Accelerating Change and Transformation in Organizations and Networks

Developing and Assessing the Validity of Claims-based Indicators of Frailty & Functional Disabilities in Electronic Health Records

FINAL REPORT

Prepared by:

Hadi Kharrazi, M.D., Ph.D., FAMIA
Johns Hopkins Center for Population Health IT
Johns Hopkins School of Public Health
Armstrong Institute for Patient Safety and Quality
Johns Hopkins School of Medicine

Douglas W. Roblin, Ph.D.
Senior Research Scientist
Kaiser Permanente Mid-Atlantic States

Jonathan P Weiner, Dr.P.H.
Johns Hopkins Center for Population Health IT
Johns Hopkins School of Public Health
Armstrong Institute for Patient Safety and Quality
Johns Hopkins School of Medicine

Disclaimer of Conflict of Interest

None of the investigators has any affiliations or financial involvement that conflict with the material presented in this report.

Funding Statement

Public Domain Notice

Suggested Citation

Developing and Assessing the Validity of Claims-based Indicators of Frailty & Functional Disabilities in Electronic Health Records. Prepared under Contract No. 75Q80120D00015. AHRQ Publication No. 22-0058. Rockville, MD: Agency for Healthcare Research and Quality; September 2022.

Acknowledgments

The authors thank and recognize the following individuals and groups, without whom this work could not have been accomplished:

Executive Summary

Frailty is conceptualized as decreased physiologic reserve and inability to withstand physical and psychological stressors. The goal of frailty measurement is to identify high-risk older adults and to target interventions to prevent adverse health outcomes. Despite the utility of frailty in identifying older adults at risk, and an abundance of frailty measures in the literature, healthcare providers continue to lack pragmatic tools to cost-effectively screen large patient populations for frailty. Screening tools for frailty may identify individuals in need of further evaluation at the point of care, but such tools still require availability of or collection of new data that is specific to the score (e.g., gait speed, chair rise, grip strength) and cannot be automatically calculated from information already in a patient's chart.

Healthcare providers and health insurance plans are actively seeking ways to measure frailty using insurance claims, electronic health records (EHRs), and on a more limited scale, health risk assessments. Applying and scaling frailty indexes across adult populations enable providers and plans to identify frail individuals at high risk for mortality, disability, and healthcare utilization. Multiple claims-based frailty indexes (CFIs) have been developed and validated over the past few years; however, healthcare providers often do not have access to the insurance claims records of their entire population of patients, thus necessitating the development of reliable EHR-based frailty indexes (EFI). Nonetheless, a challenge with developing EFI measures is the lack of frailty variables captured as structured codes within EHRs.

To address the operational gap between CFIs and EFIs, this project focused on validating an established CFI using linked claims-EHR databases of multiple large health systems: Johns Hopkins Medical Institute (JHMI); Optum Labs Data Warehouse (OLDW), which includes data from 55 health systems; and Kaiser Permanente Mid-Atlantic States (KPMAS). Task 2 of this project assessed and compared the EHR and claims data of these data sources to ensure sufficient data quality for frailty analysis. Task 3 of the project compared the EFI and CFI using EHR and claims data of each data source. Tasks 1 and 4 focused on administrative and dissemination efforts (e.g., data use agreements, scientific publications) and are not covered in this report.

The project provides a systematic approach to healthcare providers to examine the quality of the EHR data and prepare it for the application of EFI measures (Task 2). The EFI showed to be a valid measure of frailty when compared to a custom patient survey at KPMAS, and when compared to CFI measures of the same population across all data sources (subtasks 3.1 & 3.2). An acceptable concordance of EFI and CFI was found and shown to be stable across multiple health systems (subtask 3.3). The concordance of EFI and CFI was also acceptable across different patient groupings such as age, sex, and race (subtask 3.4). Finally, the EFI were found to be predictive of current and future healthcare utilization outcomes, such as inpatient hospitalization, emergency department admission, and nursing home admission (subtask 3.5).

In conclusion, the project findings demonstrated that structured EHR data can be used by healthcare providers to identify frail patients using validated EFIs; however, claims data can identify additional frailty cases compared to EHR data. Further research is needed to evaluate the role of unique EHR features, such as unstructured data in physician notes, in developing EFIs that have a higher sensitivity and specificity in identifying patients with frailty.

EFIs can also be used to improve the prediction of various healthcare utilization outcomes. Risk stratification developers may integrate EFI in their model development process, and population health managers may incorporate EFI in disease management efforts. Future studies should evaluate the interaction of comorbidity indexes with EFIs in predicting healthcare utilization outcomes and adjusting total healthcare costs.

Contents

Figures

Tables

1. Introduction

Frailty and Frailty Measures

Frailty is an age-derived vulnerability that affects 8% to 25% of older adults [1]. Frailty decreases the ability of older adults to maintain normal physiological status after a stressful event [2,3]; thus, escalating the risk of adverse health outcomes and potentially increasing utilization and cost [4-8]. Traditionally, frailty is measured by well-validated survey instruments assessing either the frailty phenotype [3] or the deficit accumulation frailty index (FI) [5]. Despite the advantages of the frailty phenotype in clinical settings [3], the FI has shown to be more useful for measuring the total burden of health among older adults [6-8].

Survey-based frailty indexes are useful for targeted care at the point of care; however, survey-based frailty indexes are often impractical to collect across large populations of older adults. Consequently, researchers have developed and validated several frailty indexes using more widely available clinical data sources such as insurance claims. Multiple claims-based frailty indexes (CFIs) have been developed and validated over the years [9-16]. Some CFIs have adopted the frailty phenotype methodology [9-14], while others have used the deficit accumulation approach [15,16]. CFIs have shown value in predicting various health outcomes such as disability, mobility impairment, and hospital days [15,16]. Due to their predictive value, insurers have shown interest in leveraging CFIs to improve care coordination and decrease unnecessary utilization [16].

Multiple claims-derived CFIs exist. Some of the common CFIs include: (1) Davidoff index [9]: This index uses HCPCS (Healthcare Common Procedure Coding System) and CPT (Current Procedural Terminology) codes in a 12-month timeframe to predict disability among oncology patients. This index does not include diagnosis codes or age. (2) Faurot index [10,11]: This index uses demographics, ICD (International Classification of Diseases), and HCPCS codes in the past 8 months to predict ADL (Activities of Daily Living) disabilities. This index excludes CPT codes. (3) Ben-Shalom index [14]: This index uses demographics and a variety of specialized diagnostic grouping codes to predict ADL disability. (4) Segal index [12,13]: This index uses demographics, ICD codes, and comorbidity scores in the past 6 months to predict the frailty phenotype. This index excludes CPT and HCPCS codes. (5) Kim index [15,16]: This index uses ICD, CPT, and HCPCS codes in the past 12 months to predict the value of deficit-accumulation frailty index. Kim's frailty index excludes demographic variables (e.g., sex and age).

EHR-based versus Claims-based Frailty Indexes

Healthcare delivery systems, specifically those operating under value-based care contracts, are usually assessed and reimbursed based on patient stratification levels that account for patient-level risk of utilization or adverse outcomes [17]. When serving older adults within such health systems, CFIs could provide a useful tool for care coordination and management. However, unlike payers, complete insurance claims data are generally not available to provider

organizations [18]. Given the increased adoption of electronic health records (EHRs) across the United States, with almost all hospitals and most outpatient practices having a certified EHR since 2017 [19,20], EHR-based frailty indexes (EFIs) have become a feasible substitute for CFIs [21-25]. Past studies, however, have not measured the accuracy and validity of EFIs relative to the more commonly used CFIs.

Several EFIs also have been developed and validated in recent years. Some of the emerging EFIs include: (1) Laken index [22]: This index focuses on inpatient encounters of the Medicare population. The index predicts mortality using diagnosis data captured in a 12-month lookback period. (2) Anzaldi/Kharrazi index [21]: This index focuses on ambulatory care among older adults to predict frailty. The index uses unstructured EHR data, such as clinical notes, to extract novel predictors of frailty. (3) Pejewski index [23]: This index is based on ambulatory care data of Medicare enrollees. The index predicts mortality using diagnosis and medication data captured in the last 24 months. (4) Shao index [24]: This index is limited to the Veterans Health Administration patients and has a 12-month assessment period. The index uses unstructured EHR data and predicts mortality. (5) Clegg index [25]: This index focuses on ambulatory care among UK residents and uses a 12-month period. The index uses diagnosis, medication, and health services data from EHRs. The model predicts both clinical frailty and mortality.

EHRs versus Claims

EHRs are inherently different from insurance claims. Despite the significant overlap between the types of data collected in claims and EHRs, they are designed for different purposes. Administrative claims are collected and reported primarily for reimbursement purposes, while EHRs are mainly used to improve clinical care and support the billing process [26]. Claims data cover a range of events, diagnoses, and procedures that are collected across all providers seen by an individual, but EHRs, unless fully interoperable within a region, contain data collected from clinical encounters that occurred between an individual and providers within a single health system, hence, missing "out-of-network" events. On the positive side, EHRs provide new data types, features, and augmentations that are not collected through the claims process, such as family history, symptoms, procedure results, lab information, vital signs, and e-prescriptions [26,27]. EHRs also provide more timely data than claims data, making them a better choice for predictive models targeting outcomes requiring quick turnaround interventions (e.g., predicting 30-day hospital readmission).

Comparing the completeness of data captured in EHRs versus claims is required before developing or validating EFIs. Contrary to the common belief that EHR data of a patient contains the same information as the patient's insurance claims, data captured in EHRs and insurance claims overlap only partially within a health system (**Figure 1**). Patients enrolled in "closed" health systems, where providers and insurers share organizational structure, generally receive most of their care internal to the system and thus have a higher overlap between their EHR and claims. Non-closed health systems, where the patient often receives substantial amounts of care

external to the system, will have lower levels of overlap between their EHRs and claims. The degree of overlap will depend on a variety of factors, including range of specialties in the system, degree of EHR interoperability, and degree of incentives or loyalty that may lead patients to seek care within the system.
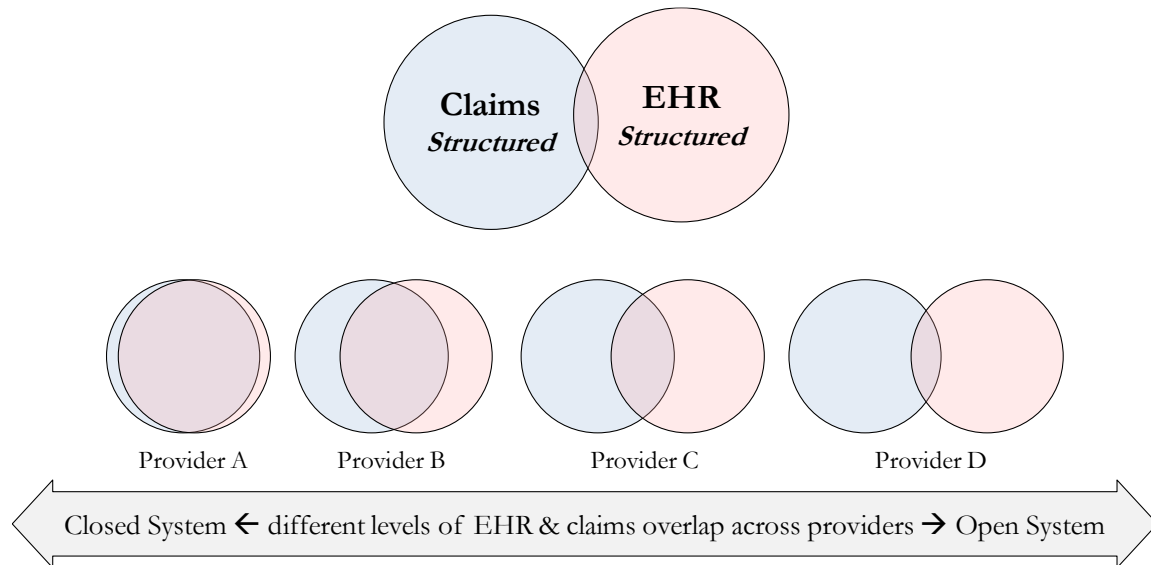
Figure 1. Schematic comparison of structured data of insurance claims

Provider A to D Venn diagrams: schematic representation of various levels of electronic health records and insurance claims data overlap across different providers

Using Frailty Measures for Risk Stratification

Population health aims to improve the health outcomes of defined populations by modifying health determinants that range from clinical to social and environmental. Within healthcare systems, population health involves the identification of patients and enrollees who are prone to having higher-than-normal risks of undesired outcomes. Depending on the anticipated risk, providers and payers place patients in different risk groups to be targeted by interventions appropriate to their risk level. For example, high-risk patients will be closely followed by a case manager, while low-risk healthy members receive an annual needs-assessment questionnaire. Given the rapid growth of population health management efforts, risk stratification activities are becoming essential for prioritizing population-level interventions [28].

In risk stratification analytics, healthcare utilization and cost are often considered dependent variables. Various determinants of health (e.g., clinical, behavioral, social) are used as independent variables in risk stratification models to predict the dependent variables. Often, the clinical determinants of health (e.g., diagnosis, medications) are grouped into higher-level concepts to facilitate the statistical modeling process. Frailty, as a unique concept of aging that is partially correlated with comorbidities [15], is increasingly being used for risk prediction. A

growing number of studies have shown the value of frailty in improving risk models of utilization that already include demographic and comorbidity scores [16,29].

Traditionally, claims data are used to train and validate risk stratification models of utilization; however, EHR-extracted determinants of health are increasingly being used for risk modeling (e.g., extracting diagnosis data from EHRs to predict cost). Although using CFI in risk prediction is well established, increasingly EHRs also are being used for risk stratification. This has propelled the potential use of EFIs in risk prediction models (e.g., using EFIs to better predict nursing home admissions [29]).

Overall Project Tasks

To address the operational gap between CFIs and EFIs, this project focused on validating an established CFI using linked claims-EHR databases of multiple large health systems. Additionally, the project aimed to refine the EFI for potential use across both older (65+) and younger adults (18-64). This project also included an assessment of the validated CFI versus its EFI version as a tool for concurrent and prospective prediction of healthcare utilization among older and younger adults.

In this project, each task included multiple subtasks. **Task-1** included administrative subtasks such as acquiring the required legal contracts and data use agreements to access the linked EHR-claims data sources, gathering information on various CFIs and EFIs, and soliciting feedback from EHR Learning Network [30] on operationalizing EFI in an EHR environment. **Task-2** included data management subtasks focusing on linking EHRs with claims, formulating and identifying the patient denominators (e.g., 18-64 vs. 65+), measuring the completeness of frailty-specific codes across EHRs and claims data, and comparing patient-level comorbidity scores across EHRs and claims data. **Task-3** included risk prediction subtasks such as validating and refining CFI and EFI against a set of functional disability outcomes, assessing the distribution and concordance of CFI vs. EFI, comparing CFI vs. EFI across multiple health systems, measuring potential disparities using CFI vs. EFI, and comparing the predictive value of CFI vs. EFI in forecasting healthcare utilization. **Task-4** included dissemination subtasks such as presentation at scientific venues, sharing the underlying code in a standardized format, and publishing the study findings in scientific journals.

This report focuses on the methods, results, and interpretation of results for Tasks 2 and 3. Tasks 1 and 4 are deemed administrative tasks, which did not generate standalone scientific results, hence are not included in this report. See **Appendix A** for additional details of the overall project tasks and subtasks.

2. Methods

Data Sources

Three data sources were used in this project: Johns Hopkins Medical Institute (JHMI), Optum Labs Data Warehouse (OLDW), and Kaiser Permanente Mid-Atlantic States (KPMAS). All data sources contained both EHR and claims data to facilitate the comparison of CFI vs. EFI for the same patient population. JHMI and KPMAS data sources represented data from individual health systems, but OLDW included data from 55 health systems. The initial denominator of JHMI, OLDW, and KPMAS were ~160k, ~29,486k, and ~161k patients. Data included anonymized patient-level (row-level) data from both EHR and claims data for each patient. Various inclusion and exclusion criteria were applied to select the denominator of interest for various tasks in this project (e.g., minimum 6 months of coverage in claims). All three data sources (i.e., JHMI, OLDW, and KPMAS) were used in Tasks 2 and 3, but some of the subtasks only used one of the data sources (i.e., subtask 3.1 used KPMAS and 3.3 used OLDW).

As a closed health system, KPMAS claims data are collected differently than typical claims collected by JHMI or other non-closed health systems represented in OLDW. KPMAS is a fully integrated delivery system providing both healthcare and insurance services for the same patient population. KPMAS' internal claims data are fully integrated and mirror the data captured in the EHR, hence making the extraction of internal claims impractical. Consequently, KPMAS claims data used in this project are mainly driven by encounters originating outside of the KPMAS network. Indeed, KPMAS' external claims data are usually collected for specific services contracted with other providers that are not offered by KPMAS. In this project, KPMAS claims data refer to out-of-KPMAS services. See **Appendix B** for additional details of the data sources used for this project.

Claims-based Frailty Index

In a prior study funded by ASPE, Kim's frailty model was identified as the measure of choice to assess frailty using claims data [30, 31]. Kim's CFI is mainly trained on Medicare data to identify frail older adults; however, Kim's model has not been formally tested using EHR data. This project aimed to identify the appropriateness of Kim's claims-derived frailty index using EHR data of several large health systems.

Timeline and Eligibility Requirements

Both EHR and claims data were extracted for encounters occurring between Jan 1, 2016, and Dec 31, 2019. Patient eligibility included the following criteria: (1) valid sex and age, (2) age of 18 or higher in any calendar year, (3) one or more eligible EHR records, and (4) 6 or more months of claims enrollment per calendar year.

Task 2: Link and Compare Claims Data with EHR

- ### Subtask 2.1: Check the linkage of EHR and claims in underlying data sources (JHMI, OLDW, KPMAS)

The accuracy of the matched EHR and claims records were examined by comparing variables used in the matching process. For example, patients' age, gender, and race were matched against the same data in the claims data. Patients with unmatching demographics were excluded from the analysis. Significant changes in the total number of patients that are caused by removing incorrectly matched patients from any of the data sources were further analyzed. Descriptive analytics were used to describe the results of subtask 2.1.

