



Developing Predictive Models to Determine Patients in End-of-Life Care in Administrative Datasets

Joel N. Swerdel^{1,2,4} · Jenna M. Reps^{1,2} · Daniel Fife¹ · Patrick B. Ryan^{1,2,3}

Published online: 14 January 2020
© The Author(s) 2020

Abstract

Introduction In observational studies with mortality endpoints, one needs to consider how to account for subjects whose interventions appear to be part of ‘end-of-life’ care.

Objective The objective of this study was to develop a diagnostic predictive model to identify those in end-of-life care at the time of a drug exposure.

Methods We used data from four administrative claims datasets from 2000 to 2017. The index date was the date of the first prescription for the last new drug subjects received during their observation period. The outcome of end-of-life care was determined by the presence of one or more codes indicating terminal or hospice care. Models were developed using regularized logistic regression. Internal validation was through examination of the area under the receiver operating characteristic curve (AUC) and through model calibration in a 25% subset of the data held back from model training. External validation was through examination of the AUC after applying the model learned on one dataset to the three other datasets.

Results The models showed excellent performance characteristics. Internal validation resulted in AUCs ranging from 0.918 (95% confidence interval [CI] 0.905–0.930) to 0.983 (95% CI 0.978–0.987) for the four different datasets. Calibration results were also very good, with slopes near unity. External validation also produced very good to excellent performance metrics, with AUCs ranging from 0.840 (95% CI 0.834–0.846) to 0.956 (95% CI 0.952–0.960).

Conclusion These results show that developing diagnostic predictive models for determining subjects in end-of-life care at the time of a drug treatment is possible and may improve the validity of the risk profile for those treatments.

1 Introduction

Studies of medication adverse effects often exclude subjects who are in end-of-life care. This exclusion is reasonable because such subjects may use medications and

receive care that is quite different from others, and because their deaths are likely to reflect their underlying risk rather than an effect of the medication being studied. Exclusion of subjects who are receiving end-of-life care is particularly important for studies of medications that may be used as part of end-of-life care and for studies that include a mortality endpoint because failure to exclude such patients would introduce confounding by indication or reverse causality bias [1, 2]. This study was started because one of the authors identified evidence of reverse causation in a study of the association of several gastrointestinal medications with the risk of sudden cardiac death (SCD). The study was susceptible to reverse causation because the medications were used for patients in terminal care, who were not initially recognized as being in terminal care, typically are at high risk for nausea and vomiting, and (presumably because their deaths were expected) had ‘heart stopped’ listed as the cause of death. Having a simple means to identify patients who are likely to be in terminal care and do a sensitivity analysis to account for these subjects

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40264-020-00906-7>) contains supplementary material, which is available to authorized users.

✉ Joel N. Swerdel
jswerdel@its.jnj.com

¹ Janssen Research and Development, LLC, 920 Route 202, Raritan, NJ 08869, USA

² OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), 622 West 168th Street, PH-20, New York, NY 10032, USA

³ Columbia University, 622 West 168th Street, PH20, New York, NY 10032, USA

⁴ 1125 Trenton-Harbourton Road, Titusville, NJ 08560, USA

Key Points

Internal validation, where the model trained on 75% of the data was tested on the remaining 25% of the data, showed excellent performance characteristics, with a mean area under the receiver operator characteristic curve (AUC) of 0.950 across four administrative claims databases.

External validation, where the model trained on one dataset was tested on the other three datasets, showed very good to excellent performance characteristics, with AUCs ranging from 0.840 to 0.956.

Accounting for subjects who received an exposure to a drug or procedure at the point in their treatment when they were in end-of-life care may improve the validity of the risk profile for those treatments.

provides a straightforward means of testing for this issue [2, 3]. There are few documented methods for retrospective studies from administrative databases to identify and exclude subjects who were receiving end-of-life care.

