**Predicting Need for Advanced Illness or Palliative Care using Electronic Health Record Data**

**Kenneth Jung, Sylvia Sudat, Nicole Kwon, Walter F. Stewart\*, Nigam H. Shah**

1 Stanford University, Palo Alto, CA;

2 Sutter Health Research, Walnut Creek, CA, USA;

\* Corresponding Author: Walter F. Stewart, StewarWF@sutterhealth.org

**ABSTRACT**

Timely outreach to individuals in an advanced stage of illness offers opportunities to exercise decision control over health care. We examined whether electronic health record (EHR) data could be used to anticipate need with enough time to engage a patient. Using EHR data, we assembled a cohort of 349,667 primary care patients between 65 and 90 years of age who sought care from Sutter Health between July 1, 2011 and June 30, 2014, of whom 2.1% died during the study period. EHR data for each patient from the 12 month observation window prior to their prediction time were extracted, comprising demographics, encounters, orders, and diagnoses. L1 regularized logistic regression and gradient boosted tree models were fit to training data and tuned by cross validation. Model performance in predicting death was assessed by the areas under the receiver operator curve (AUROC) and the precision-recall curve (AUPRC) on held-out test patients. We evaluated the effect of data density, defined by the number of encounters (i.e., varied from 1, 2, 4, or 8) in the observation window, on model performance. Patients could be identified up to one year before death with a mean AUROC of 84.8% and 80.7% (with mean AUPRCs of 9.7% and 6.4%) for the gradient boosted tree models and logistic regression, respectively. Compared to all available patients with 1+ encounter, the data density requirement reduced the number of patients for whom a model could be built to 85.9% of the total for 2 encounters, 60.3% for 4 encounters, and 31.7% for 8 encounters.. Increased data density did not change discrimination significantly (mean AUROC varied from 82.5% to 83.0% for a requirement of 1 to 8 encounters). On subgroup analyses, mean AUROC was higher (87.9%) for patients who died 0 to 3 months after their prediction time, than for those who died from 9 to 12 months after prediction time (83.6%). We conclude that EHR data can be used to build a predictive model that accurately identifies individuals with up to 12 months lead time, to allow timely outreach for advanced illness management services.

**KEYWORDS:** Predictive model; machine learning; electronic health record; palliative care

**1. Introduction**

The majority of patients who face advanced illness are known to benefit from outreach that offers decision-control over care. Timely outreach makes it possible for patients to make choices aligned with personal preferences, potentially avoiding the loss of control that often occurs when urgent clinical interventions are required. Evidence indicates that when individuals are in an advanced stage of illness and also have control of care decisions, they often choose home-based, comfort-oriented care [[1–4]](https://paperpile.com/c/QM6OBK/705n+SkxI+DDi5+KHRr). The growth in availability and use of both hospice and inpatient palliative care programs is therefore consistent with patient preferences. Between 2000 and 2013, the percentage of hospitals offering palliative care increased from 25% to 72%, and the use of hospice services among Medicare beneficiaries increased from 22% in 2000 to 42.2% in 2009 [[5]](https://paperpile.com/c/QM6OBK/MKDw). However, outreach offering such support services is largely confined to inpatient care, usually following a serious acute event. As a consequence, care at the end of life continues to be aggressive because the option to choose is unnecessarily delayed or simply not offered.

The limited outreach to individuals with advanced illness has motivated new types of services to identify those in need and coordinate service offerings outside of the hospital setting [[6]](https://paperpile.com/c/QM6OBK/PbtH). Evidence suggests a positive impact of these programs on patient satisfaction and on more effective use of care [[1,7–10]](https://paperpile.com/c/QM6OBK/705n+w0Rm+yZ7K+Z3kZ+ctaA). However, these programs typically rely on ad hoc methods or passive reporting by physicians to identify patients in need. Most eligible patients are missed or are contacted too late to be of help in supporting patient preferences. Typically, three to four months is required after an initial contact for individuals to decide on options they want to exercise [[11]](https://paperpile.com/c/QM6OBK/3FbW). Therefore, effective outreach should occur roughly six months or more before end of life care decisions must be made.

Electronic health records (EHRs), now widely adopted in U.S. healthcare, have opened a unique era in medicine where population-level data on patients can be used in real time to predict and potentially improve outcomes for a given patient [[12–14]](https://paperpile.com/c/QM6OBK/2zDv+PosZ+w594). The purpose of this study is to determine the extent to which the need for end-of-life care can be predicted with sufficient lead time to allow for effective and timely outreach. In particular, we explore how technical choices made during model development, such as data density requirements, impact model performance and utility.

**2. Methods**

**2.1. Problem Formulation**

We developed predictive models with the intention of identifying patients sufficiently far in advance of death so that patients could decide what type of care they would prefer. We approached this as a supervised learning problem using EHR data. While we define the time period when patients *may benefit from end-of-life care as occurring six or more months before death, we nonetheless examined model performance for time periods ranging from 0-3 months to 10-12 months.*  The longer the prediction window, within limits, presumably offers the benefit to patients and families in having time to decide on their care options. In this analysis, we labeled those who died as positive cases, while other patients are labeled as negative cases.