- ### Subtask 2.2: Formulate and identify denominators/cohorts of patients

Demographic, clinical, and health utilization variables were used for population stratification (e.g., age, sex, comorbidity score, encounter type, length of continuous insurance coverage, and minimum number of EHR encounters). Patients who have died during the analysis period (e.g., died during the minimum continuous insurance coverage; died in any given year of the frailty model) were excluded. Patients whose information on time to event (e.g., healthcare utilization) was not available due to loss to follow-up or nonoccurrence of outcome event during the analysis period were also censored (e.g., not having insurance coverage during the prospective year of the analysis, hence missing potential prospective hospital admissions). Descriptive analytics were used to describe the results of subtask 2.2. Results were stratified by various demographic data such as age (e.g., 18-64 years old vs. 65+ years old patients).

- ### Subtask 2.3: Measure the completeness of frailty predictors across EHR and claims data

All three data sources were included in this subtask. Predictors needed by the Kim's model were available in the underlying data sources. Each frailty-specific predictor (i.e., diagnostic and procedure codes used to calculate Kim's CFI) were classified into one of the following categories: (1) in both EHR and claims, (2) in EHR only, (3) in claims only, and (4) in neither EHR nor claims. Descriptive analytics were used to describe the distribution of each predictor of Kim's CFI (e.g., diagnostic and procedure codes) across different data sources.

- ### Subtask 2.4: Compare patient-level comorbidity scores across EHR and claims data

Comorbidity scores are different from frailty indexes; however, comorbidities may correlate with frailty. Prior studies have adjusted frailty indexes based on the underlying comorbidity levels of individual patients. Nonetheless, comorbidity scores calculated based on EHRs vs. claims data have shown meaningful differences. Thus, to disentangle the effect of comorbidities on frailty

indexes, comorbidity scores were calculated and compared in EHRs vs. claims data of all three data sources.

For each eligible individual, we calculated three different comorbidity indexes: Charlson [32, 33], Elixhauser [34], and AHRQ's Clinical Classifications Software Refined (CCSR) [35]. For each comorbidity index, we calculated two types of scores: EHR-based comorbidity score and claims-based comorbidity score. EHR-based vs. claims-based comorbidity were described among eligible individuals by data sources.

Task 3: Evaluate the Accuracy and Concordance of Claims vs. EHR-based Measures of Frailty

- **Subtask 3.1: Validate and refine claims and EHR-based frailty measures in predicting functional disability**

None of the project's data sources included a common survey-based frailty or functional disability index on a population level, thus limiting the assessment of the accuracy of Kim's frailty index in identifying survey-based frailty using the EHR data. However, KPMAS data included Medicare Total Health Assessment (MTHA) survey responses for a segment of the patient denominators. As the MHTA survey includes questions that are similar to common frailty surveys, the MTHA survey was used as a proxy for survey-based frailty to evaluate the performance of Kim's CFI in the EHR context. To achieve this retrospective analysis, Medicare Advantage members ≥ 40 years of age with an encounter with KPMAS between Jan 2018 and Jun 2020 were included. Prevalence, incident, and worsening frailty were estimated from EHR data using the Kim's CFI. Multivariable models of prevalence, incident, and worsening frailty were estimated as a function of MTHA and EHR frailty indicators, adjusted for respondent covariates such as age, gender, ethnicity/race, and residential area disadvantage index. For concurrent analysis, incident, and worsening frailty, we examined the bivariate associations of the MTHA and EHR frailty markers and respondent covariates using a chi-square test of independence. Significance was declared with p-values ≤ 0.05.

- **Subtask 3.2: Assess distribution and concordance of claims and EHR-based measures of frailty**

Data were limited to observations between Jan 1, 2016, and Dec 31, 2019, across each of the three data sources (i.e., JHMI, OLDW, and KPMAS). Each data source contained records from both EHR and claims, but patient selection for concordance necessitated the inclusion of only those patients represented in both. Additional patient inclusion criteria required sufficient data quality: (1) consistent, non-missing age and sex across EHR and claims; (2) aged 18 years or older for each year of observation; (3) having at least 6 months of continual enrollment in medical insurance for each year; and (4) having at least one valid EHR encounter for each year. The distribution of CFI and EFI at various levels was described as: (a) a binary indicator of

having frailty and (b) a continuous frailty score. The agreement between CFI and EFI at various levels were measured.

Annual frailty concordance was assessed using a combination of statistical tests that capture covariance and exact agreement between EFI and CFI. For continuous values of either index, these consisted of within-year Spearman's Rho. The concordance of derived binary markers for frailty were assessed using a threshold recommended by Kim's CFI. Statistical tests included the phi coefficient (ranging between 0 and 1, with 1 suggesting the highest association) and percent exact agreement (ranging between 0 and 100, with 100 indicating a perfect agreement), with significance testing through corresponding p-value and Cohen's kappa, respectively. Results were considered significant at p-values less than .001 due to the large sample sizes.

Statistically significant differences in the distribution and marginal frequencies of the continuous frailty score and the binary frailty marker were also evaluated. For continuous values, the statistical analysis consisted of a Wilcoxon signed rank test, coupled with mean difference in EFI-CFI to illustrate effect size. For binary frailty, McNemar's test was used to determine significant differences. In either case, tests were considered significant using an alpha of .001.

The Charlson index components were further subdivided to check the relative prevalence of several common disease categories across EHRs and claims. Because the Kim's model is a composite of different diagnosis and procedure coded events, like the Charlson comorbidity index, it was possible to evaluate whether any imbalance in EHRs versus claims recorded observation was unique to the frailty model.

- **Subtask 3.3: Compare the claims vs. EHR-based frailty measures across health systems**

CFI and EFI were compared across 55 heath systems represented in OLDW, as OLDW was the only data source to provide such information. Eligible individuals from OLDW were assigned to one of the health systems. For each health system, the distributions of EFI, CFI, and their difference (CFI vs. EFI) were described. Statistical tests were conducted to examine whether significant differences exist in EFI, CFI, and their differences across various health systems. Chi-squared tests of independence were performed for binary measures of frailty, while continuing to use .001 as the alpha level for significance testing. Due to the volume of the findings, much of the results for OLDW health systems were condensed into a set of visualizations that reflect distributions of EFI and CFI in each health system and respective proportions of frail patients in EHR or claims.

- ## Subtask 3.4: Measure potential disparities in claims- vs. EHR-based frailty indexes across subgroups

In this subtask, the potential disparities in CFI vs. EFI across subgroups were explored. Patients were assigned to one of the mutually exclusive subgroups based on a given characteristic, such as age group, sex, or race. For each level of a subgroup, the distributions of EFI, CFI, and their difference (CFI vs. EFI) were described. Age groupings were bounded such that two categories were before and after age 65, the Medicare age-eligibility criterion. Those groups included ages 18 to 49, 50 to 64, 65 to 79, and 80+. Due to variability in the race and ethnicity recording observed across data sources, it was necessary to conform our strata for racial designation into a small number of categories: African American, Asian, White, and Other/Unknown race.

Statistical tests included the Spearman's Rho, phi, percent exact agreement, and Cohen's kappa. These tests were used to inspect the variable strength of the EFI-CFI association within different subpopulations and test for statistical significance. Multiple between-strata comparisons were also performed using chi-squared tests of independence and Wilcoxon signed rank tests for binary and continuous measures of frailty, respectively.

- ## Subtask 3.5: Evaluate the value of claims- vs. EHR-based frailty measures in predicting healthcare utilization

This subtask included both descriptive and predictive modeling techniques to measure the value of Kim's CFI and its equivalent EFI in predicting healthcare utilization. Patients with pregnancies and/or deliveries in any of the data years (i.e., 2016 to 2019) were excluded from this analysis, as pregnancies and/or deliveries are usually not included in healthcare utilization analysis (i.e., pregnancy-derived utilization is primarily not due to a disease).

In the predictive modeling analysis, the predictors included Kim's frailty index and several control variables. Kim's frailty index was calculated using EHR (i.e., EFI), claims (i.e., CFI), and claims-EHR overlap (CEFI) data. Control variables included age groups, sex, race, type of insurance, and the comorbidity indexes. Three binary healthcare utilization outcomes included: having any inpatient visit (IP), having any emergency department visit (ED), and having any nursing home visit (NH). The analyses were conducted using both concurrent and prospective modeling techniques. In the prospective model, frailty indicators and control variables were derived from the first year, while outcomes were constructed using the second year. However, in concurrent modeling, the predictors and outcomes are extracted from the same year of the data.

The analysis included multiple steps. First, the characteristics of the eligible sample population were described and then stratified by data source (i.e., EHR, claims, EHR-claims overlap) and data year (i.e., 2016, 2017, 2018, and 2019). Second, separate predictive models were developed for the healthcare utilization outcomes (i.e., IP, ED, NH) with three sets of Kim's frailty indexes (i.e., EFI, CFI, and CEFI) and several control variables (i.e., age groups, sex, race, type of

insurance, and the comorbidity indexes). Then, logistic regression and generalized estimating equation (GEE) were adopted given the binary outcomes and the correlation among individuals contributing to multiple observations to the datasets across the years. Unstructured correlation structure for GEE were chosen, given its low quasi-likelihood under the independence model criterion value. A randomly selected 80% of eligible samples were used for model development while the remaining 20% were used for model validation. Model performance measures were derived from the validation set, and individuals were flagged as predicted users of a given healthcare service if the corresponding predicted probabilities were among top 5% of eligible individuals. Performance measures included area under the curve (AUC), percentage of agreement between predictive and actual users, sensitivity, specificity, positive predicted value (PPV), and negative predicted value (NPV). Additionally, the odds ratios (ORs) associated with frailty were compared across the three frailty indexes (i.e., CFI, EFI, and CEFI).

3. Results

Task 2: Link and Compare Claims Data with EHR

- ### Subtask 2.1: Check the linkage of EHR and claims in underlying data sources (JHMI, OLDW, KPMAS)

JHMI included 159,680 patients who had both EHR and claims data with no missing sex or age. Close to half of the patients had an unknown race; however, race was not used as an inclusion and exclusion condition in the selection process (**Table 1** & **Appendix C Table 1**). OLDW included 29,386,456 patients who had both EHR and claims data with no missing sex or age. Close to half of the patients had an unknown race; however, race was not used as an inclusion/exclusion criterion in the selection process (**Table 2** & **Appendix C Table 2**). KPMAS included 160,893 patients who had both EHR and claims data. Due to the unique nature of KPMAS data, and the fact that KPMAS claims data refers to external encounters, separate tables were not created for KPMAS internal claims data (**Table 3**).

Table 1. Total number of JHMI patients with both EHR and claims data across all years

Matching Process	Both EHR & Claims
Original N (Non-Missing)	159,980 (100.0%)
Same Sex	159,831 (99.91%)
Same Age *	159,828 (99.90%)
Same Race **	73,388 (45.97%)
Final N	**159,680 (99.81%)**

* "Same age" was defined as the age difference within ±0.05 years from both sources.

** No exclusions were made based on Race: "White," "African American," "Asian," and "Other/Unknown"

Table 2. Total number of OLDW patients with both EHR and claims data across all years

Matching Process	Both EHR and Claims
Original N (Non-Missing)	29,485,939 (100%)
Same Sex	29,388,582 (99.7%)
Same Age *	29,483,137 (99.9%)
Same Race **	16,509,095 (56.0%)
Final N ***	**29,386,456 (99.7%)**

* "Same age" was defined as the age difference within ±0.05 years from both sources.

** No exclusions were made based on Race: "White," "African American," "Asian," and "Other/Unknown"

*** Final denominator changed slightly in various subtasks due to the dynamic nature of the OLDW database

(i.e., OLDW data are updated regularly).

Table 3. Total number of KPMAS patients with both EHR and claims data across all years

Matching Process	Both EHR & Claims
Original N (Non-Missing)	161,200 (100.0%)
Same Sex	161,054 (99.91%)
Same Age *	161,040 (99.90%)
Same Race **	73,412 (45.54%)
Final N	**160,893 (99.81%)**

* "Same age" was defined as the age difference within ±0.05 years from both sources.

** No exclusions were made based on Race: "White," "African American," "Asian," and "Other/Unknown"

- ## Subtask 2.2: Formulate and identify denominators/cohorts of patients

The final denominator of JHMI patients were identified after applying the eligibility criteria to both EHR and claims data for each calendar year. The final denominator of JHMI patients varied between 51k and 61k across different calendar years (**Table 4**). The stratification revealed a low sample size for 65+ year old population (due to missing Medicare claims data in JHMI's claims data), hence excluding them in future data analysis of JHMI data (**Appendix C Table 3**). The stratification also showed a higher ratio for the African American race compared to the national average in 2019 (i.e., 43% instead of ~13%) among JHMI patients (**Appendix C Table 3**).

Table 4. Final denominator of JHMI patients per calendar year (2016 to 2019)

Criteria	2016	2017	2018	2019
Original N	159,680 (100.0%)	159,680 (100.0%)	159,680 (100.0%)	159,680 (100.0%)
EHR Eligibility				
Not 18+ as of Jan 1	9,853 (6.17%)	6,591 (4.13%)	3,713 (2.33%)	1,111 (0.70%)
No EHR activity	68,710 (43.03%)	70,343 (44.05%)	62,815 (39.34%)	80,105 (50.17%)
Final EHR N	84,702 (53.04%)	84,840 (53.13%)	94,013 (58.88%)	78,642 (49.25%)
Claims Eligibility				
Not 18+ as of Jan 1	9,845 (6.17%)	6,565 (4.11%)	3,701 (2.32%)	1,103 (0.69%)
< 6m med enroll.	67,229 (42.10%)	62,932 (39.41%)	70,065 (43.88%)	81,129 (50.81%)
Final Claims N	85,104 (53.30%)	91,510 (57.31%)	86,672 (54.28%)	77,693 (48.66%)
Final N	**57,797 (36.20%)**	**60,792 (38.07%)**	**60,925 (38.15%)**	**51,469 (32.23%)**

Similar to JHMI, the final denominator of OLDW patients were identified after applying the eligibility criteria to both EHR and claims data for each calendar year. The final denominator of OLDW patients varied between 2m and 2.4m across different calendar years (**Table 5**). Among

OLDW patients, the stratification showed a lower ratio for the African American race compared to the national average (i.e., 5% instead of ~13%; **Appendix C Table 4**).

Table 5. Final denominator of OLDW patients per calendar year (2016 to 2019)

Criteria	2016	2017	2018	2019
Original N	29,386,456 (100.0%)	29,386,456 (100.0%)	29,386,456 (100.0%)	29,386,456 (100.0%)
EHR Eligibility				
Not 18+ as of Jan 1	3,435,947 (11.69%)	3,153,773 (10.73%)	2,884,897 (9.82%)	2,639,005 (8.98%)
No EHR activity	17,597,078 (59.88%)	17,893,749 (60.89%)	19,043,175 (64.80%)	19,951,196 (67.89%)
Final EHR N	10,432,051 (35.50%)	10,250,471 (34.88%)	9,312,614 (31.69%)	8,553,986 (29.11%)
Claims Eligibility				
Not 18+ as of Jan 1	3,435,947 (11.69%)	3,153,773 (10.73%)	2,884,897 (9.82%)	2,639,005 (8.98%)
< 6m med enroll.	22,903,369 (77.94%)	22,789,606 (77.55%)	22,930,440 (78.03%)	23,100,289 (78.61%)
Final Claims N	5,557,515 (18.91%)	5,676,973 (19.32%)	5,595,828 (19.04%)	5,489,424 (18.68%)
Final N	**2,478,649 (8.43%)**	**2,486,447 (8.46%)**	**2,278,947 (7.76%)**	**2,076,817 (7.07%)**

The final denominator of KPMAS patients were identified after applying the eligibility criteria to both EHR and claims data for each calendar year. The final denominator of KPMAS patients varied between 110k and 122k across different calendar years (**Table 6**). Similar to JHMI, the stratification showed a higher ratio for the African American race compared to the national average (i.e., 45% instead of ~13%) among KPMAS patients (**Appendix C Table 5**).

Table 6. Final denominator of KPMAS patients per calendar year (2016 to 2019)

Criteria	2016	2017	2018	2019
Original N	796,625 (100.0%)	845,391 (100.0%)	921,141 (100.0%)	911,173 (100.0%)
EHR Eligibility				
Not 18+ as of Jan 1	158,991 (19.96%)	168,179 (19.98%)	177,733 (19.29%)	177,802 (19.51%)
No EHR activity	94,011 (11.80%)	102,830 (12.16%)	117,363 (12.74%)	110,791 (12.16%)
Final EHR N	543,623 (68.24%)	574,382 (67.94%)	626,045 (67.96%)	622,580 (68.33%)
Claims* Eligibility				
Not 18+ as of Jan 1	-	-	-	-
< 6m med enroll.	485,098 (89.23%)	520,660 (90.65%)	573,242 (91.57%)	562,718 (90.38%)
Final Claims N	110,979 (100.0)%	114,692 (100.0%)	119,006 (100.0%)	122,723 (100.0%)
Final N	**110,979 (100.0)%**	**114,692 (100.0%)**	**119,006 (100.0%)**	**122,723 (100.0%)**

* Note: KPMAS claims data only includes encounters/services provided out of the KPMAS system and does not include internal billing data

