilt University, Nashville, Tennessee, USA, j.jeffrey.carr@vumc.org

Lenore J Launer, PhD, National Institute on Aging, Baltimore, MD, USA, launerl@nia.nih.gov

Eliseo Guallar, MD, DrPH, Johns Hopkins School of Public Health, eguallar@jh.edu

João A.C. Lima, MD, Johns Hopkins Hospital, Baltimore, MD, USA, jlima@jhmi.edu

Bharath Ambale-Venkatesh, PhD, Johns Hopkins Hospital, Baltimore, MD, USA, bambale1@jhmi.edu

Short title: Multimodal Longitudinal Imaging & CVRD Prediction

Corresponding author:

Bharath Ambale-Venkatesh

Assistant Professor of Radiology and Radiological Science

MR 110, Radiology

Johns Hopkins University

Baltimore, MD 21287

Ph: 443-287-3248

bambale1@jhmi.edu

Abstract

Background: Medical examinations contain repeatedly measured data from multiple visits, including imaging variables collected from different modalities. However, the utility of such data for the prediction of time-to-event is unknown, and only a fraction of the data is typically used for risk prediction. We hypothesized that multimodal longitudinal imaging data could improve dynamic disease prognosis of cardiovascular and renal disease (CVRD).

Methods: In a multi-centered cohort of 5114 CARDIA participants, we included 166 longitudinal imaging variables from five imaging modalities: Echocardiography (Echo), Cardiac and Abdominal Computed Tomography (CT), Dual-Energy X-ray Absorptiometry (DEXA), Brain Magnetic Resonance Imaging (MRI) collected from young adulthood to mid-life over 30 years (1985-2016) to perform dynamic survival analysis of CVRD events using machine learning dynamic survival analysis (Dynamic-DeepHit, LTRCforest, and Extended Cox for Time-varying Covariates). Risk probabilities were continuously updated as new data were collected. Model performance was assessed using integrated AUC and C-index and compared to traditional risk factors.

Results: Longitudinal imaging data, even when being irregularly collected with high missing rates, improved CVRD dynamic prediction (0.03 in integrated AUC, up to 0.05 in C-index compared to traditional risk factors) from young adulthood followed up to midlife. Among imaging variables, Echo and CT variables contributed significantly to improved risk estimation. Echo measured in early adulthood predicted midlife CVRD risks almost as well as Echo measured 10-15 years later (0.01 C-index difference). The most recent CT exam provided the most accurate prediction for short-term risk estimation. Brain MRI markers provided additional information from cardiac Echo and CT variables that led to a slightly improved prediction.

Conclusions: Longitudinal multimodal imaging data readily collected from follow-up exams can improve CVRD dynamic prediction. Echocardiography measured early can provide a good long-term risk estimation, while CT/calcium scoring variables carry atherosclerotic signatures that benefit more immediate risk assessment starting in middle-age.

Keywords: multimodal, imaging, dynamic survival analysis, machine learning, dynamic prediction, cardiovascular disease, prognosis, CARDIA

Introduction

The rapidly expanding availability of large health data sets has fueled the growing research for more accurate risk prediction which holds much potential for preventive and monitoring strategies as well as improved disease understanding. In many scenarios, imaging data are collected over various modalities (multimodal) such as Echocardiography, Magnetic Resonance Imaging, and Computed Tomography, and repeatedly measured in multiple follow-up exams. Multimodal longitudinal imaging data could provide a more comprehensive description of the body and the development of organ functions and structures over time. In cardiology, numerous imaging markers for subclinical atherosclerosis have been demonstrated to be independently predictive of cardiovascular events^{1,2,3} and cardiac dysfunction.^{4,5} Many published works have focused on a few imaging variables that are low-dimensional, single-modal,^{2,4,6} or cross-sectional.⁷ The utility of high-dimensional, multimodal, and longitudinal imaging data has not been investigated.

Cox Proportional Hazards (Cox-PH) is among the most popular methods for survival analysis but Cox-PH is not suitable for high-dimensional data with repeated measures. The extended version of Cox-PH that can work with time-varying covariates is still limited because of the high number of variables, nonlinearity of variables, and requirement of data with no missingness.⁸ Machine learning (ML) approaches such as Random Survival Forest⁹ can mitigate some of Cox's limitations, but many ML methods are limited to static prediction and cannot perform dynamic survival analysis. In static prediction, the model does not automatically update as new observations are collected, (new data would require refitting an existing model or training a new model). Unlike static prediction, a dynamic survival analysis model automatically updates predicted survival probabilities as additional longitudinal observations are collected, and the model is trained only once. The ability to dynamically update risk as new information rolls in makes dynamic survival analysis attractive.^{10,11}

In this work, we demonstrated the utility of dynamic prediction of Cardiovascular and Renal Disease (CVRD) using high-dimensional, multimodal, longitudinal imaging. Data were collected in CARDIA, which is a large epidemiological study of Black and White young adults followed up over 30 years. We also identified the most important imaging predictors for CVRD in the CARDIA cohort.

Methods

Study Population and Outcome

The design of the CARDIA study (Coronary Artery Risk Development in Young Adults) has been described elsewhere.¹² Briefly, CARDIA is a prospective, observational cohort study of 5114 (originally 5115, one person withdrew consent) White and Black men and women aged 18 to 30 years, at four centers in the United States. The cohort is approximately balanced regarding age, race, sex, and educational level. Participants have been followed since 1985, with regular exam visits scheduled every 2-5 years. Each exam has collected a wide variety of variables believed to be related to heart

disease. The institutional review board of each participating institution approved the study protocol and all participants gave informed consent.

The outcome of this study is cardiovascular and renal disease (CVRD), and the first CVRD event was used as the endpoint. These events were adjudicated through August 2019. The primary composite outcome was incident cardiovascular disease and renal disease, which included coronary heart disease (CHD, myocardial infarction, acute coronary syndrome, or CHD death, including fatal myocardial infarction), stroke, transient ischemic attack, hospitalization for heart failure, intervention for peripheral arterial disease, end-stage renal disease, or death from cardiovascular or renal causes. Participants who died from a non-CVRD cause were censored at the time of death in the survival models.

Imaging Markers

CARDIA follow-up exams collect various imaging variables from different sources, such as Echocardiography (Echo), Computed Tomography (CT), Carotid Ultrasonography (CARTD), Dual-Energy X-ray Absorptiometry (DEXA), and Brain Magnetic Resonance Imaging (MRI). The extracted imaging variables have a high degree of sparsity and irregularity, reflecting real-world data. Echo was performed as part of the core study in Y5, Y25, and Y30 and as a substudy in Y10; CT was conducted in Y10 as a substudy and in Y15, Y20, and Y25 as the core study, and brain MRI was acquired in Y25 and Y30 as a substudy. Figure 1 shows an overview of imaging markers used in this study and Table S1 shows a detailed list of when the measures were collected. Data collection protocols for each imaging modality are available on the CARDIA study website.¹³ We used the longitudinal imaging CARDIA data from all exam years to develop prediction models.