Because this is a retrospective study, we must also decide for each patient *when in their timeline* we are making the prediction; we refer to this as the *prediction time*. For each patient, we pick a random date in the second year of the study period as their prediction time. We do so because in real life we never know exactly when someone is going to die; the final model will be applied to the patient’s record at some time, and some will die in the following 12 months, while others will not. Our models use as input EHR data from the twelve months prior to each patient’s chosen prediction time; we refer to this time period as the *observation window*. Thus, for a given patient, our task is to predict, at the patient’s *prediction time*, whether the patient will die within 12 months given the EHR data available in their *observation window* (Figure 1). Patients who died prior to their prediction time were removed from analysis..

**2.2. Source Population**

The study, completed as a retrospective cohort analysis of Sutter primary care patients, was approved by the Sutter Health Institutional Review Board. Sutter Health ([www.sutterhealth.org](http://www.sutterhealth.org)) is a not-for-profit open health system with a network of more than 5,000 physicians, 24 hospitals, and other healthcare services serving 23 counties in northern California, and uses a single instance of EpicCare across all of its health care delivery facilities. We used patient data from July 1, 2011 through June 30, 2014. Patients were included in the study if they: (1) had a primary care (PC) relationship with Sutter Health, defined as having at least two encounters with a primary care physician (Family Medicine, Internal Medicine or Obstetrics-Gynecology) during the study period; and (2) were 65 to 89 years of age throughout the study period. These patients were randomly partitioned once prior to analysis into training and test sets, with 70% of patients assigned to the training set.

**2.3. Data and Features**

EHR data from each patient’s observation window was processed into features as follows. The data comprised demographics (age, gender and race) and counts of individual ICD-9 diagnosis codes, procedure CPT codes, medication pharmacy subclasses, encounters, and hospital visits. Features were not created for ICD-9 and CPT codes occurring in fewer than 200 patients, or for pharmacy subclasses occurring in fewer than 10 unique patients. Intuitively, patients with complex medical histories including many distinct diagnoses, procedures and medication orders are more likely to be seriously ill. We therefore characterize the complexity of each patient’s medical history by computing - separately for diagnosis, procedure, and pharmacy subclasses - the maximum and minimum count of distinct codes over all days with a count of at least one. The number of distinct medical specialties seen during the observation window was also computed, along with the mean and maximum number of office visits and of distinct medical specialties seen on each day (again excluding days with zero counts). Supplementary Materials Table 1 lists the full set of features used.

**2.4. Experimental Design**

Many choices made during model development interact with each other to influence model performance and utility [[15,16]](https://paperpile.com/c/QM6OBK/5M6j+Dqw4). We thus performed experiments systematically varying important factors. The most important factor is the data density requirement for patient inclusion, where data density is defined as the number of clinical encounters during the patient’s observation window. Clinical encounters are defined as any patient interaction with a provider and may include ambulatory care office visits, virtual visits, ED visits, hospitalizations, and prescription orders. Patients with higher data density have more information on which to base a prediction, but requiring many encounters as an inclusion criteria reduces the eligible population for the study and directly influences the amount of data available for training and separately the patient population on which the model may eventually be used [[17]](https://paperpile.com/c/QM6OBK/p7Pj). Such patients are also more likely to be seriously ill than the general population [[18,19]](https://paperpile.com/c/QM6OBK/XrvX+o7IR). We varied the minimum number of encounters required in the observation window between 1, 2, 4, and 8 encounters. Motivated by recent work [[20]](https://paperpile.com/c/QM6OBK/BYmS) we also explored the effect of reducing dimensionality of the data by grouping ICD-9 codes into coarser categories as defined by the Healthcare Cost and Utilization project (HCUP) single level Clinical Classifications Software, or CCS, categories (see https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp).

For each combination of the above choices, we fit two types of models – L1 regularized logistic regression [[21]](https://paperpile.com/c/QM6OBK/GNS9) and gradient boosted trees [[22]](https://paperpile.com/c/QM6OBK/Z2FE) – yielding a total of 16 experimental conditions under which models were built.  Regularized logistic regression models are widely used because they are easy to interpret, straightforward to tune, and often yield performance that is close to that of more complex models; gradient boosted trees are able to model non-linearities and interactions between features. We used the glmnet [[23]](https://paperpile.com/c/QM6OBK/coxc) and gbm [[24]](https://paperpile.com/c/QM6OBK/jr7z) R packages, respectively.

**2.5. Model fitting**

We subsampled the training datasets for each experimental condition such that the prevalence of the positive class was 10% in order to avoid convergence problems when fitting linear models (without subsampling, many experimental conditions yielded linear models that contained only an intercept term due to the very low prevalence of positive cases). Hyper-parameters for each model were tuned as follows.  For logistic regression, the regularization hyper-parameter was tuned by 10-fold cross validation on the training data using the 1-s.e.m. rule [[25]](https://paperpile.com/c/QM6OBK/s70T). For gradient boosting, we fixed the maximum depth of the base models at 6 and learning rate to 0.005, and tuned the number of trees by performance on 30% of the training data reserved for this purpose. The final models were then fit on the training data with the optimal number of trees.