Across different data sources, JHMI had the lowest percentage of older adults (65+ year old). The age composition of OLDW and KPMAS were more similar than JHMI (**Figure 2**); however, racial distribution of JHMI and KPMAS were more similar than OLDW data (**Figure 3**).
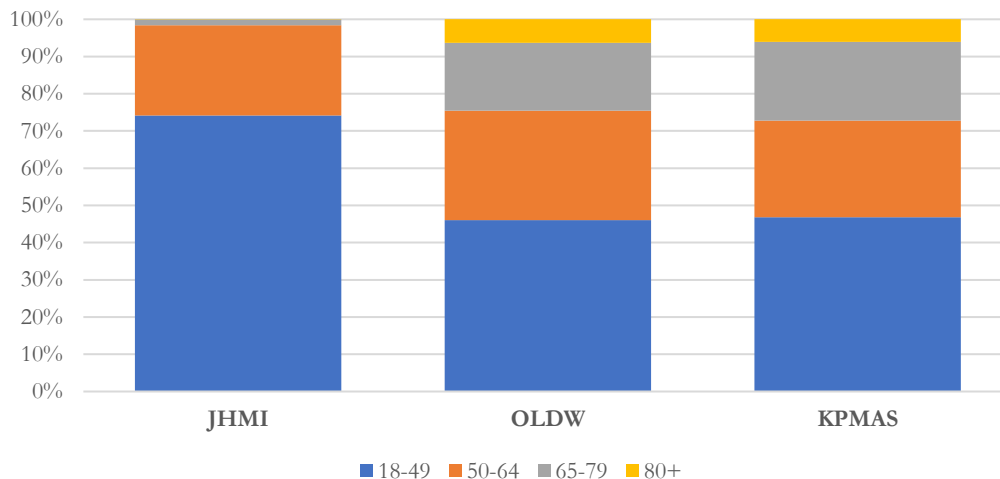
Figure 2. Composition of age groups across data sources
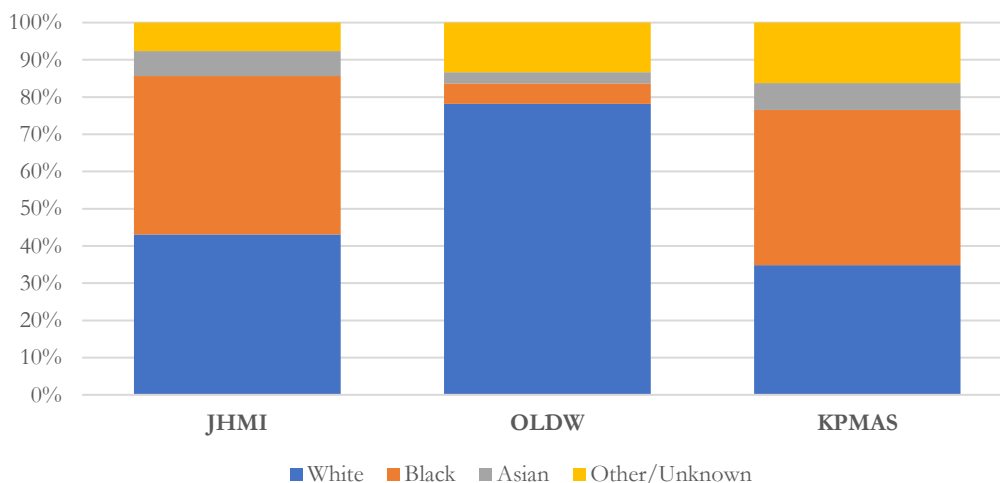
Figure 3. Composition of racial groups across data sources

- **Subtask 2.3: Measure the completeness of frailty predictors across EHR and claims data**

The percentage of patients with a given frailty predictor (i.e., predictors of Kim's model for frailty) in EHR vs. claims were calculated for each year. Then, the ratio of claims vs. EHR in identifying the frailty predictors were calculated for each data source. Comparing the predictors

14

of frailty across JHMI's EHR and claims data revealed that claims data consistently adds more value compared to EHR data in detecting frailty predictors. The ratio of patients found with a given frailty predictor in claims compared to JHMI's EHR data ranged between 5 to 1 after dropping predictors with low counts (**Appendix C Table 6**). Note that the existence of a frailty predictor does not imply the existence of frailty for a patient. Indeed, frailty predictors have different weights (and sometimes negative weights) in Kim's model. The total score of Kim's model should be used to identify a patient as frail or not. Overall, the OLDW results were similar to JHMI results with claims data showing more value in identifying frailty predictors across years (**Table 7**). KPMAS claims data mainly included data captured by external encounters, thus missing integrated internal billing-EHR records. Hence, the KPMAS results showed a lower value for claims data in identifying frailty predictors compared to EHR data (i.e., the ratio of claims vs. EHR was lower than 1 in most instances; **Appendix C Table 7**).

Table 7. Prevalence of frailty predictors across OLDW's EHR and claims data across years

	2016			2017			2018			2019			2016	2017	2018	2019
Pr.*	E%	C%	E+C%	E%	C%	E+C%	E%	C%	E+C%	E%	C%	E+C%		Ratio: C/E		
1	0.6	1.2	1.3	0.6	1.3	1.4	0.6	1.2	1.4	0.6	1.2	1.3	2.00	2.17	2.00	2.00
2	0.3	0.6	0.7	0.3	0.6	0.6	0.2	0.6	0.7	0.3	0.6	0.7	2.00	2.00	3.00	2.00
3	2.6	7.4	7.9	2.6	7.7	8.2	2.6	8	8.5	2.7	8.2	8.7	2.85	2.96	3.08	3.04
4	0.8	1.5	1.6	0.8	1.6	1.7	0.8	1.7	1.8	0.9	1.7	1.8	1.88	2.00	2.13	1.89
5	5.5	16.5	17.2	5.5	17.1	17.7	5.6	18.1	18.8	5.8	18.8	19.4	3.00	3.11	3.23	3.24
6	1.6	6.6	6.8	1.6	6.9	7.1	1.7	6.8	7.1	1.7	6.9	7.1	4.13	4.31	4.00	4.06
7	7.8	13.8	14.5	8	14.4	15.1	8.3	15.1	15.7	8.7	15.3	15.9	1.77	1.80	1.82	1.76
8	10.3	17.1	17.8	10.6	17.9	18.5	10.8	18.6	19.1	11.4	18.8	19.3	1.66	1.69	1.72	1.65
9	26.6	42.2	45.1	27.7	44.4	47.1	28.6	46.6	49.2	30.6	47.6	50	1.59	1.60	1.63	1.56
10	3.2	6.6	7.2	3.4	7	7.6	3.3	7.3	7.9	3.5	7.5	8	2.06	2.06	2.21	2.14

* Only the first 10 frailty predictors are shown (out of 93 diagnoses and procedures used as predictors in Kim's model): (1) Intestinal infectious diseases; (2) Other bacterial diseases; (3) Mycoses; (4) Malignant neoplasm of other/unspecified sites; (5) Benign neoplasms; (6) Neoplasms of uncertain behavior; (7) Disorders of thyroid; (8) Diseases of other endocrine glands; (9) Other metabolic and immunity disorders; and (10) Other and unspecified anemias.

C = Claims, Pr. = [Frailty] Predictor, and E = EHR.

- **Subtask 2.4: Compare patient-level comorbidity scores across EHR and claims data**

Comorbidity scores were calculated using three established methods: Charlson index, Elixhauser index, and CCSR (developed by AHRQ). Comparing the comorbidity scores using JHMI's EHR vs. claims data revealed that claims data usually result in higher comorbidity scores than EHR data alone (**Appendix C Table 8**). Similar to JHMI, claims data revealed a higher yield

compared to EHR data while calculating various comorbidity scores in the OLDW denominator (**Table 8**). In contrast to JHMI and OLDW data, KPMAS claims data revealed a lower value compared to EHR data while calculating various comorbidity scores (**Appendix C Table 9**). This was partly due to the fact that KPMAS claims data mainly captured encounters external to the KPMAS network.

Table 8. Comorbidity scores across OLDW's EHR and claims data across years

	2016			2017			2018			2019			2016	2017	2018	2019
Comorbidity Score	**E**	**C**	**E+C**	**E**	**C**	**E+C**	**E**	**C**	**E+C**	**E**	**C**	**E+C**	**Ratio: C/E**			
# CCSR conditions	8.4	19.2	19.9	9.0	20.4	21.1	9.1	21.5	22.1	9.9	22.2	22.8	2.27	2.26	2.36	2.25
# CCSR body systems	3.6	4.3	5.7	3.7	4.4	5.9	3.7	4.6	6.0	4.0	4.6	6.1	1.19	1.19	1.22	1.16
Mean Elixhauser	0.5	1.1	0.3	0.5	1.1	0.4	0.5	1.2	0.3	0.5	1.1	0.3	2.02	2.17	2.52	2.36
% With 0 Elixhauser	58.1	38.4	35.7	57.1	36.6	34.3	56.2	34.6	32.4	53.5	33.6	31.5	0.66	0.64	0.62	0.63
Mean Charlson	0.5	0.9	1.0	0.6	1.0	1.1	0.6	1.1	1.1	0.6	1.1	1.2	1.78	1.79	1.85	1.75
% With 0 Charlson	76.6	64.4	62.5	75.7	62.7	60.9	75.2	61.3	59.5	73.5	60.5	58.9	0.84	0.83	0.81	0.82

E = EHR; and C = Claims.

Task 3: Evaluate the Accuracy and Concordance of Claims vs. EHR-based Measures of Frailty

- **Subtask 3.1: Validate and refine claims- and EHR-based frailty measures in predicting functional disability**

Only KPMAS data were used for this subtask, as Medicare Total Health Assessment (MTHA) survey was only available in this data source (n = 15,615; 9.26% of KPMAS' study population). Using Kim's frailty index, prevalent frailty was estimated at 9.3%, 1-year incident frailty at 5.8%, and worsening frailty at 4.0%. Analyses indicated several MTHA indicators of frailty (e.g., count of ADL deficits) and EHR-based weight loss in a 12-month period were significantly ($p \leq .05$) associated with all three Kim frailty index measures. A gradient of ADL score with aOR of frailty was noted for all Kim frailty index measures (i.e., EFI, CFI, CEFI). As an instance, the aOR for frailty with an ADL score of ≥ 5 was 2.79, 2.23 for 3-4, and 1.82 for 1-2 (ADL score of 0 is reference). Self-reported cognitive impairment, impaired gait, and low leisure physical activity were also important frailty indicators (**Table 9**). Results of this subtask confirmed the validity of Kim's EFI in measuring frailty in the target population.

Table 9. Association of MTHA and EHR derived frailty markers with concurrent frailty

MTHA Item (Reference Group)	MTHA Description	Sample Percent	Percent Frail	p-value*	aOR (95% CI)	Wald Chi-square
ADL Score (0)	≥ 5	4.75	40.72	<0.01	3.62 2.93, 4.49	173.20
	3-4	6.82	26.58	<0.01	2.76 2.27, 3.39	173.20
	1-2	14.88	14.63	<0.01	1.79 1.52, 2.10	173.20
	0	73.56	4.52	<0.01	–	–
Overall Health	Poor (Fair/Good/Very Good/Excellent)	4.16	35.09	<0.01	1.55 1.25, 1.91	16.42
Sleep Quality	Very bad (Good)	5.92	22.27	<0.01	1.25 1.02, 1.52	4.56
Task Difficulty	Very / fairly often (Sometimes/Almost never/Never)	6.53	24.39	<0.01	1.37 1.14, 1.66	10.60
Cognitive Issues	Yes (No)	19.34	20.56	<0.01	1.55 1.36, 1.77	41.10
Balance Issues While Walking	Yes (No)	39.89	17.71	<0.01	2.00 1.72, 2.34	77.67
Falls While Walking	Yes (No)	25.54	18.73	<0.01	1.68 1.47, 1.91	61.93
Physical Activity	Low (Moderate/High)	29.44	18.50	<0.01	1.52 1.33, 1.73	39.20
Weight Loss ≥ 10 Pounds	Yes (No)	11.14	22.56	<0.01	2.48 2.10, 2.93	113.18
Age Group	40-64	10.67	15.77	<0.01	2.40 1.85, 3.13	95.34
	65-74	65.41	6.05	<0.01	1.72 1.40, 2.11	95.34
	75-84	20.42	12.45	<0.01	1.00 0.83, 1.20	95.34
	≥ 85	3.90	29.23	<0.01	–	95.34
Sex	Female	57.59	9.93	<0.01	1.01 0.89, 1.14	0.02
	Male	42.41	8.35	<0.01	–	0.02
Race/Ethnicity	Hispanic	3.14	8.78	<0.01	0.73 0.51, 1.05	18.86
	NH Black	27.39	9.98	<0.01	0.90 0.98, 1.04	18.86
	NH Asian	6.79	4.81	<0.01	0.53 0.38, 0.73	18.86

MTHA Item (Reference Group)	MTHA Description	Sample Percent	Percent Frail	p-value*	aOR (95% CI)	Wald Chi-square
	Other	1.58	10.16	<0.01	0.79 0.49, 1.20	18.86
	NH White	61.11	9.43	<0.01	–	18.86
ADI Percentile	Lowest SES	24.84	11.55	<0.01	1.35 1.12, 1.62	14.47
	Lower middle SES	24.23	9.88	<0.01	1.26 1.05, 1.52	14.47
	Upper middle SES	24.51	9.16	<0.01	1.39 1.16, 1.66	14.47
	Highest SES	26.42	6.66	<0.01	–	14.47
c-statistics						0.83
Percent concordant						83.0%

*aOR = adjusted odds ratio. * p-values are repeated across subgroup rows for readability.*

- ## Subtask 3.2: Assess distribution and concordance of claims- and EHR-based measures of frailty

JHMI, OLDW, and KPMAS data sources were used for this subtask. The demographic composition of each source of data had some notable differences but were found to be very consistent across all 4 years within each data source (**Table 10**). Female to male ratio ranged between 3:2 and 2:1. In the JHMI data, patients of Medicare age (i.e., age 65 years or older) made up only 2.1% of the sample annually in 2019, whereas for OLDW and KPMAS, 29.7% and 28.4% of patients were 65+ years old, respectively. The OLDW data had considerably fewer non-White racial minorities than JHMI or KPMAS, with around 19.3% of the OLDW sample being non-White on average versus 58.7% and 67.2% in JHMI and KPMAS, respectively.

Table 10. Patient characteristics by data source for all years

Characteristics	2016	2017	2018	2019
JHMI	N (%)	N (%)	N (%)	N (%)
Original N	57,797 (100)	60,792 (100%)	60,925 (100)	51,469 (100)
Male	19,350 (33.48)	20,474 (33.68)	20,644 (33.88)	17,201 (33.42)
Female	38,447 (66.52)	40,318 (66.32)	40,281 (66.12)	34,268 (66.58)
18-49	42,876 (74.18)	44,790 (73.68)	44,773 (73.49)	37,097 (72.08)
50-64	13,988 (24.2)	14,997 (24.67)	15,034 (24.68)	13,273 (25.79)
65-79	913 (1.58)	983 (1.62)	1,092 (1.79)	1,076 (2.09)
80+	20 (0.03)	22 (0.04)	26 (0.04)	23 (0.04)
Asian	3,884 (6.72)	4,147 (6.82)	4,226 (6.94)	3,668 (7.13)
Black	24,563 (42.5)	25,818 (42.47)	26,225 (43.04)	22,688 (44.08)
White	24,930 (43.13)	25,944 (42.68)	25,655 (42.11)	21,259 (41.3)

Characteristics	2016	2017	2018	2019
Other/Unknown	4,420 (7.65)	4,859 (7.99)	4,819 (7.91)	3,854 (7.49)
OLDW	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Original N	2,492,236 (100)	2,505,418 (100)	2,289,597 (100)	2,098,208 (100)
Male	1,035,171 (41.54)	1,048,993 (41.87)	962,444 (42.04)	885,846 (42.22)
Female	1,457,065 (58.46)	1,456,425 (58.13)	1,327,153 (57.96)	1,212,362 (57.78)
18-49	1,171,479 (47.01)	1,129,013 (45.06)	1,001,818 (43.76)	907,988 (43.27)
50-64	727,655 (29.2)	704,412 (28.12)	633,380 (27.66)	566,479 (27)
65-79	441,238 (17.7)	499,134 (19.92)	485,471 (21.2)	463,427 (22.09)
80+	151,864 (6.09)	172,859 (6.9)	168,928 (7.38)	160,314 (7.64)
Asian	66,884 (2.68)	57,514 (2.3)	51,076 (2.23)	48,683 (2.32)
Black	216,179 (8.67)	213,100 (8.51)	186,794 (8.16)	166,923 (7.96)
White	1,983,529 (79.59)	2,006,670 (80.09)	1,840,199 (80.37)	1,692,242 (80.65)
Other/Unknown	255,644 (10.26)	228,134 (9.11)	211,528 (9.24)	190,360 (9.07)
KPMAS	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Original N	110,979 (100)	114,692 (100)	119,006 (100)	122,723 (100)
Male	42,895 (38.65)	44,571 (38.86)	45,981 (38.64)	47,300 (38.54)
Female	68,084 (61.35)	70,121 (61.14)	73,025 (61.36)	75,423 (61.46)
18-49	51,999 (46.85)	52,824 (46.06)	54,784 (46.03)	56,167 (45.77)
50-64	28,794 (25.95)	29,507 (25.73)	30,690 (25.79)	31,594 (25.74)
65-79	23,419 (21.1)	25,155 (21.93)	25,919 (21.78)	27,035 (22.03)
80+	6,767 (6.1)	7,206 (6.28)	7,613 (6.4)	7,927 (6.46)
Asian	8,062 (7.26)	8,584 (7.48)	10,135 (8.52)	10,321 (8.41)
Black	46,275 (41.7)	48,804 (42.55)	49,518 (41.61)	52,528 (42.8)
White	38,610 (34.79)	38,763 (33.8)	39,480 (33.17)	40,198 (32.76)
Other/Unknown	7,922 (7.14)	8,316 (7.25)	8,201 (6.89)	7,470 (6.09)

Patient frailty was largely consistent across demographics and years within each data source; however, the percentage of patients found to be frail varied across data sources (**Table 11**). For example, in 2019, the percentage of frail patients varied across the data sources: 6.8% for JHMI, 8.8% for OLDW and 11.1% for KPMAS. Female patients were more often found to be frail than males, with females making around 61% of frail patients in the OLDW data across all years. In both OLDW and KPMAS data, a larger share of frail patients was aged 65+ versus those aged 18-64 (OLDW: 71.6% vs. 28.4%; KPMAS: 67.5% vs. 32.4% in 2019). This was largely consistent with the observed Medicare populations at each receptive data source in 2019 (73.6% and 69.4%, respectively). Non-White minority patients made up the majority of frail patients in the JHMI and KPMAS samples annually (56.5% and 59.7% in 2019), while in OLDW the proportion was smaller (16.3% in 2019).