Two recent studies brought this issue into sharper focus. One was a study of the risk of SCD among people exposed to domperidone, a gastrointestinal motility agent useful for the treatment of nausea and vomiting [3]. That study excluded subjects with cancer but did not initially exclude subjects who were receiving end-of-life care. Nausea and vomiting are common in end-of-life care. When the investigators developed an algorithm for identifying subjects who were receiving end-of-life care and excluded those subjects, the odds ratio estimate for the association of domperidone exposure and SCD fell from 2.1 to 1.7. The second study was a literature review to assess the association of the use of haloperidol in elderly patients with dementia with an increased risk of all-cause death [4]. It identified six studies that reported such an association but none excluded subjects who were receiving end-of-life care, although haloperidol is commonly used for its impact on delirium and control of nausea and vomiting.

Diagnostic predictive models have been developed across a number of different circumstances in recent years, particularly in research as a method to improve outcome ascertainment, and the use of these models has increased in recent years [5, 6]. Demirer et al. developed a model for tuberculous pleural effusions as a method to reduce the reliance on a diagnostic test (directly sampling plural fluid and testing for *Mycobacterium tuberculosis*) that is invasive and often unavailable in developing countries [7]. Predictive modeling was used to improve differential diagnosis of vascular parkinsonism versus Parkinson's disease

as symptoms often overlap between the two, presenting a clinical challenge [8]. Williams and colleagues used a predictive model to improve the pretest probability of the presence of occlusive coronary artery disease in order to select appropriate individuals for undergoing cardiac imaging [9].

The objective of this study was to develop a diagnostic predictive model to determine patients in end-of-life care at the time of their last prescription for a new medication in administrative databases.

2 Methods

2.1 Databases

Data for this study were from data collected between 1 January 2000 and 31 January 2017 from four data sets: IBM® MarketScan® Medicare Supplemental Beneficiaries (MDCR; data from 1 January 2000 to 31 January 2017); IBM® MarketScan® Multi-State Medicaid (MDCD; data from 1 January 2006 to 30 June 2016); IBM® MarketScan® Commercial Claims and Encounters (CCAE; data from 1 January 2000 to 31 January 2017); and Optum® De-Identified Clinformatics® Data Mart Database (Optum; data from 1 May 2000 to 30 September 2016). These datasets provided a broad range of subjects for the study, including those under age 65 years (CCAE and Optum), those over age 65 years (Optum and MDCR), and individuals of lower socioeconomic status (MDCD). The study dates were chosen based on the full set of data available for purchase from the data vendors at the time of analysis. Each dataset was converted to the Common Data Model (CDM), version 5.01, developed by the Observational Health Data Sciences and Informatics (OHDSI) interdisciplinary collaborative [10]. The Optum and IBM MarketScan databases used in this study were reviewed by the New England Institutional Review Board (IRB) and were determined to be exempt from broad IRB approval as this research project did not involve human subject research.

2.2 Prediction Model Specification

2.2.1 Target Population

Subjects were included in the study if they were 40 years of age or older and had a diagnosis of cancer, heart failure, respiratory failure, dementia, Parkinson's disease, liver failure, and/or renal failure on or prior to the index date. These conditions were chosen as a way to include those subjects more likely to be in end-of-life care and therefore of concern to a researcher conducting a study with a mortality endpoint. The administrative codes for the inclusion conditions

are shown in the Supplementary Appendix 1 spreadsheet (Table 1 labeled ‘Inclusion Concepts’). Each included subject was also required to have at least 1 year of continuous observation prior to the index date. For each subject, we identified the date of first exposure to each active ingredient, and then selected the last of those drug start dates as the subject’s index date. We chose the date of the last new drug start for the index date as this would provide the most likely drug prescribed for the treatment of end-of-life care.