Variables were pre-filtered with help from domain experts (clinicians who performed image reading daily). Other exclusion criteria include removing duplicated variables across modalities, variables available in too few subjects and variables with poor documentation. Overall, we included a total of 151 longitudinal imaging markers. In addition, we also included 15 traditional risk factors: nine variables from the AHA/ACC ASCVD risk scores and six additional risk factors (diastolic blood pressure - DBP, body-mass index - BMI, taking cholesterol-lowering medications, low-density lipoproteins - LDL, triglycerides, and fasting glucose).

Statistical Analysis

Model Training and Evaluation

Figure 1 shows the schematic of the statistical analysis procedures. All the models were trained and evaluated on the same cohort by 5-fold x 20 times cross-validation scheme. For each time the whole data was split, 20% of the data was used for testing, and the remaining 80% was further divided into training and validation sets. The training sets were used to fit the models, the validation sets were for hyperparameter tuning, and the testing sets were for assessing model performance. Stratified sampling by event was conducted to ensure the same ratio of events to non-events across the splits.

Modeling Methods

We used three algorithms to model data for dynamic survival analysis. To issue dynamically updated survival predictions the data requires methods that can incorporate high-dimensional, longitudinal data comprising various repeated measurements with varying degrees of missingness. The main algorithm we used is Dynamic-DeepHit.¹¹ Dynamic-DeepHit is a deep learning-based approach that issues dynamically updated survival predictions without making any assumptions about the underlying processes. Briefly, Dynamic-DeepHit consists of two subnetworks: a shared subnetwork with a recurrent neural network architecture that handles longitudinal measurements and predicts the next measurements of time-varying covariates, and a second subnetwork with cause-specific survival networks of fully connected layers that relates the longitudinal data to the survival outcome. Dynamic-DeepHit trains by minimizing the loss function which comprises three losses: a survival loss of log-likelihood of joint time-to-event distribution, a ranking loss that adapts the idea of concordance that encourages correct ordering of participants based on their time-to-event, and a step-ahead prediction loss that encourages correct prediction of longitudinal covariates for the next time step. Detailed description of Dynamic-DeepHit can be found elsewhere.¹¹

We employed two additional methods for dynamic prediction, namely Left-Truncated-Right-Censored Forest (LTRCforest)¹⁴ and Extended Cox for Time-dependent Covariates.⁸ Briefly, the Extended Cox is an extension of the fully parametric Cox-PH that assumes the variable values remain constant from the last observed value until updated. To handle a large number of input covariates, LASSO penalization¹⁵ was employed. LTRCforest is an extension of the non-parametric ML method conditional forest (Cforest) for time-varying covariates. Since LTRCforest and Extended Cox require fully available data, missing data were imputed before being input into these models using Multiple Imputation by Chained Equations (MICE) for multilevel data.^{16,17}

As a benchmark, we also fit static survival models at three time points (5 years, 15 years, and 25 years after baseline) to compare with the dynamic survival models.¹⁸ The idea is similar to landmarking approaches, in which a survival model is fit to the subjects who are still at risk at the landmarking time. For consistency in comparison, we used Dynamic-DeepHit and made an artificial cut-off at the landmarking time (meaning, covariate measurements after Y5, Y15, and Y25 were excluded from the static model at landmarking time 5 years, 15 years, and 25 years after baseline, respectively). Only measurements collected before the landmarking time were included.

Importance of imaging subsets and variables

To evaluate the effect of each imaging variable subset on CVRD prediction, we built five Dynamic-DeepHit models, each with traditional risk factors and imaging markers from a single modality. We also built a model with all imaging markers from all modalities and a reference model with only traditional risk factors. Additionally, to assess the complementary effects of multiple imaging subsets, we built 25 additional models representing all possible combinations of five imaging subsets. In total, 32 models were built (Table S2).

To examine the effect of imaging variables collected at different ages, we also built separate models for each imaging subset that included variables collected from each exam and excluded measurements collected outside of the exam.

Additionally, the importance of each imaging marker was quantified using permutation importance, similar to permutation testing,¹⁹ for the model trained on all variables. The longitudinal trajectories were permuted among participants, and the drop in the C-index of permuted variables to the C-index of the original dataset was used as the ranking criteria for variable importance. A bigger drop in C-index indicated a more important variable, and a minimal drop suggested that the variable was not important, as changing the variable value did not change model performance. Variables with the same difference in C-index were assigned the same ranking.

Performance Evaluation

Model performance was quantified using the time-dependent area under the receiver-operating curve (AUC) accounting for censorship²⁰ and the time-dependent concordance index that accounted for censoring distribution.²¹ In addition, the integrated AUC (iAUC) was used to quantify all time-varying AUCs as one number.²² Statistical significance was evaluated using Wilcoxon rank sum test.

Results

A total of 5114 participants were included in the analysis. Table 1 describes the characteristics and number of remaining participants in the cohort over nine follow-up exams. The mean age was 25 years old in CARDIA Y0 Exam (baseline) and 55 years old in the last exam (Y30). The cohort consisted of 46% male, 52% black, and 48% white. Over 30 years of follow-up, 3358 came back for Y30, the averaged SBP and DBP (systolic and diastolic blood pressure) increased and so was the use of hypertensive medication. BMI, total cholesterol, high-density lipoprotein (HDL), and use of cholesterol-lowering medication increased. The prevalence of diabetes also increased, while the number of smokers decreased. By the end of follow-up, 375 participants (7.3%) had developed CVRD. The cumulative incidence of CVRD is shown in Figure S1, with very few events happening before Y10 Exam, while the event rate curve is almost exponential after Y20 Exam.

Dynamic versus static prediction

Figure 2 shows the performance over time for dynamic prediction versus static prediction. The dynamic survival model using Dynamic-DeepHit trained on all 166 variables had a C-index of 0.80-0.82 before Y20 and slightly dropped to 0.78 by the last time point, 33 years after baseline. The C-index of the dynamic survival model is higher than that of the static survival models across all time points, by a margin of 0.01-0.06. Unlike the static survival models that required a separate model at each landmarking time, the dynamic model was only trained once and automatically updated survival probabilities as a new measurement updated from a follow-up exam.

Comparison of modeling methods

Figure S2 shows the performance of dynamic models trained with different algorithms (Dynamic-DeepHit, LTRCforest, Extended Cox, and Extended Cox penalized by LASSO). For both cases (trained on all variables and trained on traditional risk factors), Dynamic-DeepHit trained on unimputed data and LTRCforest trained on imputed data are consistently the

best, whereas Extended Cox consistently underperformed (~ 0.05 - 0.10 lower in C-index) for the model with all variables and 3-5% lower for model with only traditional risk factors).

Predictive gain of modalities

Figure 3 shows the predictive gain from each imaging subset over time and on average over 25 cross-validation folds. The best model was the one that included all imaging subsets, while the worst-performing was the one using only traditional risk factors (baseline) and using traditional risk factors plus CARTD variables (0.74 C-index at end of follow-up). Using all imaging markers resulted in up to a 5% increase in C-index and 3% in iAUC. The model utilizing only CT variables was only slightly below (1%) the model using all imaging variables, which helped elevate performance since Y10 Exam with a more apparent gain after Y20 Exam when more CT variables were collected. The model trained on Echo variables shows that the inclusion of Echo variables in addition to traditional risk factors helped increase prediction accuracy throughout the entire follow-up period by ~ 1.5 - 2% in C-index. DEXA variables improved performance very slightly up to Y25 by C-index (<0.01 absolute difference) and had negligible gain in terms of iAUC. Brain MRI variables, collected at Y25 and Y30 Exam, helped boost CRVD prediction performance by 0.01-0.02 C-index gain.