**2.6. Evaluation**

Models were evaluated on the held-out validation set using the area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC). Each model was evaluated for performance on four subpopulations of the validation set patients, defined by the lead time provided on each subpopulation. The subpopulations were those who died within 3 months, between 3 and 6 months, between 6 to 9 months, and between 9 and 12 months after prediction time. Note that we do not subsample the validation set to increase the prevalence of positive cases i.e. that validation set has the prevalence of death at 2.1%.

**3. Results**

We evaluated the performance of models predicting all causes mortality using the AUROC and AUPRC estimated on held out data. The results are summarized in Tables 1 and 2 respectively. Below we present detailed results as we vary experimental conditions.

**3.1. Study and Analytic Populations**

A total of 349,667 patients met inclusion criteria for the study; 2.1% died during the study period. Figure 2 summarizes the impact of different data density requirements on the size of the applicable population as a fraction of the overall study population, and on prevalence. As expected, more stringent data density requirements reduce the size of the population to which the model can be applied. In addition, because patients who are seriously ill tend to have more encounters with the healthcare system than healthy patients, prevalence also increases with data density requirements. Supplementary Materials Table 2 shows the characteristics of each of the datasets across the 16 experimental conditions.

**3.2. Model Class**

We fit gradient boosted trees and L1 regularized logistic regression models for each experimental condition. Under all 16 experimental conditions, gradient boosted trees were superior to logistic regression, suggesting that modeling interactions and non-linearities is beneficial. As measured by AUROC, boosted trees had an average advantage of 4.1% over logistic regression, with a minimum of 2% and a maximum of 7.5%. Across all experimental conditions, the mean AUROC of boosted trees was 84.8% and the mean AUROC of logistic regression was 0.807. As measured by AUPRC, which takes prevalence into account, the gap between boosted trees and logistic regression was smaller but still significant and consistent, with boosted trees outperforming logistic regression by 3.4% on average. The smallest gap in performance by AUPRC was less than 1.4%, while the largest gap in performance was 7%, and the mean AUPRC across all experimental conditions was 9.7% for boosted trees versus 6.4% for logistic regression.

**3.3. Diagnosis Code Grouping**

The features for our models are high-dimensional and sparse. It has been observed previously that grouping ICD-9 diagnosis codes improved the performance of models predicting CHF [cite Ng et al]. However, in our study, code grouping had little impact on model performance. Models using CCS diagnosis code groupings versus individual ICD-9 codes had average AUROCs of 82.6% versus 82.9%, and AUPRCs of 8.1% versus 8.2%, respectively. Thus, it appears that for this application there is little benefit to using grouped codes.

**3.4. Data Density Requirements**

Intuitively, one might expect that higher data density in the observation window (i.e., more encounters or more utilization of health care resources) would improve predictive accuracy due to increased information on which to base predictions. Furthermore, because more seriously ill patients are likely to have more encounters than healthy patients, we might expect that higher data density requirements would increase prevalence (Figure 2).

As we varied the minimum number of encounters required for patient inclusion in the study cohort from 1 to 8, we found that there was little variation in mean model performance as measured by AUROC, which varied in a relatively narrow range from 82.%5 to 83.0%. This suggests that it is not significantly easier to discriminate between positive and negative cases when higher data density requirements are enforced. However, requiring increased data density did have a significant impact on the patient population, with a 73% reduction in the size of that population as the minimum number of encounters increased from 1 to 8. In addition, the prevalence increased from 1.7% for a requirement of 1 encounter to 2.8% for a requirement of 8 encounters. The mean AUPRC varied from 7.1% under the 1 encounter requirement to 9.8% under the 8 encounter requirement.

**3.5. Lead Time**

Lead time, or how far in advance we are attempting to predict an outcome, is a critical aspect of model performance. Longer lead times generally make the prediction task more difficult, but also offer more time for effective interventions. We evaluated model performance for patients who died during various intervals after their prediction times - 0-3 months, 3-6 months, 6-9 months, and 9-12 months. We would expect it to be more difficult to discriminate patients who die farther from their prediction times. For patients who died within 3 months of their prediction times, the mean AUROC for gradient boosted trees using ICD9 codes was 87.9%, falling slightly to 83.6% for patients who died between 9 and 12 months after their prediction times. The AUPRC varied more dramatically, falling from a high of 13% for the 0-3 month group to 8.3% for the 9-12 month group even though the prevalences were roughly equal.

**4. Discussion**

The purpose of this study was to determine whether the need for advanced illness or end-of-life care can be predicted - with sufficient accuracy and with enough lead time - to allow for effective and timely outreach using data routinely collected in EHRs. We formulated the prediction problem using the surrogate outcome of all-cause mortality. Models were evaluated using AUROC, which measures the ability to discriminate between positive and negative cases regardless of prevalence, and AUPRC, which is sensitive to prevalence and may thus be more relevant when there is significant class imbalance [[26,27]](https://paperpile.com/c/QM6OBK/atuF+UARM). Each model was evaluated on the basis of its ability to identify patients who died in three month intervals after their prediction times. We systematically varied model class, diagnosis code groupings, and data density requirements to examine the impact on each of these factors on model performance and utility.

