**Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and anticipated methodologic limitations or challenges).**

***1. Methods/Study Design***

A cross-sectional predictive model comparing predictors with robust, pre-frail and frail participants developed from the baseline InCHIANTI data will be used to replicate a population predictive model in ARIC. The ARIC model is run separately with a comparison of the AUC. The results discuss the attributable AUC differences in part based on outcome derivations. The top predictive biological features for the pre-frail and frail model 1 from the InCHIANTI study with 2 rounds 10 iterations resulting in a AUC 0.9387 (95% CI 0.89-0.98). Using data from the Atherosclerosis Risk in Communities Study (ARIC), predictors from Model 1 developed in the InCHIANTI data is used to replicate the frailty model. Figure 1 represents the workflow process that is used to replicate the frailty prediction model in the ARIC dataset.

Figure 1. study approach and workflow diagram



Note. Profile of model development and validation workflow. Blue boxes indicate steps of the workflow specific to the ARIC dataset.

1.1. Model generation

The predictive clinical and laboratory biomarkers were identified in InCHIANTI Model 1 is analyzed using an Extreme Gradient Boosting (xgboost) in R1 for the validation model in ARIC. While boosting was initially developed for machine learning, ‘xgboost’ in R is based in boosted trees. Xgboost is an open source tool and a variant of the gradient boosting machine and uses a tree based model. Xgboost is used in this study for a supervised learning problem where the variables identified from a previous study are used to predict pre-frail and frail individuals.2

1.2. Evaluation of the model

With the use of any predictive model in machine learning there is a chance for inflated risk of capitalizing on chance features (overfitting) in the data. Overfitting of the model is mitigated in two ways: 1) having a distinct training and validation process for the model and 2) using xgb in R which has built-in parameter settings for selection to reduce poor predictive performance. *Internal validation:* A randomly assigned training subset is be used to validate the model within the ARIC cohort *in silico* (via simulation).

1.3. Calibration of the model

Parameter estimates for each predictive factor and associated descriptive statistics will be evaluated to provide biological insight into the underpinnings of the classification algorithm. We evaluated the calibration by partitioning the data into 5, 10, 20, 30, 40, 50, 75, 100 and 200 groups and then run the calibration test. Next, we repeated tests for all possible values between 5-200 groups and evaluated the distribution of the test statistic. The best prediction thresholds were determined using AUC.

1.4. Phenotype

The frailty phenotype defined in three categories—non-frail (0), pre-frail(1-2), and frail (3-5).3,4 Outcome measure of frailty is used from V5 in the ARIC dataset.

1.5. Predictors

1.5.1. Anticholinergic Burden Calculation

The Anticholinergic Cognitive Burden (ACB) scale is the most clinically validated scale for evaluating adverse health outcomes including cognitive and physical function5,6. The anticholinergic properties of each medication will be quantified using the ACB scale based on each drug’s serum anticholinergic activity7. ACB scores for each participants’ medications will be assigned points (0, 1, 2, 3) according to the published 2012 update and summed for a total anticholinergic burden score. Higher scores indicate higher anticholinergic properties. An example of medications with ACB scores include: Amitriptyline = 3, Amantadine = 2, and Atenolol = 1.

1.5.2. Depression Score

The CES-D self-report scale (0-60) is used to measure depressive symptoms. Reliability, validity, and factor structure have been similar across a diverse demographic and the scale has been used extensively in epidemiologic studies for depression and physical function.8 The CES-D score will be used in the predictive model.

1.5.3 Demographics

Age – used as a continuous variable, race/ethnicity, and education level

1.5.4. Table 3. represents the validated biological and clinical markers used to predict frailty in the InCHIANTI data. Table II. Describes the clinical and biological markers available for the prediction of frailty in the ARIC dataset. In ARIC, 3 biological markers were used as substitutes (serum TNF used instead of TNF-a receptor I&II, total testosterone used instead of Free testosterone, and serum creatinine used instead of Creatinine clearance, 24-hr urine/24-hour urinary creatinine). InCHAINTI model has a total of 19 frail clinical and biological markers and 13 total frail markers are available in ARIC. InCHAINTI model has a total of 22 pre-frail clinical and biological markers and 16 total prefrail markers are available in ARIC.

Table 1. InCHIANTI and ARIC participants characteristics

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | InCHIANTI | Robust | Pre-frail | Frail | ARIC | Robust | Pre-frail | Frail |
| Phenotype (n) | | 507 | 434 | 85 |  | 3025 | 3050 | 433 |
| Age,  mean(SE) | | 72(0.30) | 77(0.32) | 81(0.65) |  | 75(4.8) | 77(5.4) | 79(5.6) |
| Gender, %  Male (n)  Female (n) | | 51.5(261)  48.5(246) | 37.1(161)  62.9(273) | 37.6(32)  62.4(53) |  | 44.3(1339)  55.7(1686) | 39.3(1199)  60.7(1851) | 33.5(145)  66.5(288) |
| Race, %  Black or African American\*  White\*  American Indian or Alaskan Indian  Asian | |  |  |  |  | 21.12(639)  78.7(2380)  0.10(3)  0.10(3) | 25.2(769)  74.4(2269)  0.10(3)  0.30(9) | 27.3(118)  72.8(315)  0(0)  0(0) |
| Education, %   |  | | --- | | No Education | | Elementary Secondary | | ≧ High School | | |  |  |  |  |  |  |  |