Table S2 shows the averaged iAUC gain from each of the 32 exhaustive combinations of 5 imaging variable subsets with respect to the baseline model. Aside from the best model using all imaging markers, the second-best subset is a combination of Echo, CT, and Brain MRI markers with an averaged iAUC gain of 0.027. A combination of Echo and CT variables resulted in a 0.022 iAUC gain, and the largest gain from a single imaging subset was from CT variables (0.014 iAUC gain).

Temporal Importance

Figure 4 quantifies the importance of early versus late measurements on CRVD prediction in two imaging subsets with the most influence on prediction: Echo and CT. For Echo, the model trained on only early Echo measurements (collected in Y5 Exam as a core study and partially in Y10 Exam as a substudy) had just as good overall performance (iAUC = 0.78) as the model trained on Echo variables collected later in life (Y20 and Y25 Core Exams) (iAUC=0.78). For longer-term risk prediction (25 years after Y0), the C-index of the model trained on early Echo was only 1% less than that of the model trained on late Echo. Regarding CT variables, which were collected in Y10 as a substudy and in Y15, Y20, and Y25 as core studies, the most recent CT exam provided the most accurate prediction, as evidenced by the immediate bump in the C-index after each CT Exam. The most prominent bump is right after CT Y25 which resulted in a 3% increase in C-index compared to using CT variables from Y20 or earlier ($p < 0.001$).

Variable importance

In addition to quantifying the importance of variable subsets relative to each other, variable-level importance was quantified. Table 2 shows the top 20 ranked variables at three representative times: 15 years, 25 years, and 33 years (endpoint) after Y0. Total cholesterol and low-density lipoprotein cholesterol (LDL) were consistently the most important predictors of CVRD. Most of the top 20 variables were either collected by Echo and CT, attesting to their importance to

CVRD prediction and consistent with the results in Figure 2. Chronological age is an important variable for year 15 (mean age=40) but not among the top 20 as participants got older. The top Echo and CT variables have similar relative variable importance in the rankings.

Discussion

In this work, we investigated the utility of high-dimensional longitudinal imaging data of five modalities, separately and together, for dynamic prediction of CVRD in young adults in a multi-centered cohort followed up over 30 years. We used the entirety of imaging variables over all exam years for continuously updated predictions of risk. The results suggest that longitudinal imaging data, even when irregularly collected and having high missing rates, improved CVRD dynamic prediction (3% iAUC, up to 5% C-index in midlife). Among different subsets of imaging markers, Echo and CT contributed to most of the improved risk estimation. Brain MRI variables contributed additional information that slightly improved prediction when they were collected. DEXA and Carotid IMT contributed little to none to CVRD prediction, even though they could be helpful in other aspects of clinical and epidemiological research. In terms of the effects of imaging markers measured early or late in life, the results suggested that Echo measured in early adulthood could predict long-term CVRD risks almost as well as Echo measured 10-15 years later. For CT, the most recent CT exam provides the most accurate prediction for short-term CVRD risk estimation. The results also suggest that the prediction ability of models decreased over time, particularly so between the ages of 40 and 50 years, when only traditional risk factors were included in the models. The addition of imaging variables helped maintain the prediction ability beyond middle age.

Multimodal Imaging Markers for Dynamic Prediction

This work is unique as it is among the first that incorporates high-dimensional longitudinal imaging markers from multiple modalities collected with high levels of sparsity (a high percentage of missing values) and irregularity (non-uniform time intervals between measurements) for dynamic prediction of CVRD. Many previous studies have limited the use of imaging data in prediction models, using only cross-sectional data or a few variables, or only including complete data. Simple imputation methods are often used to deal with sparsity and irregularity (mean/median imputation or last observation carried forward),²³ but these can introduce bias and do not fully capture the information in longitudinal data. In this work, we employed Dynamic-DeepHit which was capable of dealing with data of high sparsity and irregularity,¹¹ and thus could overcome the aforementioned challenges and better capture the rich information to improve risk estimation.

We showed that the inclusion of longitudinal multimodal imaging markers led to 0.03-0.05 increase in C-index and iAUC compared to not using imaging markers. It is worth noting that the imaging data collected in CARDIA was highly sparse and irregular (only available in two or three follow-up exams, collected in a small subset of participants (Table S1). The various missingness patterns reflected the nature of real-world data as not all information from a past patient visit will be

collected in the current visit. We argue that, despite the variable missingness rates, longitudinal multimodal imaging markers still improved prediction up to 5%. More complete and frequently collected data will likely yield greater improvement in risk estimation.

This study found that the predictive accuracy of a model for cardiovascular disease risk dropped after the mean age of the participants reached 45 years old, especially when using only traditional risk factors. The inclusion of multimodal imaging markers helped stabilize the predictive accuracy and prevented a decline of 6-7% over 13 years. The decline in predictive performance when using only traditional risk factors may be due to several factors. First, traditional risk factors may be less effective for predicting 10-year cardiovascular risk in people entering midlife despite their demonstrated usage for longer-term prediction. Previous studies from our group have shown that traditional risk factors were not among the top predictors for short-term prediction in an older population (MESA cohort, mean age = 62),⁷ and also in the CARDIA population.²⁴ Second, some non-traditional risk factors such as mental health, alcohol abuse, and other lifestyle factors, were not included in our prediction models. Studies have shown that cumulative effects of stress and alcohol contributed to worsening cardiovascular health.²⁵⁻²⁷ Third, health tends to decline starting in middle-age, when many changes occur in the body, making it more challenging to predict cardiovascular disease risk at this age. For example, at this age range, menopause often begins and the aortic root could enlarge and dilate, which have been shown to negatively affect cardiovascular functions and metabolism.^{28,28} More generally, metabolic syndrome in those 40–59 years of age were about three times as likely to happen as in those 20–39 years old.²⁹

In this regard, the decline in predictive performance in using only traditional factors further highlighted the role of multimodal imaging markers. Even though the traditional risk factors are fundamental to the genesis and progression of CVRD, multimodal imaging markers can pick up physiologic signals that are closer to disease initiation and closer to adverse outcomes. Furthermore, imaging markers can capture signals from some of the cumulative effects of insults to the body that were not captured by traditional risk factors. For example, coronary calcification from CAC/CT has demonstrated the proatherogenic effects of heavy alcohol consumption since young adulthood.³⁰ CAC/CT variables consistently ranked in the top predictors of outcome in our models (Table 2). Signals signifying the changes in the body at middle age could also be recognized by longitudinal imaging, for example, aortic root enlargement captured by Echo and was among the top-6 predictors of outcome at year 33, when the average participants' age was 58. In addition, the importance of age decreased over time while the importance of the imaging markers increased (Table 2), suggesting that vascular age captured by imaging may be more relevant than chronologic age. Overall, the included multimodal longitudinal imaging markers stabilized the decrease in prediction accuracy but may not have captured all relevant information. Adding more diverse, high-quality multimodal data may be necessary to further improve prediction in this age group.