Table 11. Percent of frail patients for all data sources and years

	E & C	E & C	E & C	E & C	Ratio: C/E			
	2016 N (%)	2017 N (%)	2018 N (%)	2019 N (%)	2016	2017	2018	2019
JHMI								
N	57,797 (100)	60,792 (100)	60,925 (100)	51,469 (100)	1	1	1	1
Frail	3,479 (6.02)	4,258 (7)	3,995 (6.56)	3,492 (6.78)	2.16	1.98	1.56	1.66
Male	1,311 (37.7)	1,680 (39.5)	1,538 (38.5)	1,303 (37.3)	2.09	1.89	1.57	1.61
Female	2,168 (62.3)	2,578 (60.5)	2,457 (61.5)	2,189 (62.7)	2.21	2.04	1.55	1.69
18-49	1,719 (49.4)	2,032 (47.7)	1,850 (46.3)	1,552 (44.4)	2.37	2.04	1.62	1.72
50-64	1,690 (48.6)	2,130 (50)	2,077 (52)	1,839 (52.7)	1.98	1.96	1.52	1.65
65-79	65 (1.9)	91 (2.1)	108 (2.7)	97 (2.8)	2.38	1.98	1.33	1.25
80+	5 (0.1)	5 (0.1)	5 (0.1)	4 (0.1)	0	2	2.5	2
Asian	31 (0.9)	48 (1.1)	43 (1.1)	46 (1.3)	2.33	2.21	2.2	1.57
Black	1,767 (50.8)	2,007 (47.1)	1,912 (47.9)	1,717 (49.2)	1.99	1.7	1.34	1.49
White	1,515 (43.5)	1,952 (45.8)	1,807 (45.2)	1,520 (43.5)	2.31	2.24	1.75	1.79
Other/Unknown	166 (4.8)	251 (5.9)	233 (5.8)	209 (6)	3.02	2.91	2.34	2.47
OLDW								
N	2,492,236 (100)	250,5418 (100)	228,9597 (100)	209,8208 (100)	1	1	1	1
Frail	175,968 (7.06)	199,108 (7.95)	192,988 (8.43)	184,026 (8.77)	2.87	2.72	2.84	2.64
Male	68,460 (38.9)	76,135 (38.2)	74,471 (38.6)	70,965 (38.6)	2.86	2.72	2.85	2.63
Female	107,508 (61.1)	122,973 (61.8)	118,517 (61.4)	113,061 (61.4)	2.87	2.72	2.84	2.65
18-49	19,522 (11.1)	19,206 (9.6)	17,339 (9)	15,497 (8.4)	4.72	4.05	4.39	4.12
50-64	42,070 (23.9)	44,024 (22.1)	40,584 (21)	36,792 (20)	3.36	3.07	3.19	2.99
65-79	62,413 (35.5)	74,759 (37.5)	74,907 (38.8)	73,672 (40)	2.64	2.57	2.7	2.52
80+	51,963 (29.5)	61,119 (30.7)	60,158 (31.2)	58,065 (31.6)	2.48	2.45	2.57	2.4
Asian	1,165 (0.7)	1,175 (0.6)	1,080 (0.6)	1,103 (0.6)	3.43	3.27	3.78	3.34
Black	17,813 (10.1)	20,718 (10.4)	18,799 (9.7)	16,978 (9.2)	2.77	2.63	2.98	2.75
White	143,938 (81.8)	162,631 (81.7)	159,505 (82.7)	153,977 (83.7)	2.71	2.57	2.67	2.5
Other/Unknown	13,052 (7.4)	14,584 (7.3)	13,604 (7)	11,968 (6.5)	8.15	7.33	8.46	6.66
KPMAS								
N	110,979 (100)	114,692 (100)	119,006 (100)	122,723 (100)	1	1	1	1
Frail	10,762 (9.7)	11,514 (10.04)	12,324 (10.36)	13,615 (11.09)	1.54	1.47	1.5	1.54
Male	4,422 (41.1)	4,804 (41.7)	5,189 (42.1)	5,762 (42.3)	1.78	1.69	1.71	1.79
Female	6,340 (58.9)	6,710 (58.3)	7,135 (57.9)	7,853 (57.7)	1.39	1.33	1.36	1.38
18-49	1,123 (10.4)	1,190 (10.3)	1,325 (10.8)	1,466 (10.8)	2.96	2.65	2.68	2.58
50-64	2,544 (23.6)	2,605 (22.6)	2,822 (22.9)	3,103 (22.8)	1.91	1.84	1.91	1.8
65-79	4,251 (39.5)	4,673 (40.6)	4,995 (40.5)	5,526 (40.6)	1.34	1.23	1.26	1.33
80+	2,844 (26.4)	3,046 (26.5)	3,182 (25.8)	3,520 (25.9)	1.32	1.37	1.35	1.45

	E & C	E & C	E & C	E & C	Ratio: C/E			
	2016 N (%)	2017 N (%)	2018 N (%)	2019 N (%)	2016	2017	2018	2019
Asian	508 (4.7)	565 (4.9)	629 (5.1)	742 (5.4)	1.89	1.54	1.74	1.69
Black	4,815 (44.7)	5,195 (45.1)	5,597 (45.4)	6,280 (46.1)	1.63	1.53	1.57	1.61
White	4,584 (42.6)	4,842 (42.1)	5,058 (41)	5,482 (40.3)	1.38	1.33	1.34	1.39
Other/Unknown	446 (4.1)	468 (4.1)	467 (3.8)	460 (3.4)	2.07	2.69	2.01	2.04
Hispanic	409 (3.8)	444 (3.9)	573 (4.6)	651 (4.8)	1.63	1.5	1.69	1.77

E = EHR; and C = Claims.

The added value of EHR data vs. claims data in identifying patients with frailty was also calculated (**Appendix C Table 10**). Overall, frailty was identified more often using claims data compared to EHR data in all age groups across all data sources. Adding EHR to claims data only slightly increased the identification of frail patients in JHMI and OLDW data sources (ranging between 8% to 16% across different age groups in 2016); however, adding EHR data considerably increased the identification of frail patients in KPMAS data (27% to 32% across different age groups in 2016). The later was due to the fact that KPMAS claims data did not include billing data from encounters in KPMAS EHR data, hence mainly representing encounters occurring out of the KPMAS network of providers.

Phi coefficients and exact agreement were considerably similar across data sources and years of observation (**Table 12**). Phi coefficients and kappa statistics almost uniformly reached levels of significance. For the full sample, moderate-to-strong phi were found for EFI-CFI across each data source: .510 for JHMI, .484 for OLDW, and .430 for KPMAS in 2019. Agreement was very high for each data source and year, but the corresponding kappa values were only moderate in size. For 2019, a statistically significant kappa was found with exact agreement for 95.9% (k=.492, p < .001) of the full sample for JHMI, 93.9% (k=.426, p < .001) for OLDW, and 92.9% (k=.419, p < .001) for KPMAS.

Table 12. Concordance of EHR-based and claims-based frailty indexes/measures for each data source and year

	2016		2017		2018		2019	
Source	phi	% Exact	phi	% Exact	phi	% Exact	phi	% Exact
JHMI	0.485 *	95.9 *	0.527 *	95.5 *	0.470 *	95.9 *	0.510 *	95.9 *
OLDW	0.465 *	95.0 *	0.475 *	94.4 *	0.468 *	94.0 *	0.484 *	93.9 *
KPMAS	0.418 *	93.7 *	0.423 *	93.6 *	0.419 *	93.3 *	0.430 *	92.9 *

* p < .001

- ## Subtask 3.3: Compare the claims- vs. EHR-based frailty measures across health systems

OLDW data was used for this subtask as it contains more than one health system. A total of 55 health systems are represented in OLDW between 2016 and 2019, which were included in this subtask. By average, 3.03% of the total population were identified as having frailty using EHR data, while this rate was 8.01% using claims data in 2019. The correlation between the rate of identified frailty based on EHR data vs. claims data across the 55 health systems in 2019 was 0.526 (after excluding health systems with fewer than 1,000 patients), thus demonstrating consistency in both EHR and claims frailty rates across the health systems. Visualizing the rate of patients identified as frail in each health system using EHR and claims depicted the same findings (**Figure 4**). EHRs consistently identified lower rates of frailty compared to claims in each health system; however, at the same time, EHR rates showed a relatively stable distance to claims-derived frailty rates across all health systems. Similar results were found for frailty index/score across OLDW health systems (**Appendix C Figure 1**).
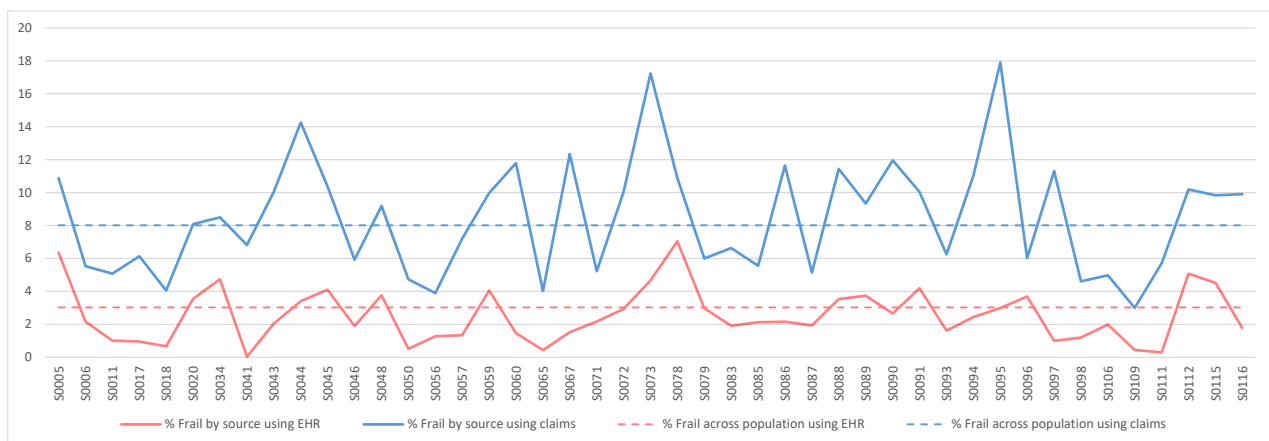
Figure 4. Parallel plot of percent frail patients identified in each health system of OLDW in 2019 using claims and EHR data

Claims data is represented by the blue continues line (i.e., top line) and EHR data is shown by the red continues line (i.e., bottom line); Blue and red dotted lines represent the overall percentage of frail patients identified using all claims and EHR data in 2019.

- ## Subtask 3.4: Measure potential disparities in claims- vs. EHR-based frailty indexes across subgroups

In this subtask, results of subtask 3.2 were stratified by age, sex, and race for all data sources (**Table 13**). When stratified by age, the within-strata phi coefficients were consistently significant in OLDW (18-49: .362; 50-64: .440; 65-79: .470; and 80+: .449) and KPMAS (18-49: .263; 50-64: .347; 65-79: .422; and 80+: .388) in 2019. The phi coefficients were significant for all age groups except 80+ year old patients in JHMI (18-49: .471; 50-64: .519; and 65-79: .594). Note that JHMI data had a small sample size for the 80+ age group. When stratified by sex, the within-strata phi coefficients were also consistently significant in the 2019 sample for JHMI (F:

.513; M: .503), OLDW (F: .486; M: .479), and KPMAS (F: .443; M: .414). Finally, when stratified by race, the within-strata phi coefficients were consistently significant in all data sources, except for Asian in JHMI. Using KPMAS data, the results were also stratified using the common insurance types (**Appendix C Table 11**).

Table 13. Concordance of EHR-based and claims-based frailty for each data source stratified by age, sex, and race

	2016		2017		2018		2019	
JHMI	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**
18-49	0.429 *	97.1 *	0.483 *	96.9 *	0.447 *	97.4 *	0.471 *	97.4 *
50-64	0.519 *	92.2 *	0.548 *	91.3 *	0.469 *	91.5 *	0.519 *	91.9 *
65-79	0.580 *	96.0 *	0.573 *	94.2 *	0.501 *	94.5	0.594 *	95.5
80+	-	75.0 *	0.261	81.8	0.592	88.5	0.673	91.3
Female	0.480 *	96.2 *	0.528 *	95.9 *	0.450 *	96.1 *	0.513 *	96.2 *
Male	0.491 *	95.4 *	0.523 *	94.7 *	0.500 *	95.5 *	0.503 *	95.4 *
Asian	0.489 *	99.4	0.528 *	99.3 *	0.356 *	99.2	0.448 *	99.1
Black	0.507 *	95.3 *	0.564 *	95.3 *	0.497 *	95.7 *	0.534 *	95.7 *
Other	0.397 *	97.2 *	0.395 *	95.9 *	0.361 *	96.4 *	0.449 *	96.3 *
White	0.462 *	95.7 *	0.500 *	95.0 *	0.456 *	95.5 *	0.490 *	95.6 *
OLDW	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**
18-49	0.340 *	98.6 *	0.354 *	98.6 *	0.351 *	98.6 *	0.362 *	98.6 *
50-64	0.413 *	95.6 *	0.431 *	95.4 *	0.426 *	95.2 *	0.440 *	95.3 *
65-79	0.462 *	90.1 *	0.465 *	89.6 *	0.455 *	89.2 *	0.470 *	89.1 *
80+	0.446 *	77.5 *	0.447 *	76.9 *	0.438 *	76.3 *	0.449 *	76.6 *
Female	0.469 *	94.8 *	0.479 *	94.1 *	0.47 *	93.7 *	0.486 *	93.6 *
Male	0.460 *	95.2 *	0.47 *	94.8 *	0.464 *	94.4 *	0.479 *	94.4 *
Asian	0.396 *	98.6 *	0.413 *	98.4 *	0.384 *	98.3 *	0.418 *	98.3 *
Black	0.474 *	94.2 *	0.478 *	93.3 *	0.458 *	92.8 *	0.476 *	92.9 *
Other	0.274 *	94.9 *	0.287 *	94.4 *	0.274 *	94.3 *	0.310 *	94.6 *
White	0.478 *	94.9 *	0.489 *	94.4 *	0.483 *	94 *	0.483 *	93.8 *
KPMAS	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**
18-49	0.241 *	98.3 *	0.221 *	98.3 *	0.234 *	98.1 *	0.263 *	98.0 *
50-64	0.330 *	93.7 *	0.353 *	93.8 *	0.354 *	93.5 *	0.347 *	93.2 *
65-79	0.424 *	88.8 *	0.424 *	88.8 *	0.406 *	87.9 *	0.422 *	87.5 *
80+	0.378 *	75.8 *	0.377 *	75.4 *	0.393 *	76.0 *	0.388 *	74.3 *
Female	0.429 *	94.2 *	0.445 *	94.2 *	0.431 *	93.9 *	0.443 *	93.6 *
Male	0.406 *	93.1 *	0.394 *	92.6 *	0.405 *	92.4 *	0.414 *	91.8 *
Asian	0.375 *	95.6 *	0.403 *	95.7 *	0.411 *	95.8 *	0.397 *	95.0 *
Black	0.413 *	93.2 *	0.432 *	93.3 *	0.407 *	92.5 *	0.433 *	92.4 *
Other**	0.366 *	96.7 *	0.402 *	95.8 *	0.449 *	96.7 *	0.304 *	95.4 *
White	0.432 *	92.6 *	0.427 *	92.2 *	0.434 *	92.0 *	0.440 *	91.6 *

*p < .001; ** "Other" race category in KPMAS data was limited to Other Non-Hispanic population.*

- **Subtask 3.5: Evaluate the value of claims- vs. EHR-based frailty measures in predicting healthcare utilization**

Patients with pregnancies were excluded from all data sources for this subtask as pregnancy and deliveries are not considered disease-driven healthcare utilization and are usually excluded in risk stratification analysis. In 2019, 16.6% of JHMI, 3.16% of OLDW, and 12.3% of KPMAS population experienced pregnancy and were excluded in this analysis (**Appendix C Table 12**). The utilization rates (i.e., IP: inpatient hospitalization, ED: emergency department admission; and NH: nursing home admission) varied across data sources but showed the same trend with ED having the highest rate, following with IP and NH. For example, in 2019, ED, IP and NH were identified in 32.9, 12.1, 1.00 percentage of the patients in JHMI data, and 26.4, 11.3 and 3.2 percentage in OLDW data, accordingly (**Appendix C Table 13**).

A total of 72 models were developed for each data source. Models were comprised of 2 different comorbidity scores (i.e., Charlson and Elixhauser comorbidity scores), 3 different frailty indexes (i.e., CFI, EFI, and CEFI), and 3 outcomes (i.e., IP, ED, and NH). Models were developed using two methods of logistic regression and general estimation equation (GEE). Finally, models were trained using the outcomes of two different years with two different methods: (1) the concurrent analysis, which predicts the outcomes from the same year of the predictors (e.g., both outcomes and predictors are from 2018), and (2) the prospective analysis, which predicts the outcomes occurring in the year after the year of the predictors (e.g., predictors are from 2018, but the outcomes are from 2019). Prospective predictive modeling often results in lower performance due to the temporal effect of unknown predictors changing the outcomes in the future year (i.e., predictors occurring in the same year of the outcome are not included in the prospective analysis). Given that 72 x 3 models were developed across the three data sources (i.e., JHMI, OLDW, KPMAS), and the fact that most models revealed the same trend in each data source, this report includes only the results of a subsample of these models. Reported models are limited to concurrent modeling techniques, logistic regressions methodology, and using the Charlson comorbidity index (CCI).

The added value of CFI, EFI, or CEFI were determined as their OR in predicting the outcomes of interest. For example, when predicting concurrent IP using logistic regression and CCI in JHMI data, the OR of EFI was 7.61, while the OR of CFI was 9.73, thus showing a higher value of CFI in predicting IP admissions (**Table 14**). Additionally, the ORs of EFI and CFI were considerably higher than the ORs of other binary predictors such as sex, race, and age group, thus showing the value of the frailty indexes in improving the concurrent prediction of IP, ED, and NH. However, frailty indexes are binary variables while the Charlson comorbidity index is a continuous (integer) variable, thus the value of frailty indexes should not be assumed to be higher than the comorbidity scores in predicting utilization outcomes.