Table 3. Biological and Clinical Predictors: InCHIANTI and ARIC (supplemental table)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **InCHIANTI** n=1,026  Baseline Data | **ARIC** | | | **Frailty**  **Feature** |
| **Clinical** | Variable | Variable | n | Time Point |  |
| Age | X | X | 6,508 | V5 | Frail/Prefrail |
| Anticholinergic Burden | X | X | 6,508 | V5 | Frail/Prefrail |
| Depression/CES-D self-report scale | X | X | 6,508 | V5 | Frail/Prefrail |
| **Inflammatory/Immunity** |  |  |  |  |  |
| 24-hour urinary cortisol (µg/24 hours) | X | NA | NA | NA | Prefrail |
| Erythrocyte sedimentation rate (ESR) (mm/hour) | X | NA | NA | NA | Frail/ Prefrail |
| Homocysteine via FPIA analysis (Âµmol/L) | X | X | 328 | V1 | Frail/ Prefrail |
| Interleukin-1B via ELISA (pg/mL) | X | X | 442 | V1 | Prefrail |
| Interleukin-6 via ELISA ultrasensitive (pg/mL) | X | X | 572 | V1 | Frail/ Prefrail |
| Monocyte chemoattractant protein-1 via Bio-Plex (pg/mL) | X | X | 604 | V1 | Prefrail |
| Soluble TNF-a receptor I via quantitative sandwich EIA (pg/mL) | X | NA\*  (serum TNF) | 170 | V4 | Frail/ Prefrail |
| Soluble TNF-a receptor II via quantitative sandwich EIA (pg/mL) | X | NA\*  (serum TNF) | 170 | V4 | Frail/ Prefrail |
| **Hematology/Liver** |  |  |  |  |  |
| Folate via RIA (ng/mL) | X | X | 6,166 | V3 | Frail/ Prefrail |
| Mean corpuscular volume (MCV) (fL) | X | X | 6,281 | V5 | Prefrail |
| Retinol via high performance liquid chromatography (µmol/L) | X | NA | NA | NA | Prefrail |
| GPT (also known as ALT) (U/L) | X | X | 5,997 | V4 | Frail |
| **Endocrine/Hormones** |  |  |  |  |  |
| 25(OH)-D (25-hydroxyvitamin D) via RIA (nmol/L) | X | X | 6,026 | V3 | Frail/ Prefrail |
| Free testosterone (ng/dL), Vermeulen | X | NA\*  (total testosterone ng/dL) | 5,555 | V4 | Prefrail |
| Blood glucose (mg/dL) | X | X | 6,108 | V5 | Frail |
| Free thyroxine, fT4 (ng/dL) | X | X | 6,409 | V5 | Frail |
| Parathyroid hormone, two-site immunoradiometric assay (pg/mL) | X | X | 6,054 | V2 | Frail |
| **Metabolomics(plasma lipids)** |  |  |  |  |  |
| Fatty acid C24:0 weight (mg/L) | X | NA | NA | NA | Prefrail |
| Lipids: HDL cholesterol (mg/dL) | X | X | 6,404 | V5 | Frail |
| **Renal/Electrolyte** |  |  |  |  |  |
| Creatine phosphokinase (U/L) | X | NA | NA | NA | Prefrail |
| Creatinine clearance, 24-hr urine (mL/minute) | X | NA\*  (urine creatinine) | 6,317 | V5 | Prefrail |
| Urine proteins (mg/dL) | X | NA | NA | NA | Frail/ Prefrail |
| 24-hour urinary creatinine (mg/24 hours) | X | NA\*  (urine creatinine) | 6,317 | V5 | Frail |
| Blood urea nitrogen (mg/dL) | X | X | 6,461 | V1 | Frail |
| **Nutrient Biomarker** |  |  |  |  |  |
| Vitamin B6 via high performance liquid chromatography (ng/mL) | X | X | 6,166 | V3 | Frail/ Prefrail |
| Vitamin E gamma tocopherol, high performance liquid chromatography (Âµmol/L) | X | X | 6,166 | V3 | Prefrail |
| Lycopene via high performance liquid chromatography (Âµmol/L) | X | NA | NA | NA | Frail |

\*substitute biomarker was used because exact biomarker match was not available, red notes >15% missing removed from model

**1.6 Limitations**

1. Several of the biomarker measurements come from different time points than the outcome measure of frailty at V5. The model is built with data as close to the outcome diagnosis (V5 Frailty) as possible. Additionally, we examine model parameters and AUC adding variables from visits 1 through 5. As variables are added, parameters (model fit and AUC) was examined for best fit. Initial rebuild in InCHIANTI with the predictors from the ARIC model maintained an AUC of .897. Although we cannot completely control for the varying temporal differences between some predictors and the frailty outcome, we find the findings to be informative and useful for future funding opportunities.

2. Frailty measure differs slightly between databases. Often there are variations in how frailty is measured between longitudinal studies which may affect the AUC and the ability of the InCHIANTI model to “fit” or accurately predicted frailty in the ARIC data. However, frailty in ARIC was based on the same frailty phenotype that is used in InCHIANTI and has been validated.

3. Demographics and ethic/racial differences between databases. The InCHIANTI databases is a geographically homogeneous population with a large white/causation European population. We consider it a strength rather than a limitation that ARIC and Health ABC are more diverse populations and will allow external validation to be conducted in different cohorts. Models explored ethic/racial differences in the model fit as the model was being built.