Importance of imaging subsets

In this work, we quantified the importance of imaging markers as whole variable subsets/imaging modalities in addition to looking at variable-level importance. We also assessed spatial importance (in one exam) and temporal importance (across exams). We found that CT and Echo variables were consistently among the most important predictors. Specifically, within

Echo variables, left ventricular dimensions, ventricular septal thickness, aortic root measurements, and circumferential peak strain were among the most important. These variables are also reportedly among the top predictors in other large-scale studies.³¹ For CT variables, markers from coronary artery calcium (CAC) scans were consistently among the top predictors, adding to the growing evidence in the literature about the importance of CAC. Abdominal aortic calcium variables such as the number and size of lesions of the abdominal aorta and common iliac aorta are also among the top 30-50 predictors. Additionally, intermuscular adipose tissue (IMAT) measured by CT also consistently presented in the top 30, agreeing with previous reports showing IMAT associated with increased subclinical atherosclerosis independent of traditional cardiovascular disease risk factors and other adipose depots.⁵

In the early years, Echo markers, specifically markers of hypertension (such as septal thickness, LV volume and dimension), contributed the most to outcome prediction. However, in later years, CT markers played a larger role in prediction (Figures 3, 4, and Table 2). This suggests that at a young age, hypertension is the main driver of CVD, whereas at middle age, markers of atherosclerosis become the main driver, which can be more efficiently captured by CT/calcium scoring. Variables from the other subsets (e.g., DEXA, CARTD) contributed weakly to the prediction. Regarding brain MRI markers, total brain volume, including gray matter, white matter and cerebral spinal fluid and abnormal tissue volumes, primarily in white matter were among the top 15-20, and overall brain markers helped improve CVRD prediction in the immediate years after they were added. Previous studies have reported that cardiovascular risk burden is associated with cognitive decline, structural brain differences, and brain age.³²⁻³⁵ However, most studies show that CVD risk factors predict or are associated with brain structure and function,^{33,35,36} and not the other direction. Therefore, the brain MRI measures may reflect already accumulated CVD risk factors and therefore provide extra information on the severity of the risk factors further improving CVRD prediction.

Algorithmic consideration

In our study, we compared several dynamic survival analysis algorithms to identify the best technique to handle sparse and irregular imaging data. Among the techniques tested, machine learning methods were superior to the Extended Cox model. The best-performing models were those using Dynamic-DeepHit trained on unimputed data, which can directly handle sparse, high dimensional, and irregular data and provide true dynamic prediction. LTRCforest trained on imputed data performed on par with Dynamic-DeepHit but was not a true dynamic prediction algorithm and required imputation and more computational time. Therefore, Dynamic-DeepHit may be the most suitable algorithm for dynamic prediction.

Limitations

Our study has several limitations. The data collection started in 1985 in a biracial population and followed through for 30 years describing a certain cohort experience. Caution must be exercised when generalizing to other races and to the current population, as there may have been shifts in population characteristics over time. Second, external validation is challenging because long-term follow-up studies of young adults with extensive phenotyping like CARDIA are sparse. Third, as noted, many imaging markers in CARDIA are highly sparse and irregularly collected, whereas quantification of longitudinal multimodal imaging utility would improve with complete data. Despite that, the inclusion of sparse and

irregular multimodal imaging data still significantly improved prediction. Finally, the collection of repeated multi-modal imaging is practically possible mainly in well-resourced health facilities.

Conclusions

We show that longitudinal multimodal imaging data readily collected from follow-ups can improve CVRD dynamic prediction. Echocardiography measured early can capture hypertension status and provide a good prediction for long-term risk estimation, while CT/calcium scoring variables carry atherosclerotic signatures that benefit more immediate risk assessment starting in middle-age.

Acknowledgments

We thank the CARDIA committee for reviewing the scientific content of this manuscript. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

Sources of Funding

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201800005I & HHSN268201800007I), Northwestern University (HHSN268201800003I), University of Minnesota (HHSN268201800006I), and Kaiser Foundation Research Institute (HHSN268201800004I). Y25 CT Exam was funded by NHLBI grant R01-HL098445 to Vanderbilt University and Wake Forest University. This manuscript has been reviewed by CARDIA for scientific content.

Disclosures

The authors declare that they have no competing interests.

References

1. Peters SAE, den Ruijter HM, Bots ML, Moons KGM. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. *Heart*. 2012;98:177–184.
2. Nwabuo CC, Moreira HT, Vasconcellos HD, Mewton N, Opdahl A, Ogunyankin KO, Ambale-Venkatesh B, Schreiner PJ, Armstrong AAC, Lewis CE. Left ventricular global function index predicts incident heart failure and cardiovascular disease in young adults: the coronary artery risk development in young adults (CARDIA) study. *European Heart Journal-Cardiovascular Imaging*. 2019;20:533–540.
3. Armstrong AC, Jacobs DR, Gidding SS, Colangelo LA, Gjesdal O, Lewis CE, Bibbins-Domingo K, Sidney S, Schreiner PJ, Williams OD, et al. Framingham score and LV mass predict events in young adults: CARDIA study. *Int J Cardiol*. 2014;
4. Yared GS, Moreira HT, Ambale-Venkatesh B, Vasconcellos HD, Nwabuo CC, Ostovaneh MR, Reis JP, Lloyd-Jones DM, Schreiner PJ, Lewis CE. Coronary artery calcium from early adulthood to middle age and left ventricular structure and function. *Circ Cardiovasc Imaging*. 2019;12:e009228.
5. Terry JG, Shay CM, Schreiner PJ, Jacobs Jr DR, Sanchez OA, Reis JP, Goff Jr DC, Gidding SS, Steffen LM, Carr JJ. Intermuscular adipose tissue and subclinical coronary artery calcification in midlife: the CARDIA study (Coronary Artery Risk Development in Young Adults). *Arterioscler Thromb Vasc Biol*. 2017;37:2370–2378.
6. Ciuffo L, Nguyen H, Marques MD, Aronis KN, Sivasambu B, de Vasconcelos HD, Tao S, Spragg DD, Marine JE, Berger RD, et al. Periatrial Fat Quality Predicts Atrial Fibrillation Ablation Outcome. *Circ Cardiovasc Imaging*. 2019;12:e008764.
7. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. *Circ Res*. 2017;121:1092–1101.

8. Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. *Annu Rev Public Health*. 1999;20:145–157.
9. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Annals of Applied Statistics*. 2008;2:841–860.
10. Rizopoulos D, Molenberghs G, Lesaffre EMEH. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal*. 2017;59:1261–1276.
11. Lee C, Yoon J, van der Schaar M. Dynamic-DeepHit: A Deep Learning Approach for Dynamic Survival Analysis with Competing Risks based on Longitudinal Data. *IEEE Trans Biomed Eng*. 2019;XX:1–1.
12. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, Liu K, Savage PJ. Cardia: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41:1105–1116.
13. The Coronary Artery Risk Development in Young Adults (CARDIA) Study Homepage [Internet]. Available from: <https://www.cardia.dopm.uab.edu/>
14. Fu W, Simonoff JS. Survival trees for left-truncated and right-censored data, with application to time-varying covariate data. *Biostatistics*. 2017;18:352–369.
15. Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med*. 1997;
16. Resche-Rigon M, White IR. Multiple imputation by chained equations for systematically and sporadically missing multilevel data. *Stat Methods Med Res*. 2018;27:1634–1649.
17. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1–67.
18. Nguyen HT, Vasconcellos HD, Keck K, Reis JP, Lewis CE, Sidney S, Lloyd-Jones DM, Schreiner PJ, Guallar E, Wu CO. Multivariate longitudinal data for survival analysis of cardiovascular event prediction in young adults: insights from a comparative explainable study. *BMC Med Res Methodol*. 2023;23:23.
19. Ishwaran H. Variable importance in binary regression trees and forests. *Electron J Stat*. 2007;1:519–537.
20. Liang CJ, Heagerty PJ. A risk-based measure of time-varying prognostic discrimination for survival models. *Biometrics*. 2017;73:725–734.
21. Gerds TA, Kattan MW, Schumacher M, Yu C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Stat Med*. 2013;32:2173–2184.
22. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61:92–105.
23. Zhao J, Feng QP, Wu P, Lupu RA, Wilke RA, Wells QS, Denny JC, Wei WQ. Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Sci Rep* [Internet]. 2019;9:1–10. Available from: <http://dx.doi.org/10.1038/s41598-018-36745-x>
24. Nguyen HT, Venkatesh BA, Reis JP, Wu CO, Carr J, Nwabuo C, Gidding SS, Guallar E, Lima JAC. Lifetime vs 10-year Cardiovascular Disease Prediction in Young Adults Using Statistical Machine Learning and Deep Learning: The CARDIA Study. *medRxiv*. 2022;
25. Chaddha A, Robinson EA, Kline-Rogers E, Alexandris-Souphis T, Rubenfire M. Mental health and cardiovascular disease. *Am J Med*. 2016;129:1145–1148.
26. Albert MA, Durazo EM, Slopen N, Zaslavsky AM, Buring JE, Silva T, Chasman D, Williams DR. Cumulative psychological stress and cardiovascular disease risk in middle aged and older women: Rationale, design, and baseline characteristics. *Am Heart J*. 2017;192:1–12.

27. Kiechl S, Willeit J, Rungger G, Egger G, Oberhollenzer F, Bonora E. Alcohol consumption and atherosclerosis: what is the relation? Prospective results from the Bruneck Study. *Stroke*. 1998;29:900–907.
28. Rosano GMC, Vitale C, Marazzi G, Volterrani M. Menopause and cardiovascular disease: the evidence. *Climacteric*. 2007;10:19–24.
29. Ervin RB. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003–2006. 2009;
30. Pletcher MJ, Varosy P, Kiefe CI, Lewis CE, Sidney S, Hulley SB. Alcohol consumption, binge drinking, and early coronary calcification: findings from the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Epidemiol*. 2005;161:423–433.
31. Li Z, Yang Y, Zheng L, Sun G, Guo X, Sun Y. It's Time to Add Electrocardiography and Echocardiography to CVD Risk Prediction Models: Results from a Prospective Cohort Study. *Risk Manag Healthc Policy*. 2021;14:4657.
32. Song R, Xu H, Dintica CS, Pan K-Y, Qi X, Buchman AS, Bennett DA, Xu W. Associations between cardiovascular risk, structural brain changes, and cognitive decline. *J Am Coll Cardiol*. 2020;75:2525–2534.
33. Srinivasa RN, Rossetti HC, Gupta MK, Rosenberg RN, Weiner MF, Peshock RM, McColl RW, Hynan LS, Lucarelli RT, King KS. Cardiovascular risk factors associated with smaller brain volumes in regions identified as early predictors of cognitive decline. *Radiology*. 2016;278:198.
34. Kharabian Masouleh S, Beyer F, Lampe L, Loeffler M, Luck T, Riedel-Heller SG, Schroeter ML, Stumvoll M, Villringer A, Witte AV. Gray matter structural networks are associated with cardiovascular risk factors in healthy older adults. *Journal of Cerebral Blood Flow & Metabolism*. 2018;38:360–372.
35. Pase MP, Davis-Plourde K, Himali JJ, Satizabal CL, Aparicio H, Seshadri S, Beiser AS, DeCarli C. Vascular risk at younger ages most strongly associates with current and future brain volume. *Neurology*. 2018;91:e1479–e1486.
36. Armstrong AC, Muller M, Ambale-Ventakesh B, Halstead M, Kishi S, Bryan N, Sidney S, Correia LCL, Gidding SS, Launer LJ. Association of early left ventricular dysfunction with advanced magnetic resonance white matter and gray matter brain measures: The CARDIA study. *Echocardiography*. 2017;34:1617–1622.

Tables

Table 1. Characteristics of the study population over time

	Exam Y0	Exam Y2	Exam Y5	Exam Y7	Exam Y10	Exam Y15	Exam Y20	Exam Y25	Exam Y30
No. subjects examined	(N=5114)	(N=4624)	(N=4352)	(N=4085)	(N=3948)	(N=3671)	(N=3549)	(N=3499)	(N=3358)
Age	24.8 (3.63)	26.8 (3.60)	29.9 (3.60)	31.9 (3.58)	34.9 (3.61)	40.0 (3.60)	45.0 (3.59)	50.0 (3.59)	55.0 (3.55)
Sex (%Male)	2327 (46%)	2089 (45%)	1958 (45%)	1836 (45%)	1755 (44%)	1619 (44%)	1527 (43%)	1505 (43%)	1352 (43%)
Race									
Black	2651 (52%)	2300 (50%)	2132 (49%)	1987 (49%)	1938 (49%)	1741 (47%)	1649 (47%)	1632 (47%)	1510 (48%)
White	2463 (48%)	2323 (50%)	2219 (51%)	2098 (51%)	2010 (51%)	1930 (53%)	1884 (53%)	1842 (53%)	1639 (52%)
Systolic blood pressure (SBP)	110 (10.9)	108 (10.8)	108 (11.6)	109 (12.4)	110 (12.8)	113 (14.9)	117 (15.2)	120 (16.2)	121 (16.6)
Diastolic blood pressure (DBP)	68.6 (9.62)	67.4 (9.67)	69.2 (10.2)	69.3 (10.3)	72.4 (10.2)	74.5 (11.6)	73.1 (11.5)	74.9 (11.2)	74.0 (11.1)
Use of hypertensive medication	115 (2%)	41 (1%)	70 (2%)	81 (2%)	135 (3%)	292 (8%)	615 (17%)	936 (27%)	1033 (33%)
Body Mass Index (BMI)	24.5 (5.05)	25.2 (5.38)	26.2 (5.91)	26.8 (6.13)	27.5 (6.54)	28.8 (6.84)	29.5 (7.25)	30.2 (7.19)	30.5 (7.19)
Total cholesterol (mg/dL)	177 (33.5)	183 (35.2)	178 (34.3)	177 (34.3)	178 (34.6)	185 (35.8)	186 (34.9)	192 (36.9)	191 (38.1)
Use of cholesterol- lowering medication	0 (0%)	0 (0%)	11 (0%)	10 (0%)	19 (0%)	88 (2%)	311 (9%)	540 (16%)	632 (20%)
Total HDL cholesterol (mg/dl)	53.2 (13.2)	54.8 (14.1)	53.3 (14.2)	52.1 (14.2)	50.3 (14.0)	50.7 (14.6)	54.2 (16.7)	58.0 (18.0)	59.9 (19.0)
Total LDL cholesterol (mg/dl)	109 (31.3)	113 (33.3)	108 (32.1)	108 (31.6)	109 (32.1)	113 (32.3)	110 (32.0)	112 (32.8)	110 (33.2)
Triglycerides (mg/dl)	72.9 (48.5)	78.9 (53.4)	80.8 (72.2)	86.4 (75.7)	92.1 (74.7)	105 (92.8)	109 (79.9)	114 (85.9)	108 (99.5)
Fasting glucose (mg/100 ml)	82.6 (16.3)	NA (NA)	NA (NA)	90.1 (19.4)	88.2 (20.4)	86.7 (21.0)	98.0 (26.4)	99.5 (28.7)	102 (32.0)
Diabetes Mellitus	43 (1%)	55 (1%)	83 (2%)	142 (3%)	173 (4%)	212 (6%)	273 (8%)	373 (11%)	442 (14%)
Smoking now	1546 (30%)	1359 (29%)	1243 (29%)	1096 (27%)	1005 (25%)	808 (22%)	677 (19%)	588 (17%)	436 (14%)
CVRD event	0 (0%)	0 (0%)	2 (0%)	7 (0.2%)	14 (0.4%)	56 (1.5%)	112 (3%)	208 (6%)	313 (10%)