Table 14. Odds ratios and 95% confidence intervals of concurrent logistic regression models using JHMI data

Predictor	IP		ED		NH	
	EFI* OR (95% CI)	CFI* OR (95% CI)	EFI* OR (95% CI)	CFI* OR (95% CI)	EFI* OR (95% CI)	CFI* OR (95% CI)
Frailty Index	7.61 (7.09 - 8.17)	9.73 (9.23 - 10.27)	6.57 (6.03 - 7.16)	5.63 (5.3 - 5.97)	3.47 (3.04 - 3.96)	40.2 (34.35 - 47.03)
Sex: Male	1.19 (1.15 - 1.24)	1.24 (1.19 - 1.28)	1.01 (0.98 - 1.03)	1.01 (0.99 - 1.04)	1.3 (1.16 - 1.44)	1.45 (1.3 - 1.62)
Race: Asian	0.51 (0.46 - 0.57)	0.57 (0.52 - 0.64)	0.66 (0.62 - 0.7)	0.69 (0.65 - 0.73)	0.31 (0.18 - 0.52)	0.63 (0.37 - 1.08)
Race: Black	0.69 (0.66 - 0.71)	0.75 (0.72 - 0.78)	1.56 (1.53 - 1.61)	1.63 (1.59 - 1.67)	0.63 (0.56 - 0.7)	0.85 (0.76 - 0.96)
Race: Others	0.83 (0.77 - 0.88)	0.85 (0.79 - 0.91)	1.06 (1.02 - 1.11)	1.08 (1.03 - 1.13)	1.07 (0.88 - 1.31)	1.32 (1.07 - 1.63)
Age Group: 50-64	1.05 (1.01 - 1.09)	0.98 (0.95 - 1.02)	0.74 (0.72 - 0.76)	0.72 (0.7 - 0.74)	2.1 (1.87 - 2.37)	1.63 (1.44 - 1.83)
Age Group: 65+	0.45 (0.4 - 0.5)	0.47 (0.41 - 0.53)	0.22 (0.2 - 0.24)	0.23 (0.21 - 0.25)	0.86 (0.62 - 1.2)	0.98 (0.69 - 1.39)
Charlson Comorbidity Index	1.35 (1.34 - 1.36)	1.3 (1.29 - 1.31)	1.17 (1.16 - 1.18)	1.14 (1.13 - 1.15)	1.27 (1.25 - 1.29)	1.13 (1.11 - 1.14)
Year: 2017	0.94 (0.9 - 0.99)	0.93 (0.88 - 0.97)	0.99 (0.96 - 1.02)	0.99 (0.95 - 1.02)	1.27 (1.09 - 1.49)	1.2 (1.02 - 1.41)
Year: 2018	0.89 (0.85 - 0.93)	0.94 (0.89 - 0.98)	0.88 (0.85 - 0.91)	0.9 (0.87 - 0.93)	1.35 (1.16 - 1.58)	1.51 (1.28 - 1.77)
Year: 2019	0.9 (0.86 - 0.95)	0.93 (0.88 - 0.98)	0.89 (0.86 - 0.92)	0.9 (0.87 - 0.93)	1.29 (1.1 - 1.52)	1.32 (1.12 - 1.56)
Insurance: Commercial	0.27 (0.26 - 0.28)	0.3 (0.29 - 0.31)	0.26 (0.25 - 0.27)	0.27 (0.26 - 0.27)	0.13 (0.1 - 0.16)	0.25 (0.2 - 0.31)
Intercept	0.14 (0.14 - 0.15)	0.12 (0.11 - 0.13)	0.68 (0.66 - 0.71)	0.65 (0.63 - 0.67)	0 (0 - 0.01)	0 (0 - 0)
Model's AUC	*0.813*	*0.825*	*0.754*	*0.758*	*0.913*	*0.954*

* Note that CEFI is not shown

AUC: Area under the curve; CEFI: Claims-EHR-based frailty index; CFI: Claims-based frailty index;

ED: Emergency Department; EFI: EHR-based frailty index; IP: Inpatient; and NH: Nursing home

Overall, CFI and CEFI showed an improvement in the OR of predicting concurrent IP and NH admissions compared to ED admissions across all data sources (**Figure 5**). CFI often had a higher OR in predicting the utilization outcomes compared to EFI; however, EFI had a higher OR in predicting ED in JHMI and OLDW. Interestingly, CEFI's OR in predicting various outcomes was often lower than CFI, indicating that EHR data did not add additional value to CFI in predicting those outcomes. Further research is needed to understand why adding EHR data to claims data decreases the OR of frailty indexes in predicting certain utilization outcomes (e.g., OR of CEFI is lower than CFI in predicting NH admission in JHMI and KPMAS).
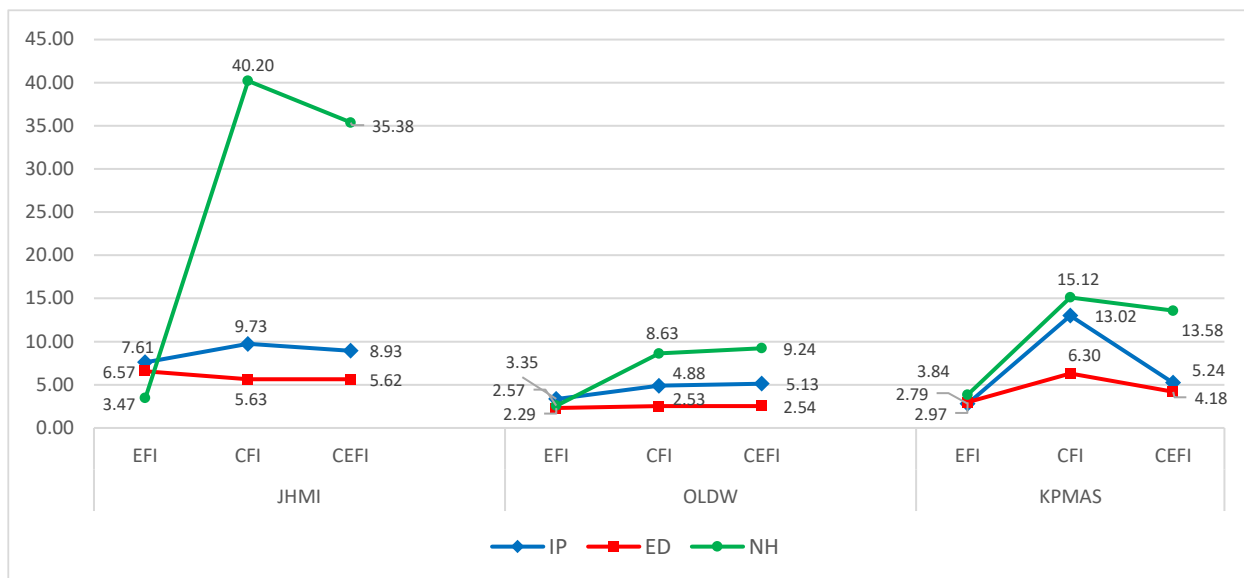
Figure 5. Parallel plot of ORs of frailty indexes in predicting concurrent healthcare utilization outcomes using JHMI (left), OLDW (middle), and KPMAS (right) data

Blue lines with diamond markers: IP; Red lines with square markers: ED; and, Green lines with round markers:
NH Continues lines are shown for visualization purposes only (i.e., to ease depicting an increase or decrease in OR across frailty
indexes); EFI: EHR-based frailty index; CFI: claims-based frailty index; CEFI: claims-EHR-based frailty index;
IP: inpatient admission/hospitalization; ED: emergency department admission; NH: nursing home admission

Overall, the ORs of the frailty indexes were lower in the prospective models compared to concurrent models. However, prospective predictive models of healthcare utilization using EFI, CFI, and CEFI showed similar trends with NH benefiting the most from incorporating the frailty indexes in the models (**Appendix C Table 14**). CFI achieved the highest OR among frailty indexes in prospectively predicting NH using JHMI and KPMAS (OR = 7.10 and 4.98, respectively), while CEFI showed the highest OR among frailty indexes in predicting NH using OLDW data (OR = 3.91). The same trend was detected in predicting IP and ED admissions, with CFI being the highest performing frailty index in JHMI and KPMAS (OR IP = 3.93 and 2.53, OR ED = 2.95 and 2.48, respectively), but CEFI being the best performing index in OLDW data (OR IP = 2.21, OR ED = 1.81).

Both concurrent and prospective models were stratified using race (i.e., White, Black, Asian, Other) and insurance type (i.e., Medicaid, Medicare (if available), non-Medicaid/Medicare). The overall direction of the ORs across outcomes and within data sources did not change and were similar across race or insurance strata (**Appendix C Figure 2, Appendix C Figure 3**).

4. Discussion

Frailty is a vulnerability that affects up to 11% of older adults [1]. Survey-based frailty indexes are useful for targeting care at the point of care [2-8]; however, survey-based frailty indexes are often impractical to collect across large populations of older adults [12]. Consequently, researchers have developed and validated several frailty indexes using more widely available clinical data sources such as insurance claims and EHRs [9-16, 21-25]; however, past studies have not compared the concordance of CFI vs. EHR. To address the operational gap between CFIs and EFIs, this project assessed the reliability of a validated CFI using linked claims-EHR databases of multiple large health systems. Additionally, the task order aimed to refine the EFI for potential use across both older (65+) and younger adults (18-64). This project also included an assessment of Kim's CFI versus its EFI version as a predictor for concurrent and prospective prediction of healthcare utilization among older and younger adults. Findings of this project can be used as preliminary results to promote the use of EFI by healthcare providers and population health managers, who are more likely to have access to EHR data than claims.

Interpretation and Potential Implications of the Findings

- ### Task 2: Link and Compare Claims Data with EHR

Findings of Task 2 revealed a few steps that should be addressed before EFIs are applied to local EHR data. Healthcare providers should make decisions on what is considered an appropriate level of data quality and/or denominator of patients before applying EFIs to EHRs. For example, patients with low-quality EHR data may not have the breadth of diagnostic and procedural data needed to identify frailty. Potential remedies can be selecting a denominator of patients who have sufficient levels of quality for EHR data, or patients who have claims in addition to EHR data (**Tables 3-5**). Measuring the prevalence of underlying diagnoses and procedures used by an EFI in the EHR can also help with identifying potential coding issues if specific groups of diagnoses or procedure codes are missing (**Table 6**). In summary, it is recommended that EHR data be properly prepared and examined for data quality issues before applying any of the EFI measures. Complementing EHR data with additional data types, such as patient-reported outcomes, may increase the value of EHRs to measure frailty; however, more research is needed to assess the value of such data in improving the accuracy of EFIs.

- ### Task 3: Evaluate the Accuracy and Concordance of Claims vs. EHR-based Measures of Frailty

Kim's CFI has been validated in prior studies with respect to survey-based frailty measures [15]. However, Kim's frailty index has not been validated using EHR data. In Task 3, Kim's EFI (i.e., CFI applied to EHR) was validated against a custom survey (i.e., MTHA at KPMAS) as well as against Kim's CFI using claims data for the same patient population across three large data sources (i.e., JHMI, OLDW, and KPMAS). More specifically, subtask 3.1 findings demonstrated

an acceptable rate of agreement between Kim's EFI and MTHA items that were selected as possible markers of frailty. Although MTHA is not a standard survey of frailty, the high correlation between Kim's EFI and select MTHA ADL-related questions/responses depicts an acceptable validity of Kim's frailty index when applied to EHR data (i.e., Kim's EFI; **Table 8**). Additionally, subtask 3.2 showed a high concordance of Kim's EFI vs. CFI with phi ranging between .43 and .51, and exact percentage agreement ranging between 92.9% and 95.9% across the study data sources in 2019 (**Table 11**). Given the prior validation of Kim's CFI against frailty-specific surveys, the high correlation of EFI and CFI can be interpreted as EFI identifying mostly valid frailty cases, but potentially undercounting frailty given the lack of data on care from other providers. Despite the shortcoming of EFIs in identifying all frail patients, EFIs can enable healthcare providers to identify frailty cases in large retrospective EHR data that can be used for population health management purposes and potentially at the point of care.

The findings of subtask 3.3 are key in assessing if CFI and EFI measures are generalizable across different health systems (e.g., generating similar levels of correlation in different health systems). When applied to OLDW data, which included 55 health systems, the results showed a fairly consistent mean of difference between CFI and EFI results across the health system (**Figure 4**). Given the variety of health systems and EHR platforms represented in OLDW data, these results can be interpreted that most health systems should receive similar results (i.e., consistent differences between EFI and CFI) when applying Kim's frailty index to their EHR data. Of course, most health systems do not have access to the claims data of their entire patient population; however, in instances of such data access (e.g., Medicare data in an ACO setting), the providers are encouraged to compare their EFI results with the CFI results to ensure validity of their EFI results.

Potential bias and disparities in various clinical measures is of concern to patients, healthcare providers, and policy makers. Skewed data quality issues, often derived from disparities in healthcare access among minority and vulnerable populations, can lead to biased measurement of clinical outcomes, including frailty. The results of subtask 3.4 depicted an acceptable within- and between-group concordance between EFI and CFI measures (**Table 12**). These results can be interpreted that significant bias was not found when EFI and CFI were compared within and across various age groups, sexes, and races. However, despite these promising results, further research is warranted to explore the potential skewness of EFI vs. CFI in other minority groups (e.g., gender minorities such as LGBTQ+ groups), across other population denominators (e.g., different geographic regions), and different clinical settings (e.g., outpatient clinics, rural health systems).

Prior research has shown that frailty can lead to adverse health outcomes and higher rates of healthcare utilization, such as higher nursing home admissions. [36] Results of subtask 3.5 confirm these past findings by showing the value of EFI, CFI, or the combination of both (CEFI) in predicting various health utilization outcomes (**Table 13**). More specifically, the findings

showed that CFI, EFI, and CEFI statistically significantly improve the prediction of inpatient hospitalizations (IP), emergency department visits (ED), and nursing home (NH) admissions both concurrently and prospectively. The largest improvement in prediction utilization occurs in NH admissions (i.e., highest OR). Utilization prediction improvement was often the highest when using CFI as a predictor, followed by CEFI and EFI (**Figure 5**). As expected with predictive models of healthcare utilization, the ORs of improvements were lower in prospective models compared to concurrent models. Overall, these results can be interpreted that the frailty index improves healthcare utilization prediction; however, the improvements in prediction performance might be different based on the data sources used to calculate the frailty index (e.g., claims vs. EHR) and outcomes of interest (e.g., IP prediction vs. NH prediction). Interestingly, in most predictions, CEFI did not perform better than CFI, thus questioning the value of adding EHR's diagnosis and procedure data to claims data to increase the identification of frailty cases and improve the prediction of healthcare utilization outcomes when claims data are available. Finally, subtask 3.5 findings showed that the value of EFI and CFI in improving IP, ED, and NH predictions is somewhat similar across different patient characteristics such as race (**Appendix C Figures C2**) and insurance plans (e.g., Medicaid vs. Medicare vs. non-CMS; **Appendix C Figures C3**). Thus, the frailty index, either EFI, CFI, or their combination, is recommended to be explored as a potential predictor in future risk stratification or healthcare utilization prediction models.

Key Findings and Implications

- Recommend EHR data be properly prepared and examined for data quality issues before applying EFI measures.

- EFIs can enable healthcare providers to identify frailty cases in large retrospective EHR data and potentially at the point of care.

- CFI and EFI measures are generalizable across different health systems, with consistent differences between data sources across systems.

- Providers are encouraged to compare their EFI results with the CFI results to ensure validity of their EFI results.

- Frailty index improves healthcare utilization prediction, especially nursing home admissions.

- When claims data are available for CFI, adding EHR data does not significantly improve identification of frailty.

Challenges and Opportunities of Using EHRs for Frailty Measurement

Generally, various challenges exist with using EHR data for frailty measurement; however, EHRs also offer potential opportunities that may remedy those challenges. Some of the EHR challenges include the lack of routine collection of data types used for frailty measures in EHRs, lack of formal coding to enter functional status measures as structured fields, and incomplete clinical diagnoses captured by EHRs compared to claims (i.e., data quality issues). Despite these challenges, EHRs offer new data types such as unstructured clinical notes that may be used to extract frailty constructs.

Most frailty variables used across different frailty instruments are clinical variables. In a recent structured review of commonly used frailty instruments, nonclinical determinants of health (e.g., social, behavioral, and environmental factors) were found to be often missing in frailty measures [37]. Indeed, more than 87% of all frailty variables used across the reviewed frailty instruments were categorized as clinical variables, and only 12% of these variables were considered behavioral, health services, or social determinants of health. Moreover, some of the social determinants of health may affect frailty more than others, and older adults may have a heterogenic response to various social determinants of health, which requires further research. The inadequacy of frailty variables represented by social determinants of health may limit the use of frailty instruments across larger populations of older adults with different underlying social and behavioral needs.

A major challenge associated with developing EFI measures using EHRs is the lack of frailty-related variables such as functional status, mobility, or cognitive status captured as structured codes within EHRs (e.g., ICD codes). A review found that most variables measured in the frailty instruments are not typically captured within a structured format in EHRs [37], thus limiting their use for population level applications. To help assess the potential feasibility of frailty measures derived using only secondary electronic data (i.e., EHR or claims), the review identified 135 distinct frailty instruments from which multiple unique variables were isolated across behavioral, clinical, social, and environmental health determinants. Out of the 135 instruments, only 22 instruments used variables that could entirely be extracted from EHRs or insurance claims without the need for a survey. However, efforts to integrate various surveys/instruments such as patient-reported outcomes as coded data in EHRs [38] may provide opportunities to increase the value of structured EHR data in capturing additional indicators of frailty.

Unstructured data within EHRs (e.g., clinical notes indicating walking difficulty) can be algorithmically mined to enhance the measurement of frailty on a population level. A prior study assessed the value of unstructured EHR data in identifying several constructs of frailty [39]. The study used a pragmatic natural language processing (NLP) algorithm to identify individuals at high risk of experiencing frailty. The study found that claims and structured EHR data give an incomplete picture of burden related to frailty constructs, and frailty variables are likely to be missed if unstructured data are not analyzed. Using structured claims data, results showed frailty

prevalence ranging from .03% for lack of social support to 8.3% for walking difficulty. Using structured EHR data resulted in similar prevalence rates, ranging from .03% for malnutrition to 7.85% for walking difficulty. However, incorporating unstructured EHR notes, enabled by applying the NLP algorithm, identified considerably higher rates of frailty constructs: absence of fecal control (2.1%, 2.3 times as much as structured claims and EHR data combined), decubitus ulcer (1.4%, 1.7 times), dementia (6.7%, 1.5 times), falls (23.6%, 3.2 times), malnutrition (2.5%, 18.0 times), lack of social support (29.8%, 455.9 times), urinary retention (4.2%, 3.9 times), vision impairment (6.2%, 7.4 times), weight loss (19.2%, 2.9 times), and walking difficulty (36.34%, 3.4 times). Finally, the frailty construct rates extracted from structured data were substantially lower than published epidemiological rates, although adding the NLP results considerably closed this gap [39].

Although not assessed in this project, using unstructured data of EHRs remains a critical step in enhancing the accuracy and predictive power of EFIs in forecasting healthcare outcomes, utilization, and cost. Future projects shall focus on the utility of pragmatic NLP methods in extracting frailty constructs from EHR-embedded clinical notes and survey data. Past studies have shown generalizable NLP methods in extracting various clinical and socio-behavioral constructs of frailty that can be used to increase the completeness of EFIs and consequently improve the prediction of healthcare outcomes and utilization [36, 40, 41].

Using EHR-based Frailty Measures for Risk Stratification

Population health management is an evolving concept within healthcare. Population health management aims to improve the health outcomes of defined populations by identifying, targeting, and modifying health determinants that range from clinical to social and environmental. Within healthcare systems, population health involves the identification of patients and enrollees who are prone to having higher-than-normal risk of undesired outcomes. This may include identification of frail or pre-frail patients using EHR-based frailty measures for early interventions to prevent adverse health outcomes. Depending on the anticipated risk, providers and payers place patients in different risk groups to be targeted by interventions appropriate to their risk level, also known as risk stratification. For example, high-risk patients will be closely followed by a case manager, while low-risk healthy members receive an annual needs-assessment questionnaire. Given the rapid growth of population health management efforts, "risk stratification" activities are becoming essential for prioritizing population-level interventions [42]. EHR-based frailty measures may offer another dimension that can be incorporated into risk stratification approaches.