We found that the most important of these factors was data density, i.e., the minimum number of encounters required during a patient’s observation period to be included in the study. Increasing the data density requirement from 1 to 8 encounters significantly restricted the population to which the model could be applied and increased the prevalence of positive cases. We found that the AUROC varied only slightly over the range of data densities examined, while the AUPRC varied more significantly, suggesting that any gains in performance were due primarily to the increased prevalence of positive cases among patients with many encounters. We further discovered, consistent with intuition, that it is harder to discriminate patients who died farther in the future, but reasonable discrimination was possible even 9 months in advance of death. Gradient boosted trees consistently outperformed regularized logistic regression, suggesting that non-linearities and interactions are important for this application. Finally, dimensionality reduction by aggregation of ICD-9 diagnosis codes into CCS categories did not prove to be very beneficial in this setting.

These results suggest that predictive models using EHR data can be applied to a broad patient population, and can effectively identify the need for advanced illness or end of life care far enough in advance for effective interventions.

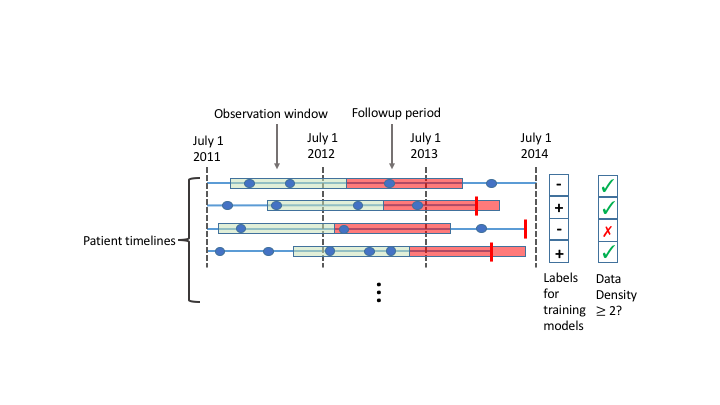
Surveillance is typically considered when an important health need is identified for which there is a solution available.Evidence indicates that when individuals are in an advanced stage of illness and also have control of care decisions, they often choose home-based, comfort-oriented care. However, conducting surveillance for patients who are approaching the end of life is difficult as there is no single set of ICD-9 or other digital codes for identifying such patients. Hence, accurate predictive models such as those we have developed create opportunities to proactively reach individuals in need.

Our work has important limitations with respect to addressing these needs. There are a host of cultural issues with predicting end of life including, but not limited to, ethical challenges as to how such models might be used, communication challenges around how providers, patients, and families are contacted and engaged to discuss options, and management challenges to ensure that patient autonomy is not usurped and that fairness and patient control of care access is maintained. Furthermore, there are challenges in operationalizing such a predictive model, especially given that a physician must ultimately decide whether and what actions are warranted, recognizing that models are imperfect. Decisions must therefore be balanced with caution in managing false positives, the benefits of patient control in care decision-making, and the risks from future care that offers little benefit.