Table 2. Variable Importance Ranking (top 20)

Year 15				Year 25				Year 33			
Rank	RVI	Modality	Variable	Rank	RVI	Modality	Variable	Rank	RVI	Modality	Variable
1	1.00	Traditional	Total Cholesterol	1	1.00	Traditional	Total Cholesterol	1	1.00	Traditional	Total Cholesterol
2	0.94	Traditional	LDL	2	0.94	Traditional	LDL	2	0.91	Traditional	LDL
3	0.63	Traditional	Age	3	0.55	Echo	Circumferential Peak Strain (%)	3	0.74	Echo	Mitral Regurgitation (Moderate Or Severe)
4	0.59	Echo	Vent Septal Thickness - Systole	4	0.55	Echo	LV Internal Dim In Diastole	4	0.64	Echo	Tricuspid Regurgitation (RVsp > 40 Mmhg)
5	0.59	Echo	Vent Septal Thickness - Diastole	5	0.54	Echo	LV Internal Dim In Systole	5	0.63	Echo	Mitral Valve E-wave To A-wave Ratio
5	0.59	Echo	2D: LV Volume In Systole	6	0.51	Echo	Tricuspid Annular Peak Systolic Excursion	6	0.63	Echo	Aortic Root Dimension
6	0.58	Traditional	Smoking status	7	0.51	Echo	Aortic Root Dimension	7	0.62	Echo	Global LV Longitudinal Peak Strain (%)
6	0.58	CT	Agatston Score Left Anterior Descending	8	0.50	Echo	Vent Septal Thickness - Systole	8	0.61	Echo	LV Ejection Fraction
7	0.58	Echo	LV Internal Dim In Systole	9	0.49	Echo	LV Volume In Systole	9	0.61	MRI	Volume Of Total Brain
8	0.58	CT	Volume Of Lesions LAD	10	0.48	CT	Standard Deviation Attenuation Of Liver	10	0.61	Echo	LV Stroke Volume
9	0.58	Echo	Aortic Root Dimension	11	0.48	Echo	Vent Septal Thickness - Diastole	11	0.60	Echo	LV Internal Dim In Systole
10	0.58	Echo	AFVI/EFVI Ratio	12	0.48	Echo	LV Post Wall Thickness - Diastole	12	0.60	Echo	Circumferential Peak Strain (%)
10	0.58	CT	No. Of Lesions Left Anterior Descending	13	0.47	Echo	LV Internal Dimension Systole	13	0.60	Echo	LV Volume In Systole
11	0.58	DEXA	Hip: Total BMD	14	0.47	CT	Calcium Mass Of Lesions Infrarenal Abd. Aorta	14	0.59	MRI	Abnormal Tissue Volume In Gray Matter
12	0.58	CT	No. Of Lesions All Coronary	15	0.47	CT	Volume Of Lesions Right Coronary	15	0.59	Echo	Vent Septal Thickness - Diastole
12	0.58	Echo	LV Ejection Fraction	16	0.47	MRI	Volume Of Total Brain	16	0.59	CT	Stent Present
13	0.58	CT	Volume Of Lesions All Coronary	16	0.47	CT	Valve Replacement Present		0.59	CT	Calcium Mass Of Lesions LAD
14	0.58	CT	Mean Of Lesions Of Whole Heart	17	0.46	CT	No. Of Lesions Right Common Iliac	18	0.59	CT	Coronary Artery Bypass Graft
14	0.58	CT	Calcium Mass Of Lesions LAD	18	0.46	Echo	Doppler Tissue Doppler RV S-wave Velocity	19	0.59	CT	Cardiac Surgery Present
14	0.58	CT	Mean Of Lesions Left Circumflex	19	0.46	MRI	Abnormal Tissue Volume In Gray Matter	20	0.58	MRI	CSF Volume

RVI: relative variable importance; LDL: low-density lipoprotein; LV: left ventricular; RV: right ventricular; AFVI/EFVI: atrial to early diastolic flow velocity integral; BMD: bone mass density; LAD: left anterior descending artery; CSF: cerebrospinal fluid.