In risk stratification analytics, healthcare utilization and cost are often considered dependent variables. Various determinants of health (e.g., clinical, behavioral, social) are used as independent variables in risk stratification models to predict the dependent variables. Often, the clinical determinants of health (e.g., diagnosis, medications) are grouped into higher-level concepts to facilitate the statistical modeling process. Traditionally, claims data are used to train

and validate risk stratification models of utilization; however, EHR-extracted determinants of health are increasingly being used for risk modeling [43-48] (e.g., extracting diagnosis data from EHRs to predict cost). Although CFI use in risk prediction is well established, increasingly EHRs are also being used for risk stratification. This has propelled the potential use of EFIs in risk-prediction models (e.g., using EFIs to better predict nursing home admissions [36]).

Although frailty and comorbidities correlate moderately, frailty adds information beyond comorbidities (e.g., functional status) [37]. As frailty measures are increasingly considered a potential predictor of healthcare utilization, [36] including nursing home admission as well as loss of functional independence, frailty indexes are being considered as a risk adjustment factor for clinical quality measurement and reimbursement. The frailty-derived risk adjustment is particularly important to healthcare providers caring for older adult populations (e.g., Medicare patients). Both CFI and EFI can be used, in addition to comorbidity scores, to adjust or control for the underlying 'frailty' and 'functional disability' of a provider's patient population. Similar to comorbidity indexes, the use of a validated CFI and/or EFI will enable incentivizing providers who provide care to a patient population with higher incidence of frailty in addition to the underlying comorbidity scores [49-54]. Of course, like other outcomes used to adjust payments, reimbursing providers based on frailty measures should not turn into incentivizing providers who have worse functional health outcomes while penalizing providers who are doing well in managing and reducing frailty among their patients.

Patient-centered outcomes and health systems researchers should also explore the use of CFI and/or EFI in risk adjusting clinical and operational quality measures. Medical researchers may consider frailty as a factor for adjusting clinical outcomes of a study intervention. Patient-centered outcomes researchers may use frailty as an outcome of interest or a controlling factor. Health services researchers may also use frailty indexes to adjust quality measures for functional disability on a health system level. However, similar to frailty-driven reimbursement adjustments, adjusting quality measures should not lead to penalizing providers who are doing well in controlling frailty among their patient population.

Study Limitations & Potential Factors Influencing the Results

Multiple factors have potentially affected different aspects of this project that may limit the generalizability and implications of the findings. Following are some potential study limitations that may have negatively affected the study findings:

Lack of a Gold Standard Survey: This study did not include a population-level standard survey of frailty, hence the validity of EFI was compared with CFI of the same population. The lack of population-level survey-based frailty measures in EHRs is indeed a common issue with measuring the accuracy of EFI. Only a few population-level claims data sources (e.g., Medicare's Current Beneficiary Survey, or Medicare's Outcome and Assessment Information Set) include individual-level surveys of frailty, and thus can be considered the closest

population-level gold standard measurement of frailty. However, such data sources are often limited in size, usually lack EHR data, and are difficult to access for research purposes (e.g., not all EFI developers can use such data to evaluate the accuracy of their measures).

Closedness of the Studied Health System: As discussed in the Introduction of the report, linked claims and EHRs never overlap perfectly (**Figure 1**). All health systems experience some degree of "EHR data leakage" due to their patients receiving care outside the core group of providers that share a fully interoperable EHR system. Thus, variability in the degree of "closedness" across health systems will lead to variability in "EHR leakage". Given that in the United States, most providers are far less "closed" than those included in many EHR-based measurement development studies (including EFIs), from a policy perspective, one needs to be cautious about measures that have unrealistic data availability expectations. Hence, a one-size-fits-all frailty index will be misleading, as the accuracy of this project's CFI and EFI measures could be highly dependent on the closedness of the health systems represented in this study. For example, in a closed health system (e.g., KPMAS) the EHR data incompleteness is often low and the EHR is expected to contain most encounters found in the internal claims/billing data. Thus, the EFI in such a setting could be highly correlated with CFI, which is generated using internal billing data (if external claims are not used). However, in most "open" health systems (e.g., JHMI) the incompleteness of EHR data could be higher and thus, EFI may not correlate as closely with CFI. More research is needed to identify the relationship of a health system's closedness versus the gap found between EFI and CFI measures.

Uncontrolled Moderator Effects: Many moderating factors may affect the performance of EFI compared to CFI. Although controlling for all possible moderating effects is impossible, any validation of CFI vs. EFI should clearly delineate these potential external factors. For example, the quality of EHR data (e.g., completeness, accuracy, timelines, provenance) may impact the performance of the EFI. EFIs will naturally perform better in health systems with higher-quality data on frailty constructs (e.g., having internal policies incentivizing the capture of frailty constructs in the EHR at each visit). Similarly, Federal and State policies incentivizing the collection of frailty variables may affect the performance of EFI, and perhaps CFI, on a geographic and/or temporal basis. For example, collecting social determinants of health data in EHRs, which can be informative in identifying frailty, has been incentivized in certain regions and/or in specific populations [55,56]. Finally, major disruptions to the traditional healthcare delivery process (e.g., COVID-19 replacing some in-person visits with telehealth visits, thus affecting the collection of frailty constructs in EHRs) can potentially affect frailty indexes, and thus, should be further assessed. Future studies should measure potential effects of external moderators not captured in claims or EHRs on the sensitivity and specificity of EFI and CFI measures in identifying frail patients.

Additional Opportunities for Research

- Assess the value of additional data types collected within EHRs, such as patient-reported outcomes, in identifying frailty.

- Explore the potential skewness of EFI vs. CFI in other patient minority groups, geographies, and care settings.

- Examine how social determinants of health and frailty may be correlated and could increase risk of frailty at an earlier age for socially disadvantaged populations.

- Assess the utility of pragmatic NLP methods in extracting frailty constructs from EHR-embedded clinical notes and survey data.

- Capture additional indicators of frailty through efforts to integrate various surveys/instruments as coded data in EHRs.

- Incorporate frailty indices into risk-stratification approaches for population health management.

- Explore the use of CFI and/or EFI in risk-adjusting clinical and operational quality measures as well as for patient-centered outcomes research.

---

5. Conclusion

This project assessed the feasibility and accuracy of measuring frailty using claims and EHR data across multiple health systems. The findings showed that an established CFI can be applied to EHR's structured data to identify frail patients. The results were replicated in several data sources, multiple health systems, and across different subpopulations (e.g., age and race groups). Overall, the EFI identified a reliable level of frailty cases in EHRs when compared to claims data; however, CFI usually found a higher number of frail patients, given the more complete diagnostic and procedure codes captured in claims.

In conclusion, structured EHR data can be used by healthcare providers to identify frail patients using validated EFIs; however, claims data can identify additional frailty cases on a population level compared to EHR data. Adding claims to EHR data, or using claims data standalone, is recommended when such data exists; however, the advantage of claims data will depend on the closedness of the health system and the completeness of EHR data in capturing all patient encounters. Further research is needed to evaluate the role of the unique EHR features, such as unstructured data, in developing EFIs that have a higher accuracy and recall in identifying patients with frailty.

EFIs can also be used to improve the prediction of various healthcare utilization outcomes, such as inpatient hospitalization, emergency department admission, and nursing home admission. Risk-stratification developers may integrate EFI in their model development process, and population health managers may incorporate EFI in disease management efforts. Future studies should evaluate the interaction of comorbidity indexes with EFIs in predicting various healthcare utilization outcomes, as well as adjusting healthcare costs.

Appendix

Appendix A. Overall Project Tasks and Subtasks

This project included four major tasks, with each task consisting of multiple subtasks. The major tasks included (1) execute legal contracts and collect feedback; (2) link and compare EHR versus claims in multiple data sources; (3) validate and compare the claims-based frailty index (CFI) versus EHR-based frailty index (EFI) using EHR and claims data, and then assessing their value in predicting utilization; and (4) share and disseminate the findings. In summary, Task-1 completed the legal contracts, Task-2 covered the data management tasks, Task 3 included the analytics, and Task 4 focused on dissemination. The tasks and subtasks are summarized as:

- **Task 1. Execute Legal Contracts & Collate Feedback on Study Findings**

    o   Subtask 1.1. Apply and acquire IRB clearance and establish DUAs (Data Use Agreements)

    o   Subtask 1.2. Gather information on the design and implementation of the claims- and EHR-based frailty models

    o   Subtask 1.3. Solicit feedback on the feasibility of EHR-based frailty measures from the EHR Learning Network

- **Task 2. Link and Compare Claims Data with EHR Data**

    o   Subtask 2.1. Check the linkage of EHR and claims in underlying data sources

    o   Subtask 2.2. Formulate and identify denominators/cohorts of patients

    o   Subtask 2.3. Measure the completeness of frailty predictors across EHR and claims data

    o   Subtask 2.4. Compare patient-level comorbidity scores across EHR and claims data

- **Task 3. Evaluate the Accuracy and Concordance of Claims vs. EHR-based Measures of Frailty (CFI vs. EFI)**

    o   Subtask 3.1. Validate and refine claims- and EHR-based frailty measures in predicting functional disability

    o   Subtask 3.2. Assess distribution and concordance of validated claims- and EHR-based measures of frailty

    o   Subtask 3.3. Compare the claims- vs. EHR-based frailty measures across health systems

    o   Subtask 3.4. Measure potential disparities in claims- vs. EHR-based frailty indexes across subgroups

    o   Subtask 3.5. Evaluate the value of claims- vs. EHR-based frailty measures in predicting healthcare utilization

- **Task 4. Dissemination and Final Report**

  - o Subtask 4.1. Presentation at scientific venues

  - o Subtask 4.2. Disseminate standardized code and algorithms

  - o Subtask 4.3. Disseminate the findings in peer-reviewed publications

  - o Subtask 4.4. Prepare and submit the final report

The project tasks followed a specific temporal order, representing both parallel and series of subtasks (Appendix A Figure 1). Task 1 subtasks were executed at different times during the project. Subtasks 1.1 and 1.2 were conducted early in the project (i.e., acquiring all legal requirements such as IRBs and DUAs and understanding the CFI and EFI furnished by AHRQ), while subtask 1.3 was conducted throughout the project timeline (e.g., soliciting feedback from project's EHR Learning Network). Task 2 started as soon as subtask 1.1 was approved/confirmed by the IRB officers and the legal counsel. Three data sources were used for Task 2: Johns Hopkins Medical Institute (JHMI), Optum Labs Data Warehouse (OLDW), and Kaiser Permanente MidAtlantic States (KPMAS). Subtasks 2.1, 2.2, 2.3 and 2.4 were executed consecutively for each data source. The three data sources enabled the project team to start Task 2 as soon as one data source became available, hence reducing the risk of a potential critical pathway in executing the project. Task 3 started as soon as subtask 1.2 was completed. Subtasks 3.1, 3.2, 3.3., 3.4, and 3.5 were completed in series. Finally, Task 4 was initiated after Task 3's subtasks were completed. Subtasks 4.1, 4.2, 4.3, and 4.4. occurred in parallel. Subtask 4.3 (i.e., peer reviewed publication) will continue beyond the end of the project until the manuscripts are reviewed, accepted, and published by scientific journals.
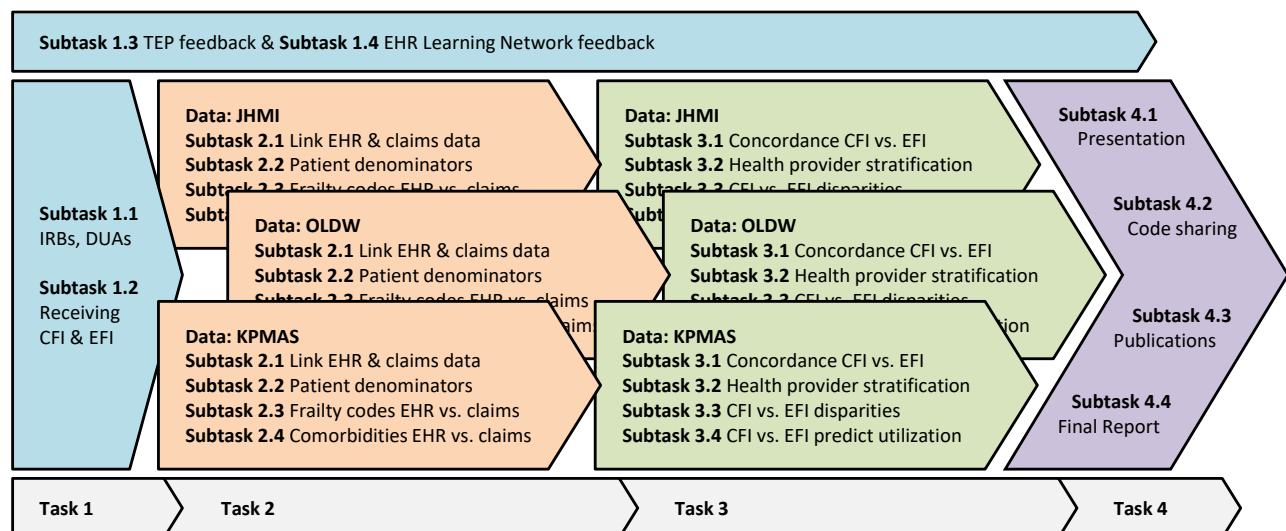
Appendix A Figure 1. Temporal order of project tasks from start (left) to finish (right)

Appendix B. Data Sources

This project included three data sources, with one representing more than one health system. These data sources included:

- **JHMI**

    o *Health System Specs*: The Johns Hopkins Medical Institute (JHMI) includes 5 hospitals and 30+ outpatient clinics in Maryland. Despite the number of outpatient clinics, JHMI is often used as a tertiary care provider by a majority of patients. Most patients with an inpatient admission lack outpatient care provided by the JHMI network (i.e., JHMI is not a locked-in system; Appendix B Figure 1).

    o *Denominator Notes*: JHMI data used in this project was comprised of ~160k patients who had both EHR and claims records.

    o *Data Notes*: JHMI's EHR data included clinical data captured across all hospitals and outpatient clinics of JHMI in Maryland. JHMI's claims data were limited to data provided by the Johns Hopkins Health Care (JHHC; the insurance arm of JHMI), which included a select subpopulation of JHMI patients with Medicaid (i.e., Priority Partners) and commercial (i.e., Employer Health Plan) health plans. JHMI claims data represent all services provided to JHMI patients across all healthcare providers, both inside and outside of JHMI. JHMI's claims data did not include Medicare data, hence results for 65+ year old patients are not reliable for the JHMI population.

- **OLDW**

    o *Health System Specs*: The Optum Labs Data Warehouse (OLDW) included EHR and claims records of more than 55 healthcare provider networks. Health systems represented in OLDW have varying levels of patient lock-in features, with some having minimal outpatient data for their patients and others having much broader data collected from their patients across various health settings, including outpatient care (Appendix B Figure 1).

    o *Denominator Notes*: More than 29m patients have both the EHR and claims data in OLDW.

    o *Data Notes*: OLDW claims data included only Medicare Advantage (MA) and commercial claims. Some MA patients were dual eligible and had Medicaid coverage. None of the patients had only Medicaid coverage. OLDW claims data represent all services provided to OLDW patients across all healthcare providers, both inside and outside of OLDW health systems.

- **KPMAS**

    o *Health System Specs*: The Kaiser Permanente Mid-Atlantic States (KPMAS) is a health system and a health plan. KPMAS patients receive the majority of healthcare services

from KPMAS providers, except select inpatient and consulting services in Maryland that are not provided by KPMAS (Appendix B Figure 1).

o   *Denominator Specs*: More than 200k patients had both the EHR and external (non-KPMAS) claims data.

o   *Data Notes*: KPMAS claims data only includes encounters originated out of the KP network. KPMAS claims data do not include encounters that occur within the KP system, as KP is both a provider and a payer, thus uses an internal billing mechanism for internal encounters. The internal billing data are not included in KPMAS claims data. This makes KPMAS claims data inherently different from JHMI or OLDW claims data, which include encounters occurring within JHMI and OLDW health systems.
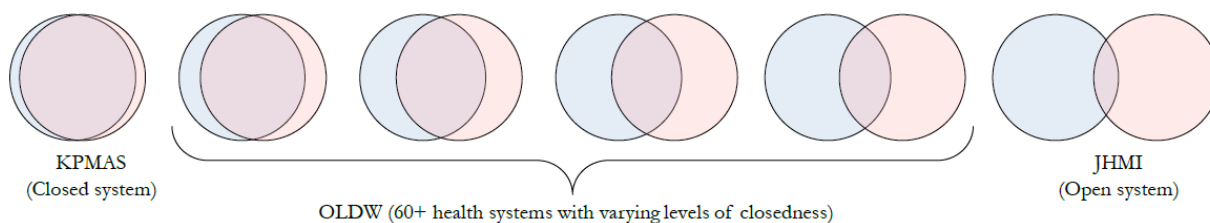
Appendix B Figure 1. Schematic representation of the level of closedness across health systems included in this project

(red: EHR; blue: claims; purple: overlap of EHR and claims)

Appendix C. Additional Results

- **Task 2: Link and Compare Claims Data with EHR**

Subtask 2.1: Check the linkage of EHR and claims in underlying data sources

**Appendix C Table 1. Checking the data quality of demographic data between
EHR and claims for JHMI data across all years**

Data Cleaning Processes	Claims Patients	EHR Patients
Original N	160,048 (100%)	2,349,255 (100%)
Missing Birth Year	0 (0%)	0 (0%)
Missing or Unknown or Other Sex	0 (0%)	1,909 (0.08%)
Other Race *	908 (0.56%)	262,383 (11.17%)
Unknown Race *	79,566 (49.71%)	129,324 (5.50%)
Final N	160,048 (100%)	2,347,346 (99.92%)

* No exclusions were made based on Race: "White," "African American," "Asian," and "Other/Unknown"

**Appendix C Table 2. Checking the data quality of demographic data between
EHR and claims for OLDW data across all years**