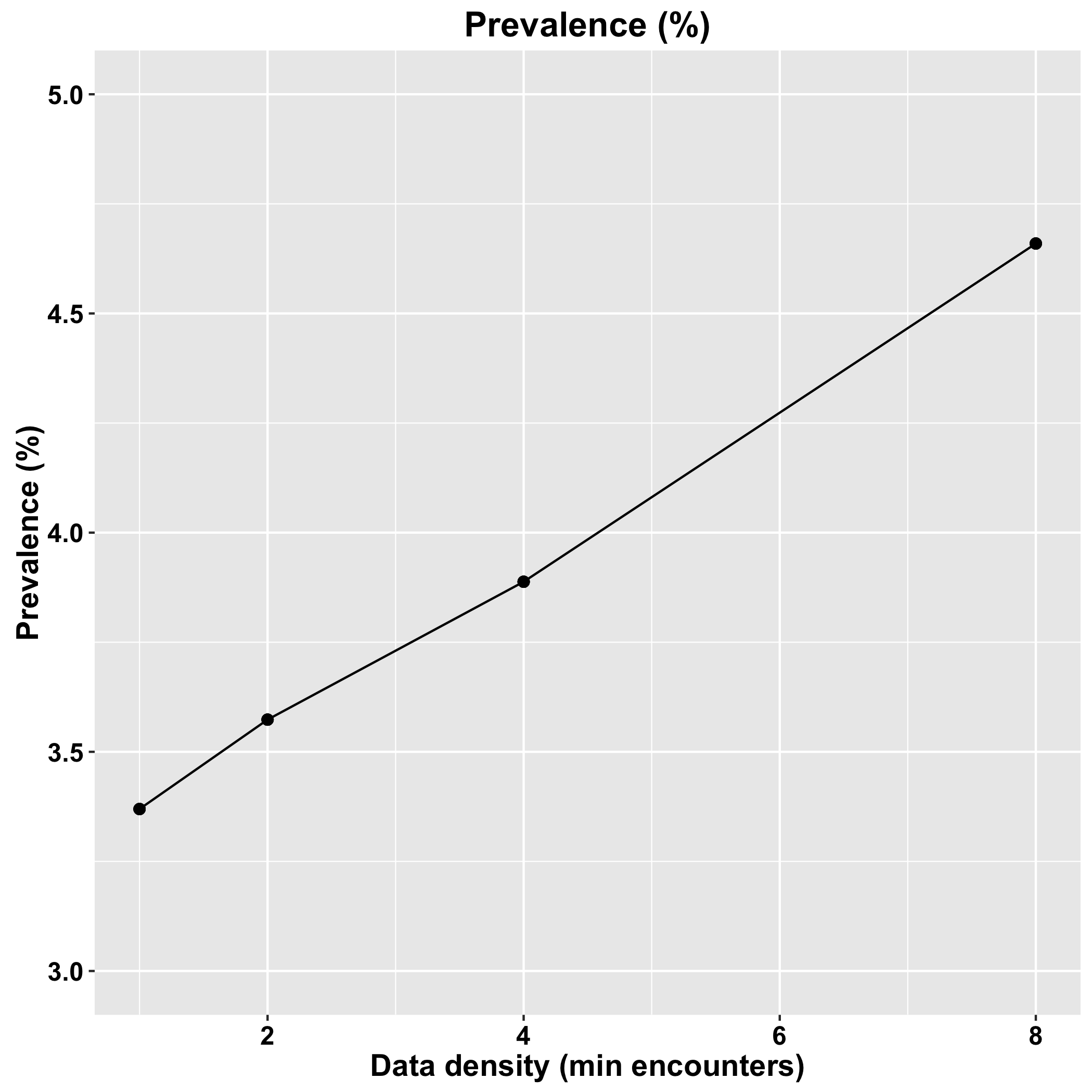
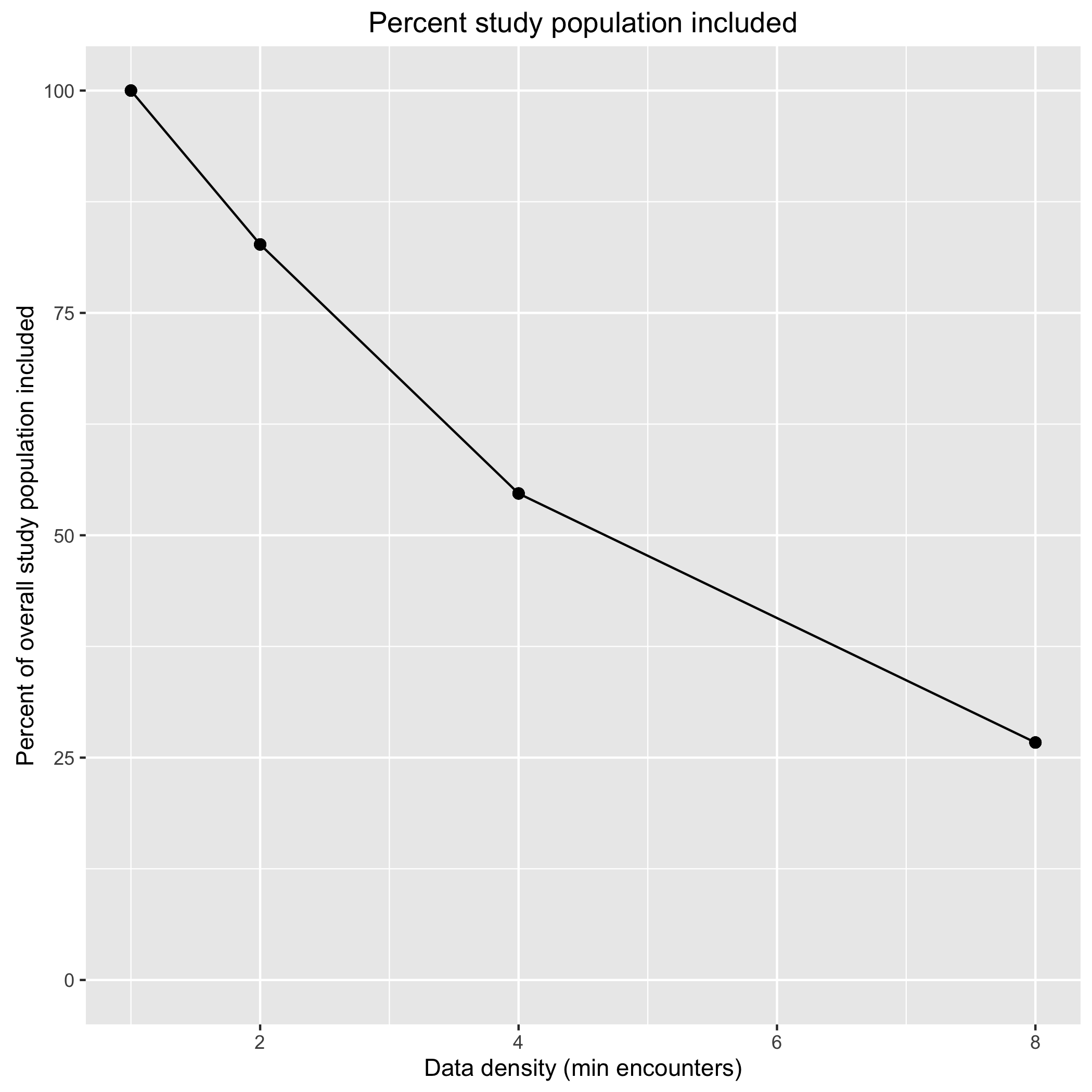
In summary, we have demonstrated that it is technically possible to predict the need for end of life care, and have laid the groundwork for future work to determine whether or not predictive models can be effectively used to actually improve patient outcomes. Our future work will focus on designing, and evaluating the utility of offering such care. Our development of predictive models for surveillance of advanced illness care need is motivated by the goal of improving patient access to decision control and the options to make choices aligned with personal preferences. Implementation of predictive models for patients at risk could potentially reduce the loss of control that often occurs when urgent clinical interventions are required.