Figures with Figure Legends

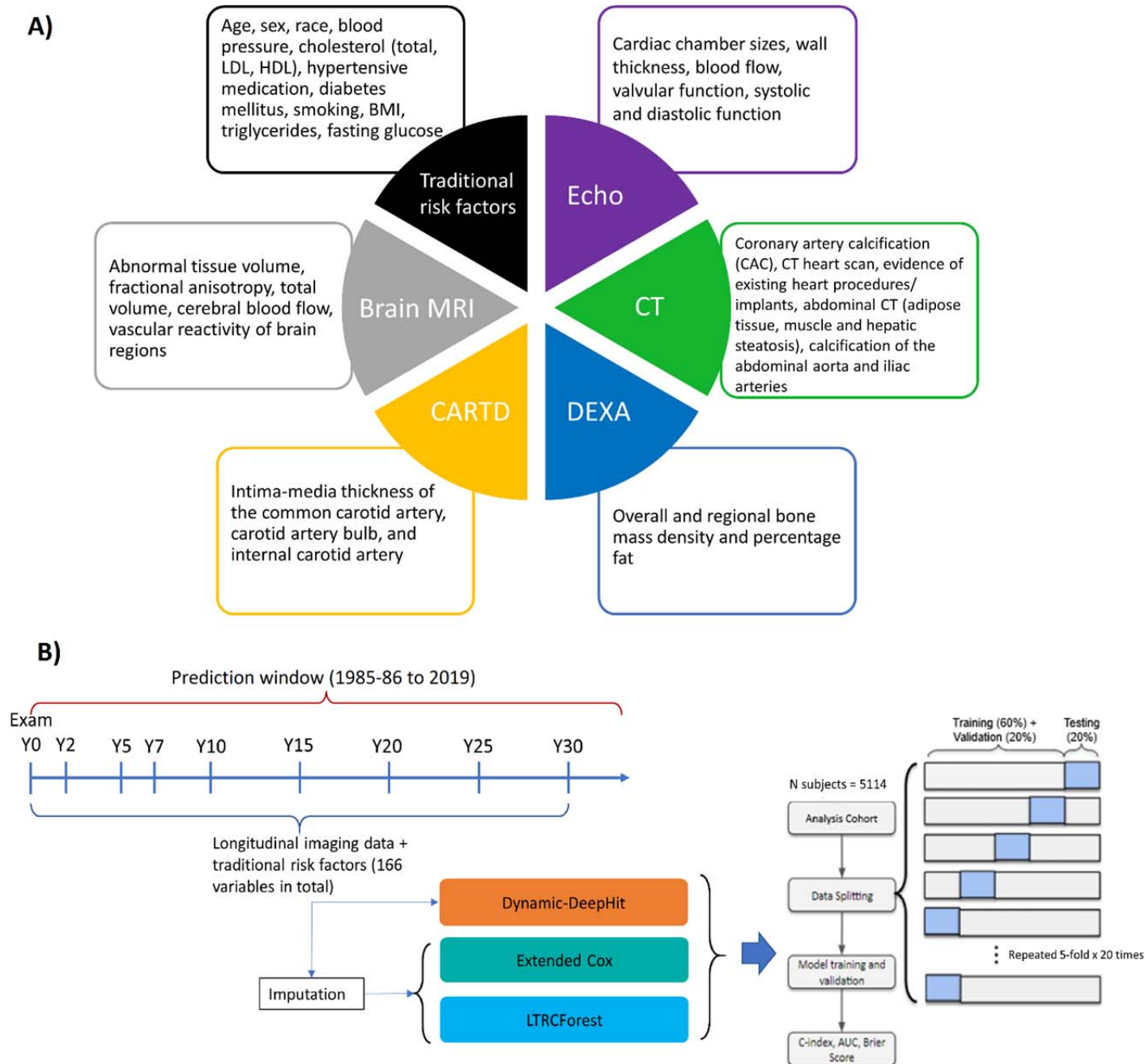


Figure 1. A) Overview of the multimodal data used for prediction. B) Methodology framework visualization for dynamic survival analysis. Echo: echocardiography, CT: Computed Tomography, CARTD: Carotid Artery Ultrasonography, DEXA: Dual-Energy X-ray Absorptiometry, MRI: Magnetic Resonance Imaging.