Data Cleaning Processes	Claims Patients	EHR Patients
Original N	126,363,281 (100%)	100,184,788 (100%)
Missing Birth Year	47,371 (0%)	637,841 (0.6%)
Missing or Unknown or Other Sex	488,141 (0.4%)	724,898 (0.7%)
Other Race *	-	-
Unknown Race *	21,954,770 (17.4%)	29,745,309 (29.7%)
Final N	125,863,403 (99.6%)	99,128,861 (98.9%)

* No exclusions were made based on Race: "White," "African American," "Asian," and "Other/Unknown"

Subtask 2.2: Formulate and identify denominators/cohorts of patients

Appendix C Table 3. Age and race stratification of the final JHMI denominator per calendar year (2016 to 2019)

Strata	2016	2017	2018	2019
Total N	57,797 (100.0%)	60,792 (100.0%)	60,925 (100.0%)	51,469 (100.0%)
Male	19,350 (33.48%)	20,474 (33.68%)	20,644 (33.88%)	17,201 (33.42%)
Age 18-49	42,876 (74.18%)	44,790 (73.68%)	44,773 (73.49%)	37,097 (72.08%)
Age 50-64	13,988 (24.20%)	14,997 (24.67%)	15,034 (24.68%)	13,273 (25.79%)
Age 65-79*	913 (1.58%)	983 (1.62%)	1,092 (1.79%)	1,076 (2.09%)
Age 80+*	20 (0.03%)	22 (0.04%)	26 (0.04%)	23 (0.40%)
Race White	24,930 (43.13%)	25,944 (42.68%)	25,655 (42.11%)	21,259 (41.30%)
Race Black	24,563 (42.50%)	25,818 (42.47%)	26,225 (43.04%)	22,688 (44.08%)
Race Asian	3,884 (6.72%)	4,171 (6.86%)	4,226 (6.94%)	3,668 (7.13%)
Race Other	4,148 (7.18%)	4,486 (7.38%)	4,422 (7.26%)	3,628 (7.05%)
Race Unknown	272 (0.47%)	373 (0.61%)	397 (0.65%)	226 (0.44%)

* Note that 65+ age groups have a small sample size in JHMI data due to missing Medicare data; thus,

should not be interpreted as generalizable data

Appendix C Table 4. Age and race stratification of the final OLDW denominator per calendar year (2016 to 2019)

Strata	2016	2017	2018	2019
Total N	2,478,649 (100.0%)	2,486,447 (100.0%)	2,278,947 (100.0%)	2,076,817 (100.0%)
Male	1,028,100 (41.50%)	1,039,222 (41.80%)	956,050 (42.00%)	874,729 (42.10%)
Age 18-49	1,082,121 (46.00%)	1,046,505 (44.20%)	942,965 (43.00%)	856,625 (42.80%)
Age 50-64	691,809 (29.40%)	672,951 (28.40%)	611,140 (27.90%)	544,404 (27.20%)
Age 65-79	427,737 (18.20%)	482,281 (20.40%)	472,112 (21.50%)	446,886 (22.30%)
Age 80+	148,636 (6.30%)	167,768 (7.10%)	164,653 (7.50%)	154,895 (7.70%)
Race White	1,938,910 (78.20%)	1,955,236 (78.60%)	1,803,059 (79.10%)	1,652,117 (79.60%)
Race Black	135,918 (5.50%)	137,621 (5.50%)	124,588 (5.50%)	113,871 (5.50%)
Race Asian	73,773 (3.00%)	66,145 (2.70%)	58,754 (2.60%)	55,093 (2.70%)
Race Other/Unknown*	330,048 (13.30%)	327,445 (13.20%)	292,546 (12.80%)	255,736 (12.30%)

* Note that "Other" and "Unknown" races are combined in OLDW data

**Appendix C Table 5. Age and race stratification of the final KPMAS denominator
per calendar year (2016 to 2019)**

Strata	2016	2017	2018	2019
Total N*	110,979 (100.0%)	114,692 (100.0%)	119,006 (100.0%)	122,723 (100.0%)
Male	42,895 (38.65%)	44,571 (38.86%)	45,981 (38.64%)	47,300 (38.54%)
Age 18-49	51,999 (46.85%)	52,824 (46.06%)	54,784 (46.03%)	56,167 (45.77%)
Age 50-64	28,794 (25.95%)	29,507 (25.73%)	30,690 (25.79%)	31,594 (25.74%)
Age 65-79	23,419 (21.10%)	25,155 (21.93%)	25,919 (21.78%)	27,035 (22.03%)
Age 80+	6,767 (6.10%)	7,206 (6.28%)	7,613 (6.40%)	7,927 (6.46%)
Race White	38,610 (34.79%)	38,763 (33.80%)	39,480 (33.17%)	40,198 (32.76%)
Race Black	46,275 (41.70%)	48,804 (42.55%)	49,518 (41.61%)	52,528 (42.80%)
Race Asian	8,062 (7.26%)	8,584 (7.48%)	10,135 (8.52%)	10,321 (8.41%)
Race Other	14,037 (12.65%)	14,513 (12.65%)	16,395 (13.78%)	16,917 (13.78%)
Race Unknown	3,942 (3.55%)	3,962 (3.45%)	3,419 (2.87%)	2,705 (2.20%)

*Note: KPMAS claims data only includes encounters/services provided out of the KPMAS system and
does not include internal billing data*

**Appendix C Table 6. Prevalence of frailty predictors across JHMI's EHR and claims data
per calendar year (2016 to 2019)**

	2016			2017			2018			2019			2016	2017	2018	2019
Pr.*	E%	C%	E+C%	E%	C%	E+C%	E%	C%	E+C%	E%	C%	E+C%		Ratio: C/E		
1	0.8	1.7	1.9	1.0	2.2	2.4	0.8	1.9	2.2	0.9	1.9	2.3	2.17	2.12	2.32	2.07
2	1.8	1.1	1.4	0.5	0.7	0.8	0.4	0.6	0.7	0.5	0.7	0.8	1.06	1.45	1.48	1.48
3	4.9	8.1	9.3	4.6	8.2	9.3	4.8	8.1	9.5	4.8	8.6	9.7	1.66	1.78	1.68	1.76
4	0.8	1.0	1.2	0.9	1.2	1.3	0.9	1.1	1.3	1.0	1.1	1.3	1.26	1.35	1.25	1.18
5	7.3	11.7	12.4	8.1	12.5	13.2	7.8	11.6	12.9	9.1	13.3	14.3	1.61	1.53	1.50	1.46
6	2.1	3.5	3.7	2.1	3.6	3.8	2.1	3.5	3.7	2.4	3.7	3.9	1.68	1.69	1.68	1.57
7	6.5	9.2	9.6	6.7	9.7	10.2	6.7	8.9	10.2	7.3	9.1	10.2	1.41	1.44	1.32	1.25
8	10.1	13.6	14.1	10.9	14.3	14.9	10.7	13.3	14.8	11.8	14.2	15.5	1.34	1.31	1.24	1.20
9	23.7	34.0	36.3	26.3	36.8	39.2	26.4	32.9	38.9	28.7	35.9	40.5	1.43	1.40	1.25	1.25
10	5.0	7.9	8.7	6.0	9.2	10.1	5.9	7.5	9.8	5.8	7.7	9.5	1.56	1.54	1.27	1.33

*Only the first 10 predictors are shown (out of 93 diagnoses and procedures used as predictors in Kim's model): (1) Intestinal
infectious diseases; (2) Other bacterial diseases; (3) Mycoses; (4) Malignant neoplasm of other/unspecified sites; (5) Benign
neoplasms; (6) Neoplasms of uncertain behavior; (7) Disorders of thyroid; (8) Diseases of other endocrine glands;
(9) Other metabolic and immunity disorders; and (10) Other and unspecified anemias.
C = Claims, Pr. = [Frailty] Predictor, and E = EHR.*

Appendix C Table 7. Prevalence of frailty predictors across KPMAS' EHR and claims data per calendar year (2016 to 2019)

Pr.*	2016 E%	2016 C%	2016 E+C%	2017 E%	2017 C%	2017 E+C%	2018 E%	2018 C%	2018 E+C%	2019 E%	2019 C%	2019 E+C%	Ratio: C/E 2016	Ratio: C/E 2017	Ratio: C/E 2018	Ratio: C/E 2019
1	1.3	0.9	2.0	1.4	0.9	2.0	1.4	0.8	2.0	1.4	0.9	2.0	0.69	0.68	0.59	0.64
2	0.5	0.5	0.9	0.6	0.5	0.9	0.6	0.5	1.0	0.5	0.5	0.9	0.90	0.83	0.89	1.02
3	8.7	2.2	10.4	8.7	2.1	10.3	8.8	2.0	10.3	8.7	1.6	9.9	0.25	0.24	0.23	0.19
4	2.3	1.7	2.6	2.3	1.7	2.6	2.3	1.7	2.7	2.3	1.7	2.8	0.75	0.76	0.71	0.74
5	8.2	4.3	10.8	8.5	4.7	11.3	8.4	4.6	11.2	8.5	4.0	10.6	0.52	0.55	0.55	0.47
6	2.9	1.2	3.7	2.7	1.5	3.8	2.9	1.4	3.9	2.9	0.9	3.4	0.42	0.56	0.49	0.32
7	10.9	5.5	12.2	11.3	5.3	12.5	11.9	5.5	13.1	12.3	5.5	13.4	0.50	0.47	0.46	0.45
8	20.3	14.4	22.6	20.9	14.3	22.9	21.1	14.5	23.2	22.6	14.2	24.5	0.71	0.68	0.68	0.63
9	42.2	25.6	50.0	44.1	26.0	51.6	45.6	26.7	53.0	45.0	26.0	52.6	0.61	0.59	0.59	0.58
10	6.4	6.8	10.4	6.6	6.8	10.6	7.3	7.5	11.6	8.3	7.8	12.5	1.05	1.02	1.01	0.94

Only the first 10 predictors are shown (out of 93 diagnoses and procedures used as predictors in Kim's model): (1) Intestinal infectious diseases; (2) Other bacterial diseases; (3) Mycoses; (4) Malignant neoplasm of other/unspecified sites; (5) Benign neoplasms; (6) Neoplasms of uncertain behavior; (7) Disorders of thyroid; (8) Diseases of other endocrine glands; (9) Other metabolic and immunity disorders; and (10) Other and unspecified anemias.

C = Claims, Pr. = [Frailty] Predictor, and E = EHR.

Appendix C Table 8. Comorbidity scores across JHMI's EHR and claims data per calendar year (2016 to 2019)

Comorbidity Score	2016 E	2016 C	2016 E+C	2017 E	2017 C	2017 E+C	2018 E	2018 C	2018 E+C	2019 E	2019 C	2019 E+C	Ratio: C/E 2016	Ratio: C/E 2017	Ratio: C/E 2018	Ratio: C/E 2019
# CCSR conditions	9.0	13.6	15.1	10.2	14.8	16.5	10.0	13.1	16.3	10.8	13.9	16.8	1.51	1.46	1.30	1.29
# CCSR body systems	4.5	6.2	6.7	4.9	6.6	7.1	4.7	6.1	7.0	5.1	6.5	7.2	1.38	1.35	1.30	1.28
Mean Elixhauser	-0.4	-0.3	-0.4	-0.4	-0.3	-0.4	-0.5	-0.1	-0.4	-0.6	-0.2	-0.5	0.61	0.57	0.13	0.28
% With 0 Elixhauser	51.3	39.1	36.2	48.3	36.9	34.0	49.9	41.6	34.8	45.8	38.7	32.2	0.76	0.76	0.83	0.85
Mean Charlson	0.6	0.8	0.9	0.7	0.9	1.0	0.7	0.8	0.9	0.7	0.8	1.0	1.34	1.34	1.20	1.20
% With 0 Charlson	73.8	66.4	64.6	71.6	64.8	63.0	72.5	68.6	63.6	70.7	66.9	62.4	0.90	0.90	0.95	0.95

E = EHR; and C = Claims.

Appendix C Table 9. Comorbidity scores across KPMAS' EHR and claims data per calendar year (2016 to 2019)

Comorbidity Score	2016 E	2016 C	2016 E+C	2017 E	2017 C	2017 E+C	2018 E	2018 C	2018 E+C	2019 E	2019 C	2019 E+C	2016 Ratio: C/E	2017 Ratio: C/E	2018 Ratio: C/E	2019 Ratio: C/E
# CCSR conditions	14.8	8.8	19.0	15.5	8.9	19.6	15.9	9.2	20.2	16.2	9.6	20.5	0.59	0.57	0.58	0.59
# CCSR body sys.	6.7	4.3	7.7	6.9	4.3	7.9	7.00	4.4	8.00	7.1	4.6	8.0	0.63	0.62	0.62	0.63
Mean Elixhauser	-0.1	-0.0	-0.0	-0.2	-0.0	-0.1	-0.2	-0.0	-0.1	-0.3	-0.0	-0.2	0.50	0.22	0.15	0.08
% With 0 Elixhauser	36.2	51.7	27.8	34.3	51.7	26.7	32.3	50.8	25.2	30.7	52.4	24.1	1.43	1.51	1.57	1.70
Mean Charlson	1.2	0.9	1.5	1.3	1.0	1.6	1.3	1.0	1.6	1.4	1.0	1.7	0.75	0.75	0.74	0.72
% With 0 Charlson	58.1	68.5	51.9	57.4	68.4	51.8	56.9	68.3	51.3	55.6	69.0	50.5	1.18	1.19	1.20	1.24

E = EHR; and C = Claims.

- ## Task 3: Evaluate the Accuracy and Concordance of Claims vs. EHR-based Measures of Frailty

Subtask 3.2: Assess distribution and concordance of claims- and EHR-based measures of frailty

Appendix C Table 10. Rate of frailty identified using EHR, claims, and EHR plus claims in 2016 across different age groups

Group	E	C	E+C	EHR Added Value
JHMI				
18-49	1.4	3.4	4.0	16%
50-64	5.5	10.8	12.1	10%
65-79	-	-	-	-
80+	-	-	-	-
OLDW				
18-49	0.3	1.4	1.6	13%
50-64	1.6	5.1	5.7	11%
65-79	4.9	12.9	14.1	9%
80+	12.8	31.5	34.1	8%
KPMAS				
18-49	0.5	1.6	2.2	27%
50-64	3.3	6.3	8.8	28%
65-79	9.2	12.3	18.2	32%
80+	22.2	29.3	42.0	30%

E = EHR; and C = Claims.

Subtask 3.3: Compare the claims- vs. EHR-based frailty measures across health systems
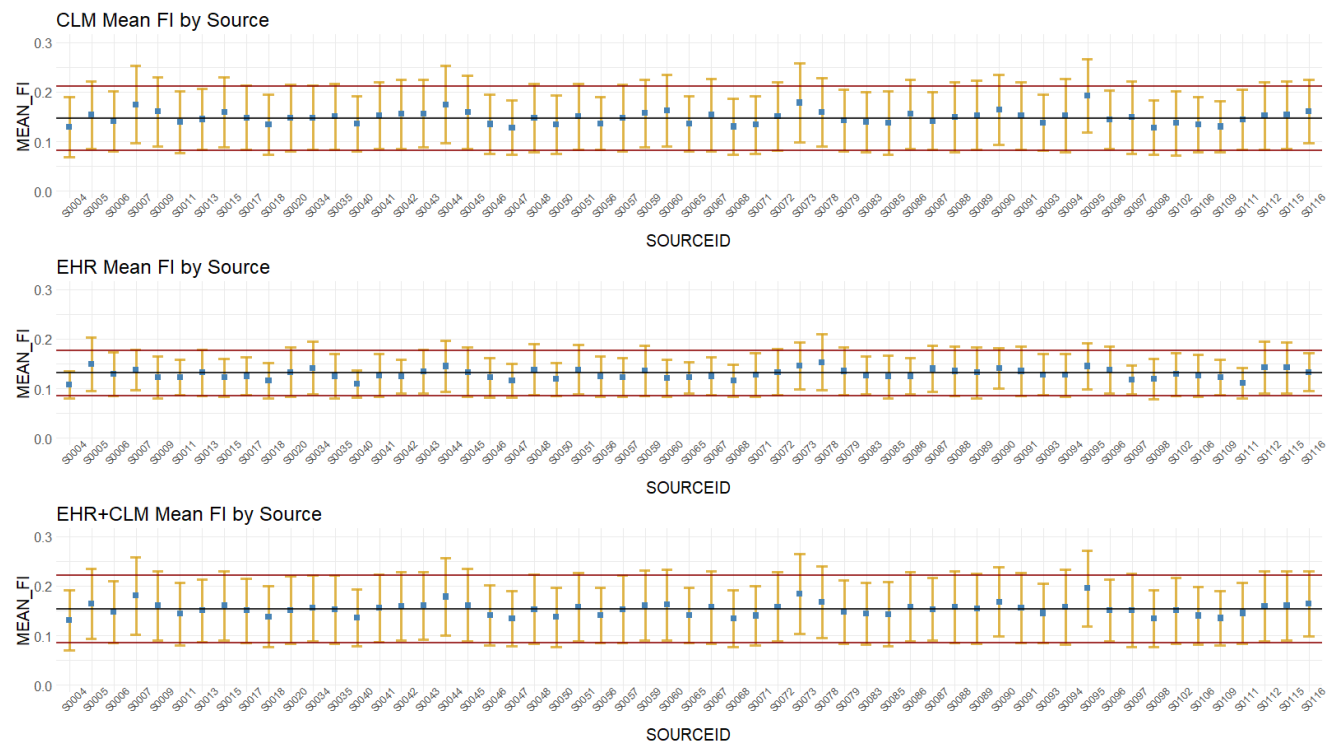
Appendix C Figure 1. Mean and standard deviation of frailty index calculated using claims data (top), EHR data (middle), and both EHR and claims (bottom) across 55 health systems

FI: Frailty Index (score); red and blue lines indicate the average and standard deviation of frailty index within each data type across health systems. Note: This diagram depicts the raw frailty index/score; please refer to the main text for the percentage of frail patients identified in each data source.