**5. Conclusion**

Constructing and validating a predictive model for end-of-life care needs is the first step in identifying which patients should be a priority for timely advanced illness or end-of-life care outreach. We have shown that EHR data can be used to build a predictive model that accurately identifies individuals with up to 12 months of lead time.Given the high accuracy and long lead time offered by our models, we believe it is possible to devise effective clinical workflows that offer timely outreach for effective advanced illness management.



**Figure 1.** We construct a supervised learning problem for the task of predicting all causes mortality within a year. The dataset comprises patient level EHR data from July 1, 2011 through July 1, 2014. We represent each patient’s timeline as a horizontal line, with events shown as blue dots (for clinical encounters) or vertical red bars (for mortality). We construct a supervised learning problem as follows. For each patient we pick a random time in the second year of the study period. We then use the data from the year prior to these times (green segments) to predict the probability of mortality in the follow-up period (red segments). Positive cases are patients whose timelines end during their follow-up period (2nd and 4th rows). Data density requirements may exclude patients from the study. For instance, using a data density requirement of 2 encounters, the patient represented by the 3rd row would be excluded because they have only 1 encounter during in the year preceding their prediction time.



**Figure 2.** Impact of data density on applicable population and prevalence. As data density requirements increase, the fraction of the population included (and for whom the model would be applicable later) falls. Because patients with higher data density tend to be sicker, the prevalence of the outcome (death) in tandem with as data density. This increase directly impacts performance measures such as AUPRC.

**Table 1.** Model performance measured by AUROC across experimental conditions. Performance is shown separately for positive cases who died in different time intervals relative to their prediction times. The time intervals are denoted in months. We systematically varied minimum data density requirements and diagnosis coding.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **GBM** | | | | **LASSO** | | | |
| **Data Density** | **Dx coding** | **0-3 mo** | **3-6 mo** | **6-9 mo** | **9-12 mo** | **0-3 mo** | **3-6 mo** | **6-9 mo** | **9-12 mo** |
| **1** | **ICD-9** | 0.876 | 0.856 | 0.822 | 0.832 | 0.842 | 0.82 | 0.791 | 0.809 |
| **2** | **ICD-9** | 0.878 | 0.863 | 0.822 | 0.832 | 0.828 | 0.812 | 0.783 | 0.8 |
| **4** | **ICD-9** | 0.877 | 0.874 | 0.818 | 0.839 | 0.822 | 0.831 | 0.772 | 0.801 |
| **8** | **ICD-9** | 0.885 | 0.869 | 0.836 | 0.841 | 0.831 | 0.811 | 0.761 | 0.797 |
|  | | | | | | | | | |
| **1** | **CCS** | 0.877 | 0.855 | 0.816 | 0.819 | 0.821 | 0.803 | 0.769 | 0.793 |
| **2** | **CCS** | 0.878 | 0.862 | 0.815 | 0.819 | 0.837 | 0.818 | 0.781 | 0.799 |
| **4** | **CCS** | 0.876 | 0.872 | 0.809 | 0.828 | 0.836 | 0.838 | 0.774 | 0.807 |
| **8** | **CCS** | 0.881 | 0.863 | 0.827 | 0.833 | 0.839 | 0.819 | 0.767 | 0.802 |

**Table 2.** Model performance measured by AUPRC across experimental conditions. Performance is shown separately for positive cases who died in different time intervals relative to their prediction times. The time intervals are denoted in months. We systematically varied minimum data density requirements and diagnosis coding.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **GBM** | | | | **LASSO** | | | |
| **Data Density** | **Dx coding** | **0-3 mo** | **3-6 mo** | **6-9 mo** | **9-12 mo** | **0-3 mo** | **3-6 mo** | **6-9 mo** | **9-12 mo** |
| **1** | **ICD-9** | 0.111 | 0.104 | 0.063 | 0.072 | 0.11 | 0.095 | 0.062 | 0.069 |
| **2** | **ICD-9** | 0.114 | 0.111 | 0.064 | 0.074 | 0.111 | 0.1 | 0.064 | 0.069 |
| **4** | **ICD-9** | 0.134 | 0.13 | 0.069 | 0.079 | 0.132 | 0.117 | 0.067 | 0.076 |
| **8** | **ICD-9** | 0.161 | 0.153 | 0.079 | 0.105 | 0.148 | 0.135 | 0.074 | 0.104 |
|  | | | | | | | | | |
| **1** | **CCS** | 0.069 | 0.067 | 0.041 | 0.05 | 0.065 | 0.066 | 0.037 | 0.05 |
| **2** | **CCS** | 0.07 | 0.069 | 0.041 | 0.05 | 0.076 | 0.071 | 0.04 | 0.054 |
| **4** | **CCS** | 0.078 | 0.078 | 0.044 | 0.059 | 0.084 | 0.08 | 0.044 | 0.062 |
| **8** | **CCS** | 0.091 | 0.087 | 0.046 | 0.069 | 0.097 | 0.1 | 0.045 | 0.076 |

