

# HDAI 1-year Mortality Models

## Overview

This README file describes how to use the provided coefficient files to calculate 1-year mortality for patients 18-99 years old from demographic metrics and individual or categories of ICD-10-CM billing codes. Categories of codes were defined using the first revision of the 2021 Clinical Classifications Software Refined (CCSR) key (<https://www.hcup-us.ahrq.gov/toolssoftware/ccsr/prccsr.jsp>.) Coefficient files for one year mortality models were developed on 2018 CMS Part-A and Part-B population. Models were trained using logistic regression with both logistic and log-log link. Models were split by sex, as indicated in the coefficient files. The log-log model is the public version of the 1-year model referenced in **Covid-19 and Excess Morality in Medicare Beneficiaries** (doi: <https://doi.org/10.1101/2021.04.07.21254793>).

## Files

*morx\_1yr\_logistic\_model.csv* - coefficient file for logistic model.

*morx\_1yr\_log\_log\_model.csv* - coefficient file for log-log model.

## Field Description

**PARAMETER** - Name of coefficient. See section below for interpretation.

**CODE\_TYPE** - Field to indicate whether parameter is ICD-10-CM or CCSR code. Blank otherwise.

**ESTIMATE** - Trained model coefficient.

**STDERR** - Standard error of trained model coefficient.

**TVALUE** - T-value of trained model coefficient.

**YEAR** - Reference year for model training.

**SEX\_IDENT\_CD** - Medicare sex identification code. 1 = Male. 2 = Female.

## How-to Apply

### Encoding Billing Codes

All ICD10-CM and CCSR codes are set to 1 if present in the medical record during the past 365 days and to 0 otherwise. For CCSR mappings, use the first revision of the 2021 CCSR key. Also, note that an ICD10-CM code with a "\*" as the last character indicates that all ICD-10-CM codes in that family should be considered when encoding that variable.

### Encoding Age + Sex

Sex should be used to select the set of coefficients to use but does not have a coefficient in the models. Select coefficients where *SEX\_IDENT\_CD* = 1 for men, and where *SEX\_IDENT\_CD* = 2 for women. Age is categorically encoded. To encode age from the **PARAMETER** field, set field to 1 if *<age\_lower> <= age <= <age\_upper>*, otherwise set to 0. All age parameters are in the format, *AGECLASS <age\_lower><age\_upper>*. Predictions should be set to system missing (i.e., not calculated) for ages outside the collective range of the AGECLASS parameter (i.e., outside of 18-99).

## Encoding Special Variables

The parameter *LIS\_DIS 1* indicates whether or not the patient has low-income or disabled status (i.e., CMS dual eligible status). Set to 1 for low-income or disabled status and to 0 otherwise. The parameter *SNF\_LTC\_EH* indicates whether or not the patient is currently (past 30 days) receiving services in long-term-care or skilled nursing settings. Set to 1 for patients receiving services in long-term-care or skilled nursing settings within the past 30 days, and to 0 otherwise.

## Link Functions

Let  $X\beta$  represent the sum of all of the coefficients and the intercept; then the predicted 1-year mortality for the model using the log-log link function is  $\exp(-\exp(-X\beta))$  and for the model using the logit link function is  $1/(1+\exp(-X\beta))$ .