Subtask 3.4: Measure potential disparities in claims- vs. EHR-based frailty indexes across subgroups

Appendix C Table 11. Concordance of EHR-based and claims-based frailty for each data source stratified by major insurance type

Group		2016		2017		2018		2019
KPMAS	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**	**phi**	**% Exact**
Non-CMS	0.301 *	97.0 *	0.309 *	97.1 *	0.321 *	96.8 *	0.348 *	96.6 *
Medicaid	0.291 *	96.5 *	0.231 *	96.2 *	0.274 *	95.8 *	0.294 *	95.9 *
Medicare	0.422 *	85.9 *	0.427 *	85.9 *	0.418 *	85.2 *	0.421 *	84.5 *

* p < .001; Non-CMS includes commercial and other types of non-CMS insurance plans

Subtask 3.5: Evaluate the value of claims- vs. EHR-based frailty measures in predicting healthcare utilization

Appendix C Table 12. Denominators used for predictive analysis across data sources and years

Criteria	2016	2017	2018	2019
JHMI				
Original N	57,797	60,792	60,925	51,469
Pregnancy N	8936 (15.46%)	10,072 (16.57%)	10,159 (16.67%)	8521 (16.56%)
N	48,861	50,720	50,766	42,948
OLDW				
Original N	2,472,846	2,466,412	2,247,574	2,055,789
Pregnancy N	87,289 (3.53%)	81,196 (3.29%)	71,395 (3.18%)	64,953 (3.16%)
N	2,385,557	2,385,216	2,176,179	1,990,836
KPMAS				
Original N	110,979	114,692	119,006	122,723
Pregnancy N	12,988 (11.70%)	13,662 (11.91%)	14,451 (12.14%)	15,088 (12.29%)
N	97,991	101,030	104,555	107,635

**Appendix C Table 13. Utilization markers across data sources and years**

	2016			2017			2018			2019		
JHMI	E	C	E+C	E	C	E+C	E	C	E+C	E	C	E+C
% With 1+ ED admit	20.3	31.6	34.5	20.7	32.3	35.1	18.7	30.6	33.1	19.1	30.1	32.9
% With 1+ IP admit	6.2	11.6	12.4	6.7	11.8	12.6	5.8	11.4	12.0	6.00	11.5	12.1
% With 1+ NH admit	0.0	0.8	0.8	0.0	1.1	1.1	0.0	1.1	1.1	0.0	1.0	1.0
OLDW	E	C	E+C	E	C	E+C	E	C	E+C	E	C	E+C
% With 1+ ED admit	11.7	17.6	25.3	12.4	18.0	25.7	11.2	18.4	25.2	11.5	19.6	26.4
% With 1+ IP admit	5.3	10.3	11.2	5.6	10.7	11.6	5.1	10.7	11.4	5.1	10.5	11.3
% With 1+ NH admit	0.4	2.7	2.7	0.5	3.0	3.1	0.5	3.1	3.1	0.5	3.2	3.2
KPMAS	E	C	E+C	E	C	E+C	E	C	E+C	E	C	E+C
% With 1+ ED admit	0.1	49.0	49.0	0.1	50.3	50.3	0.1	53.1	53.1	0.0	49.5	49.5
With 1+ IP admit	8.9	24.6	27.1	9.7	24.5	27.4	9.9	25.0	28.1	9.5	23.7	26.9
With 1+ NH admit	1.7	0.9	2.3	1.8	0.9	2.4	2.0	0.8	2.6	2.0	0.8	2.6

E = EHR; and C = Claims.

**Appendix C Table 14. ORs of frailty indexes in predicting prospective outcomes using JHMI, OLDW, and KPMAS data**

	JHMI			OLDW			KPMAS		
	EFI	**CFI**	**CEFI**	**EFI**	**CFI**	**CEFI**	**EFI**	**CFI**	**CEFI**
IP	3.20	3.93	3.75	1.53	2.09	2.12	2.03	2.53	2.50
ED	2.77	2.95	2.92	1.65	1.80	1.81	2.19	2.48	2.47
NH	2.67	7.10	6.69	2.02	3.78	3.91	2.28	4.98	4.80

CFI = claims-based frailty index; EFI = EHR-based frailty index; and CEFI: claims-EHR-based frailty index
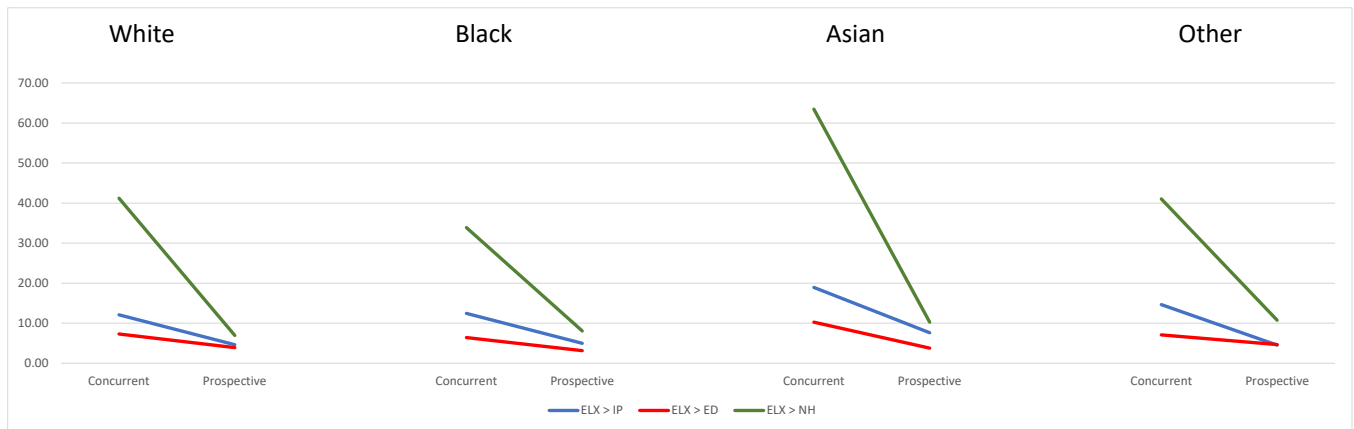
Appendix C Figure 2. Parallel plot of odds ratios of frailty index in race stratified concurrent and prospective logistic regression models of utilization prediction using JHMI data

ED: Emergency department; ELX: Elixhauser comorbidity index; IP: Inpatient; and NH: Nursing home.
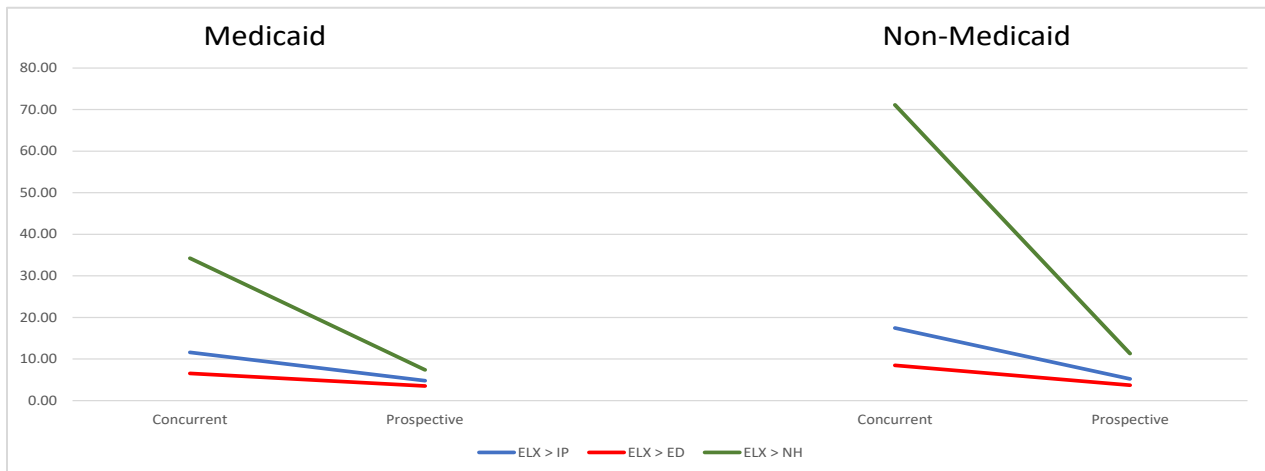
Appendix C Figure 3. Parallel plot of odds ratios of frailty index in insurance type stratified concurrent and prospective models of utilization prediction using JHMI data

ED: Emergency department; ELX: Elixhauser comorbidity index; IP: Inpatient; and NH: Nursing home.

Appendix D. Bibliography

1. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc. 2012; 60: 1487-1492

2. Song X, Mitnitski A, Rockwood K. Prevalence and 10-year outcomes of frailty in older adults in relation to deficit accumulation. J Am Geriatr Soc. 2010; 58: 681-687

3. Fried LP, Tangen CM, Walston J, et al.; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001; 56: M146–M156

4. Rockwood K, Mitnitski A. Frailty defined by deficit accumulation and geriatric medicine defined by frailty. Clin Geriatr Med. 2011; 27: 17-26

5. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. Scientific World J. 2001; 1: 323-336

6. Mitnitski A, Song X, Skoog I, et al. Relative fitness and frailty of elderly men and women in developed countries and their relationship with mortality. J Am Geriatr Soc. 2005; 53: 2184-2189

7. Rockwood K, Mitnitski A, Song X, Steen B, Skoog I. Long-term risks of death and institutionalization of elderly people in relation to deficit accumulation at age 70. J Am Geriatr Soc. 2006; 54: 975-979

8. Rockwood K, Andrew M, Mitnitski A. A comparison of two approaches to measuring frailty in elderly people. J Gerontol A Biol Sci Med Sci. 2007; 62: 738-743

9. Davidoff AJ, Zuckerman IH, Pandya N, et al. A novel approach to improve health status measurement in observational claims-based studies of cancer treatment and outcomes. J Geriatr Oncol. 2013; 4(2): 157-165

10. Faurot KR, Jonsson Funk M, Pate V, et al. Using claims data to predict dependency in activities of daily living as a proxy for frailty. Pharmacoepidemiol Drug Saf. 2015; 24(1): 59-66

11. Cuthbertson CC, Kucharska-Newton A, Faurot KR, et al. Controlling for frailty in pharmacoepidemiologic studies of older adults: validation of an existing Medicare claims-based algorithm. Epidemiology. 2018; 29(4): 556-561

12. Segal JB, Chang HY, Du Y, Walston JD, Carlson MC, Varadhan R. Development of a claims-based frailty indicator anchored to a well-established frailty phenotype. Med Care. 2017; 55(7): 716-722

13. Segal JB, Huang J, Roth DL, Varadhan R. External validation of the claims-based frailty index in the national health and aging trends study cohort. Am J Epidemiol. 2017; 186(6): 745-747

14. Ben-Shalom Y, Stapleton DC. Predicting Disability among Community-Dwelling Medicare Beneficiaries Using Claims-Based Indicators. Health Serv Res. 2016; 51(1): 262-281

15. Kim DH, Schneeweiss S, Glynn RJ, Lipsitz LA, Rockwood K, Avorn J. Measuring frailty in Medicare data: development and validation of a claims-based frailty index. J Gerontol A Biol Sci Med Sci. 2018; 73(7): 980-987

16. Kim DH, Glynn RJ, Avorn J, et al. Validation of a claims-based frailty index against physical performance and adverse health outcomes in the health and retirement study. J Gerontol A Biol Sci Med Sci. 2019; 74: 1271-1276

17. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. Ann Intern Med. 2008; 149(10): 751-760

18. Heintzman J, Bailey SR, Hoopes MJ, et al. Agreement of Medicaid claims and electronic health records for assessing preventive care quality among adults. J Am Med Inform Assoc. 2014; 21(4): 720-724

19. Office of the National Coordinator for Health Information Technology. Non-federal Acute Care Hospital Electronic Health Record Adoption. 2017. Accessed Jul 11, 2020. https://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-EHR-Adoption.php

20. Office of the National Coordinator for Health Information Technology. Office-based Physician Electronic Health Record Adoption. Health IT Quick-Stat #50 [Internet]. 2019. Available from: https://dashboard.healthit.gov/quickstats/pages/physician-ehr-adoption-trends.php.

21. Anzaldi LJ, Davison A, Boyd CM, Leff B, Kharrazi H. Comparing clinician descriptions of frailty and geriatric syndromes using electronic health records: a retrospective cohort study. BMC geriatrics. 2017; 17(1): 248

22. Lekan DA, Wallace DC, McCoy TP, Hu J, Silva SG, Whitson HE. Frailty assessment in hospitalized older adults using the electronic health record. Biol Res Nurs. 2017; 19(2): 213-228

23. Pajewski NM, Lenoir K, Wells BJ, Williamson JD, Callahan KE. Frailty screening using the electronic health record within a Medicare accountable care organization. J Gerontol A Biol Sci Med Sci. 2019; 74(11): 1771-1777

24. Shao Y, Mohanty AF, Ahmed A, et al. Identification and Use of Frailty Indicators from Text to Examine Associations with Clinical Outcomes Among Patients with Heart Failure. AMIA Annu Symp Proc. 2017; 2016: 1110-1118

25. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age and ageing. 2016; 45(3): 353-360

26. Optum. The benefit of using both claims data and electronic medical record data in health care analysis. 2012. Accessed Jul 9, 2020. https://www.optum.com/content/dam/optum/resources/whitePapers/Benefits-of-using-both-claims-and-EMR-data-in-HC-analysis-WhitePaper-ACS.pdf

27. West SL, Johnson W, Visscher W, et al. The challenges of linking health insurer claims with electronic medical records. Health Informatics J. 2014; 20: 22-34

28. Kharrazi H, Lasser EC, Yasnoff WA, Loonsk J, Advani A, Lehmann HP, et al. A proposed national research and development agenda for population health informatics: summary recommendations from a national expert workshop. J Am Med Inform Assoc. 2017; 24(1): 2-12.

29. Kan HJ, Kharrazi H, Leff B, Boyd C, Davison A, Chang HY, et al. Defining and assessing geriatric risk factors and associated health care utilization among older adults using claims and electronic health records. Medical care. 2018; 56(3): 233-239

30. Office of the Assistant Secretary for Planning and Evaluation (ASPE). Validating and expanding claims-based algorithms of frailty and functional disability for value-based care and payment. Accessed Jul 12, 2020. https://aspe.hhs.gov/validating-and-expanding-claims-based-algorithms-frailty-and-functional-disability-value-based-care-and-payment

31. Office of the Assistant Secretary for Planning and Evaluation (ASPE). Electronic Health Record (EHR) Learning Network on Algorithms to Identify Frailty and Functional Disability. Accessed Aug 8, 2022. https://aspe.hhs.gov/electronic-health-record-ehr-learning-network-algorithms-identify-frailty-functional-disability

32. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. 2011; 64: 749-759

33. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987; 40: 373-383

34. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care. 1998; 36: 8-27

35. AHRQ HCUP. Clinical Classifications Software (CCS) for ICD-9-CM. Accessed Jul 11, 2020. https://hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp

36. Kan HJ, Kharrazi H, Leff B, Boyd C, Davison A, Chang HY, et al. Defining and assessing geriatric risk factors and associated health care utilization among older adults using claims and electronic health records. Medical care. 2018; 56(3): 233-239

37. Bery AK, Anzaldi LJ, Boyd CM, Leff B, Kharrazi H. Potential Value of electronic health records in capturing data on geriatric frailty for population health. Archives of Gerontology and Geriatrics. 2020; 91(104224)

38. Wu AW, Kharrazi H, Boulware LE, Snyder CF. Measure once, cut twice--adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. J Clin Epidemiol. 2013 Aug; 66 (8 Suppl): S12-20.

39. Kharrazi H, Anzaldi LJ, Hernandez L, Davison A, Boyd CM, Leff B, et al. The value of unstructured electronic health record data in geriatric syndrome case identification. Journal of the American Geriatrics Society. 2018; 66(8): 1499-1507

40. Chen T, Dredze M, Weiner JP, Kharrazi H. Identifying vulnerable older adult populations by contextualizing geriatric syndrome information in clinical notes of electronic health records. J Am Med Inform Assoc. 2019; 26(8-9): 787-795

41. Chen T, Dredze M, Weiner JP, Hernandez L, Kimura J, Kharrazi H. Extraction of geriatric syndromes from electronic health record clinical notes: assessment of statistical natural language processing methods. JMIR Med Inform. 2019; 7(1): e13039

42. Kharrazi H, Lasser EC, Yasnoff WA, Loonsk J, Advani A, Lehmann HP, et al. A proposed national research and development agenda for population health informatics: summary recommendations from a national expert workshop. J Am Med Inform Assoc. 2017; 24(1): 2-12.

43. Kharrazi H, Chi W, Chang HY, Richards TM, Gallagher JM, Knudson SM, et al. Comparing population-based risk-stratification model performance using demographic, diagnosis and medication data extracted from outpatient electronic health records versus administrative claims. Medical care. 2017; 55(8): 789-796

44. Kharrazi H, Weiner JP. A practical comparison between the predictive power of population-based risk stratification models using data from electronic health records versus administrative claims: Setting a baseline for future EHR-derived risk stratification models. Medical care. 2018; 56(2): 202-203

45. Chang HY, Richards TM, Shermock KM, Elder Dalpoas S, H JK, Alexander GC, et al. Evaluating the impact of prescription fill rates on risk stratification model performance. Medical care. 2017; 55(12): 1052-1060

46. Lemke KW, Gudzune KA, Kharrazi H, Weiner JP. Assessing markers from ambulatory laboratory tests for predicting high-risk patients. The American journal of managed care. 2018; 24(6): e190-e195

47. Kharrazi H, Chang HY, Heins S, Weiner JP, Gudzune K. Enhancing the prediction of healthcare costs and utilization by including outpatient BMI values to diagnosis-based risk models. Med Care. 2018; 56 (12): 1042-1050

48. Kharrazi H, Chang HY, Weiner JP, Gudzune KA. Assessing the added value of blood pressure information derived from electronic health records in predicting health care cost and utilization. Popul Health Manag. 2022 Jun; 25(3): 323-334.

49. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet (London, England). 2013; 381(9868): 752-762

50. Studenski S, Hayes RP, Leibowitz RQ, Bode R, Lavery L, Walston J, et al. Clinical Global Impression of Change in Physical Frailty: development of a measure based on clinical judgment. Journal of the American Geriatrics Society. 2004; 52(9): 1560-1566

51. Sheehan OC, Leff B, Ritchie CS. Frailty tools are not yet ready for prime time in high-risk identification. J Hosp Med. 2019; 14(7): 450

52. Rockwood K, Theou O, Mitnitski A. What are frailty instruments for? Age and ageing. 2015; 44(4): 545-547

53. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. The journal of nutrition, health & aging. 2012; 16(7): 601-608

54. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age and ageing. 2016; 45(3): 353-360

55. Cantor MN, Thorpe L. Integrating data on social determinants of health into electronic health records. Health Aff (Millwood). 2018 Apr; 37(4): 585-590.

56. Pandya CJ, Hatef E, Wu J, Richards T, Weiner JP, Kharrazi H. Impact of social needs in electronic health records and claims on health care utilization and costs risk-adjustment models within Medicaid population. Popul Health Manag. 2022