**FUNDING SOURCES**

Funding for N.S. and K.J. was provided by National Library of Medicine grant 2R01LM011369-0. S.S., N.K., and W.F.S. were supported by AIM HCIA Grant 1C1CMS331005-01-00 from the U.S. Department of Health and Human Services, Centers for Medicare and Medicaid Services.

**REFERENCES**

1. [Sudat SE, Franco A, Pressman AR, Rosenfeld K, Gornet E, Stewart W. Impact of home-based, patient-centered support for people with advanced illness in an open health system: A retrospective claims analysis of health expenditures, utilization, and quality of care at end of life. Palliat Med. 2017; 269216317711824. doi:](http://paperpile.com/b/QM6OBK/705n)[10.1177/0269216317711824](http://dx.doi.org/10.1177/0269216317711824)

2. [Teno JM, Gozalo PL, Bynum JPW, Leland NE, Miller SC, Morden NE, et al. Change in end-of-life care for Medicare beneficiaries: site of death, place of care, and health care transitions in 2000, 2005, and 2009. JAMA. 2013;309: 470–477. doi:](http://paperpile.com/b/QM6OBK/SkxI)[10.1001/jama.2012.207624](http://dx.doi.org/10.1001/jama.2012.207624)

3. [Dumanovsky T, Augustin R, Rogers - Journal of palliative … M, 2016. The growth of palliative care in US hospitals: a status report. online.liebertpub.com. 2016; Available:](http://paperpile.com/b/QM6OBK/DDi5) <http://online.liebertpub.com/doi/abs/10.1089/jpm.2015.0351>

4. [Vaughn J, Szekendi M. Gaps in the Use of Palliative Care in US Hospitals (FR461D). J Pain Symptom Manage. 2017;53: 379. doi:](http://paperpile.com/b/QM6OBK/KHRr)[10.1016/j.jpainsymman.2016.12.152](http://dx.doi.org/10.1016/j.jpainsymman.2016.12.152)

5. [Teno JM, Christian TJ, Gozalo P, Plotzke M. Proportion and Patterns of Hospice Discharges in Medicare Advantage Compared to Medicare Fee-for-Service. J Palliat Med. 2017; doi:](http://paperpile.com/b/QM6OBK/MKDw)[10.1089/jpm.2017.0046](http://dx.doi.org/10.1089/jpm.2017.0046)

6. [Sylvia Sudat, Sutter Health, Walnut Creek, CA, Kathy Blanton, Sutter Health, Walnut Creek, CA, Jessica Hanserd, Sutter Care at Home, Emeryville, CA, Kenneth Rosenfeld, California Pacific Medical Center, San Francisco, CA, Anjali Franco, Sutter Health, Walnut Creek, CA, Shruti Vaidya, Sutter Health, Walnut Creek, CA, et al. Development of an Algorithm to Prospectively Identify Palliative Care-Eligible Patients From the Electronic Health Record. Journal of Patient-Centered Research and Reviews. Aurora Health Care, Inc.; 2017;4: 182. doi:](http://paperpile.com/b/QM6OBK/PbtH)[10.17294/2330-0698.1537](http://dx.doi.org/10.17294/2330-0698.1537)

7. [Chen CY, Thorsteinsdottir B, Cha SS, Hanson GJ, Peterson SM, Rahman PA, et al. Health care outcomes and advance care planning in older adults who receive home-based palliative care: a pilot cohort study. J Palliat Med. 2015;18: 38–44. doi:](http://paperpile.com/b/QM6OBK/w0Rm)[10.1089/jpm.2014.0150](http://dx.doi.org/10.1089/jpm.2014.0150)

8. [Kerr CW, Tangeman JC, Rudra CB, Grant PC, Luczkiewicz DL, Mylotte KM, et al. Clinical impact of a home-based palliative care program: a hospice-private payer partnership. J Pain Symptom Manage. 2014;48: 883–92.e1. doi:](http://paperpile.com/b/QM6OBK/yZ7K)[10.1016/j.jpainsymman.2014.02.003](http://dx.doi.org/10.1016/j.jpainsymman.2014.02.003)

9. [Engelhardt JB, McClive-Reed KP, Toseland RW, Smith TL, Larson DG, Tobin DR. Effects of a program for coordinated care of advanced illness on patients, surrogates, and healthcare costs: a randomized trial. Am J Manag Care. 2006;12: 93–100. Available:](http://paperpile.com/b/QM6OBK/Z3kZ) <https://www.ncbi.nlm.nih.gov/pubmed/16464138>