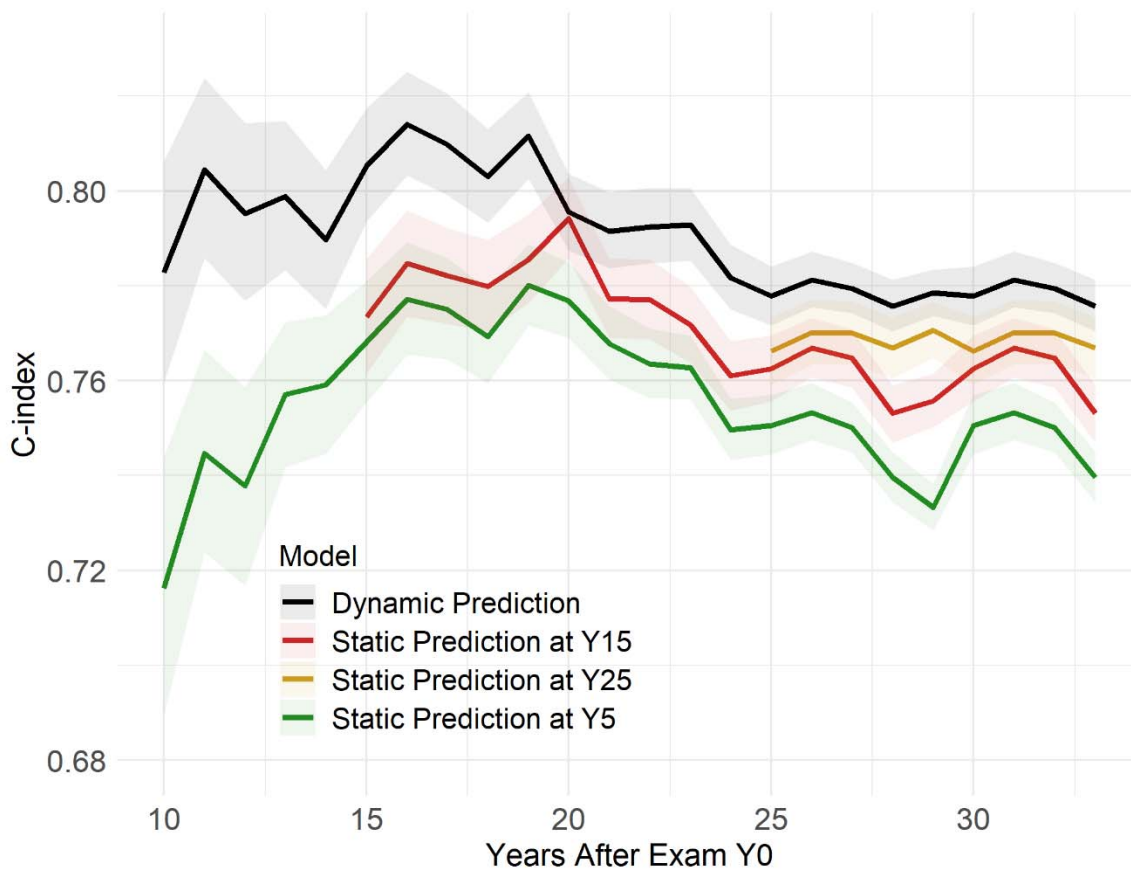


Figure 2. Dynamic vs. static prediction

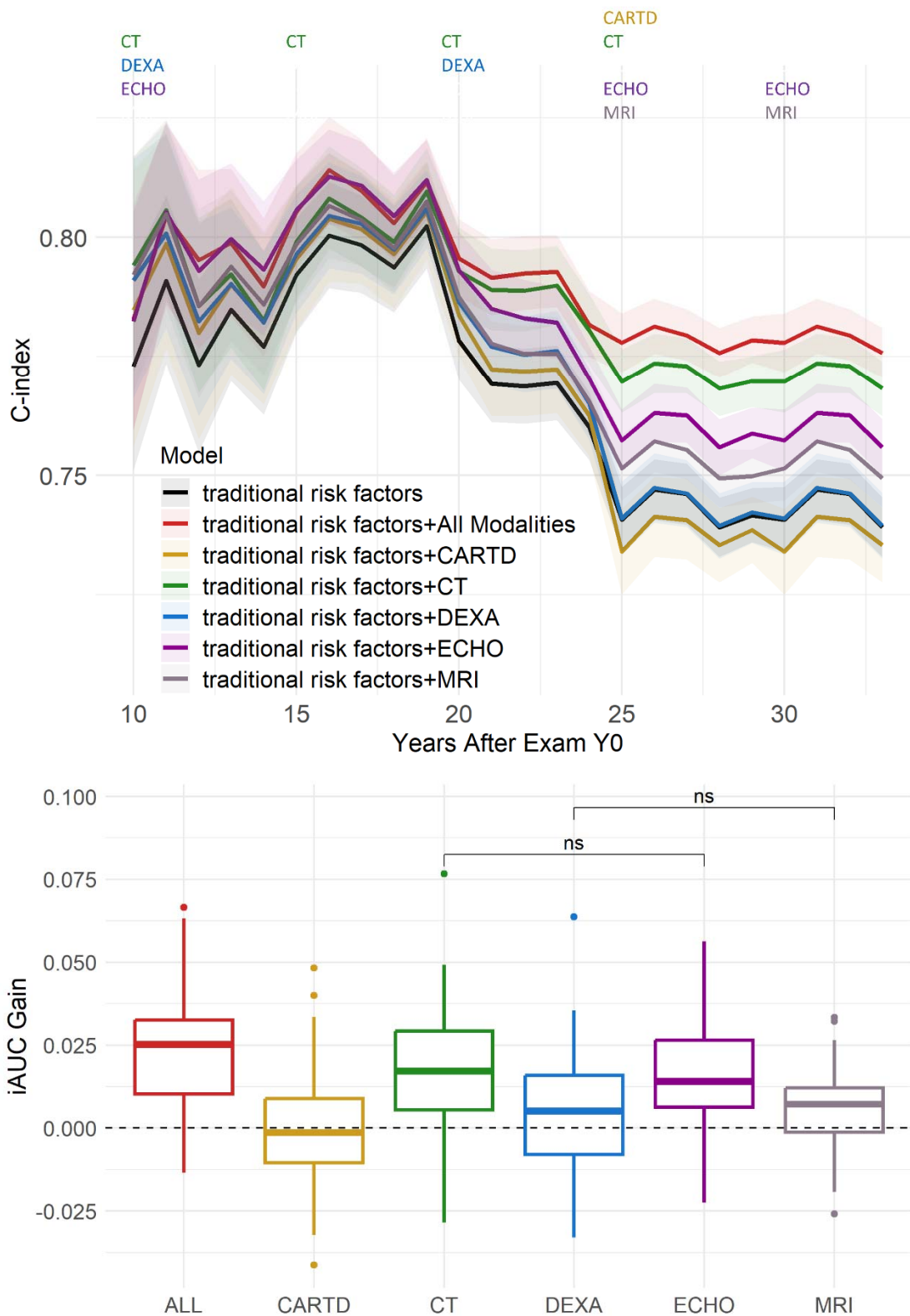


Figure 3. Predictive Gain from Imaging Variables of Different Modalities. Top: performance over time. The colored texts at the top indicate which imaging modalities were collected at each exam. Bottom: integrated AUC (iAUC) Gain with respect to the baseline model trained on only traditional risk factors. All pairwise hypothesis tests are significant (including All vs. ECHO and All vs. CT), unless otherwise denoted with 'ns' ('non-significant').

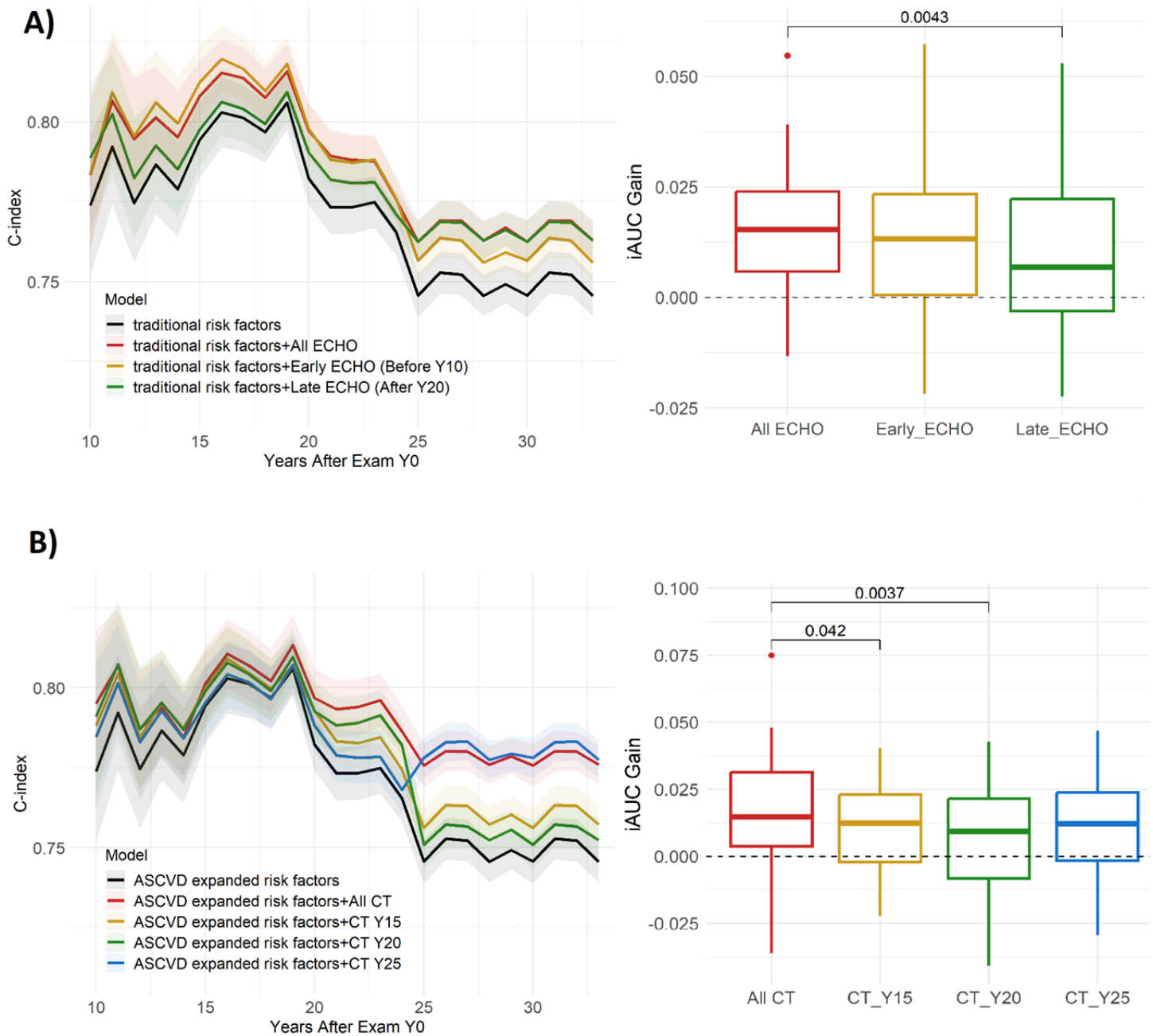


Figure 4. Effects of early vs. late imaging measurement. A) Early vs. Late Echo, B) Early vs. Late CT. Early Echo provided good overall gain for long-term risk estimation, compared to Late Echo. For CT, the most recent CT exam provides the most accurate prediction, evident by immediate bumps after each CT Exam, especially CT Y25. P-values of significant pairwise hypothesis tests are shown.