10. [Seow H, Piet L, Kenworthy CM, Jones S, Fagan PJ, Dy SM. Evaluating a palliative care case management program for cancer patients: the Omega Life Program. J Palliat Med. 2008;11: 1314–1318. doi:](http://paperpile.com/b/QM6OBK/ctaA)[10.1089/jpm.2008.0140](http://dx.doi.org/10.1089/jpm.2008.0140)

11. [Scibetta C, Kerr K, Mcguire J, Rabow MW. The Costs of Waiting: Implications of the Timing of Palliative Care Consultation among a Cohort of Decedents at a Comprehensive Cancer Center. J Palliat Med. 2016;19: 69–75. doi:](http://paperpile.com/b/QM6OBK/3FbW)[10.1089/jpm.2015.0119](http://dx.doi.org/10.1089/jpm.2015.0119)

12. [Jung K, Covington S, Sen CK, Januszyk M, Kirsner RS, Gurtner GC, et al. Rapid identification of slow healing wounds. Wound Repair Regen. 2016;24: 181–188. doi:](http://paperpile.com/b/QM6OBK/2zDv)[10.1111/wrr.12384](http://dx.doi.org/10.1111/wrr.12384)

13. [Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015;7: 299ra122. doi:](http://paperpile.com/b/QM6OBK/PosZ)[10.1126/scitranslmed.aab3719](http://dx.doi.org/10.1126/scitranslmed.aab3719)

14. [Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc. 2017;24: 198–208. doi:](http://paperpile.com/b/QM6OBK/w594)[10.1093/jamia/ocw042](http://dx.doi.org/10.1093/jamia/ocw042)

15. [Botsis T, Hartvigsen G, Chen F, Weng C. Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. AMIA Jt Summits Transl Sci Proc. 2010;2010: 1–5. Available:](http://paperpile.com/b/QM6OBK/5M6j) <https://www.ncbi.nlm.nih.gov/pubmed/21347133>

16. [Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. J Biomed Inform. 2013;46: 830–836. doi:](http://paperpile.com/b/QM6OBK/Dqw4)[10.1016/j.jbi.2013.06.010](http://dx.doi.org/10.1016/j.jbi.2013.06.010)

17. [Weber GM, Adams WG, Bernstam EV, Bickel JP, Fox KP, Marsolo K, et al. Biases introduced by filtering electronic health records for patients with “complete data.” J Am Med Inform Assoc. 2017;24: 1134–1141. doi:](http://paperpile.com/b/QM6OBK/p7Pj)[10.1093/jamia/ocx071](http://dx.doi.org/10.1093/jamia/ocx071)

18. [Rusanov A, Weiskopf NG, Wang S, Weng C. Hidden in plain sight: bias towards sick patients when sampling patients with sufficient electronic health record data for research. BMC Med Inform Decis Mak. 2014;14: 51. doi:](http://paperpile.com/b/QM6OBK/XrvX)[10.1186/1472-6947-14-51](http://dx.doi.org/10.1186/1472-6947-14-51)

19. [Weiskopf NG, Rusanov A, Weng C. Sick patients have more data: the non-random completeness of electronic health records. AMIA Annu Symp Proc. 2013;2013: 1472–1477. Available:](http://paperpile.com/b/QM6OBK/o7IR) <https://www.ncbi.nlm.nih.gov/pubmed/24551421>

20. [Ng K, Steinhubl SR, deFilippi C, Dey S, Stewart WF. Early Detection of Heart Failure Using Electronic Health Records: Practical Implications for Time Before Diagnosis, Data Diversity, Data Quantity, and Data Density. Circ Cardiovasc Qual Outcomes. 2016;9: 649–658. doi:](http://paperpile.com/b/QM6OBK/BYmS)[10.1161/CIRCOUTCOMES.116.002797](http://dx.doi.org/10.1161/CIRCOUTCOMES.116.002797)

21. [Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33: 1–22. Available:](http://paperpile.com/b/QM6OBK/GNS9) <https://www.ncbi.nlm.nih.gov/pubmed/20808728>

22. [Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Stat. Institute of Mathematical Statistics; 2001;29: 1189–1232. Available:](http://paperpile.com/b/QM6OBK/Z2FE) <http://www.jstor.org/stable/2699986>

23. [http://cran.r-project.org/web/packages/glmnet/index.html.](http://paperpile.com/b/QM6OBK/coxc)

24. [gbm: Generalized Boosted Regression Models. R Package Ver. 2.1. http://cran.r-project.org/web/packages/gbm/. 2007;](http://paperpile.com/b/QM6OBK/jr7z)

25. [Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Springer-Verlag; 2009.](http://paperpile.com/b/QM6OBK/s70T)

26. [Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One. 2015;10: e0118432. doi:](http://paperpile.com/b/QM6OBK/atuF)[10.1371/journal.pone.0118432](http://dx.doi.org/10.1371/journal.pone.0118432)

27. [Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning. ACM; 2006. pp. 233–240. doi:](http://paperpile.com/b/QM6OBK/UARM)[10.1145/1143844.1143874](http://dx.doi.org/10.1145/1143844.1143874)