2.2.2 Outcome

The outcome of end-of-life care was determined by a set of diagnostic conditions and clinical observations, procedures, and measures codes indicative of terminal and/or hospice care (see Supplementary Appendix 1 spreadsheet, Table 2 labeled ‘Outcome Concepts’). These included diagnostic conditions such as ‘death imminent’ and ‘terminal illness’, and clinical observations such as ‘hospice care supervision’ and ‘routine admission to hospice’. These codes were developed with the aid of two physicians (co-author DF and Soledad Cepeda, MD, PhD). It should be noted that a code for death was not required in order to improve the portability of the model as many health services databases do not reliably capture out-of-hospital death.

2.2.3 Time at Risk

Subjects were required to have at least one of the above-mentioned codes indicative of terminal and/or hospice care within 180 days prior to the index date to be considered to have the outcome and thus be considered a ‘case’ of end-of-life care at the time of the prescription. End-of-life care is generally considered to be for those with < 180 days to live, which was the rationale for deciding the period prior to the index date [11].

2.2.4 Covariates

Covariates used in the diagnostic predictive model included demographics (sex, age, index year and month), prior conditions, drugs, procedures and measurements observed during 30, 180, 365, or all days prior to the index date, Charlson comorbidity index [12], Diabetes Complications Severity Index (DCSI) score [13], CHADS2 score [14], and the number of distinct conditions, drugs, procedures and visits observed in the 365 days prior to the index date. Procedures included actions such as ‘intravenous injection’ and ‘continuous infusion of therapeutic substance’. Measurements included testing such as ‘albumin; urine, microalbumin, quantitative’ and ‘body weight’.

Covariates were included in the model if they were present in at least 25 subjects. We excluded all terminal,

hospice, and palliative care concepts from model development (see Supplementary Appendix 1 spreadsheet, Table 3 labeled ‘Model Exclusions’).

We used the OHDSI Patient-Level Prediction software, an open source R package, to develop the model. The software implements a framework for developing diagnostic models while addressing existing best practices towards ensuring that models are clinically useful and transparent [15]. We developed and tested models for each of the four databases. All prediction models were fitted using regularized logistic regression with an L_1 (LASSO) prior [15–17]. The optimal regularization hyperparameters were estimated using tenfold cross-validation. Each model was trained on a randomly selected sample of 75% of the subjects (the ‘training set’) and internal validation was performed on the remaining 25% (the ‘test set’). Subjects were excluded from both the test and training sets if they had the outcome prior to the index date. Internal validation was through examination of the area under the receiver operating characteristic (ROC) curve and through calibration of the model by examining the observed fraction of subjects with the outcome of end-of-life care throughout the range of predicted probabilities. External validation was through examination of the area under the ROC curve after applying the model learned on one dataset to the other three datasets.

3 Results

The number of participants, number of outcomes, and the demographic and comorbid characteristics for the four datasets analyzed in this study are displayed in Table 1. Continuous variables are presented as median (interquartile range [IQR]) and/or mean (standard deviation [SD]), while categorical variables are presented as number (%). The portion of the cohort with the outcome of being in end-of-life care was older than those subjects not in end-of-life care. In the 180 days prior to the index date, those with the outcome had higher rates of each of the seven comorbid conditions used for inclusion into the cohort. The largest differences between those with the outcome and those without were in the proportion of the cohort with a recorded condition of cancer in the prior 180 days. We did not test for statistical differences between proportions of those with or without the outcome.

Each of the models showed excellent performance based on our internal validation (Table 2). The AUCs after applying the model developed on the training set to the test set ranged from 0.918 in MDCR, the smallest dataset, to 0.983 in CCAE, the largest dataset. The intercepts for each of the fitted calibration curves were near 0 and the slopes of the curves were all 8% or less from unity, the ideal value indicating that the models were well calibrated [18]. The AUC and calibration curves for each of the models are shown in Fig. 1.

Table 1 Subject demographic and prior comorbid condition data for those in end-of-life care ('with outcome') and those not in end-of-life care ('no outcome') for the four datasets used in this study

Name	CCAЕ		MDCD		MDCR		Optum	
	With outcome	No outcome	With outcome	No outcome	With outcome	No outcome	With outcome	No outcome
Participants (N)	2199	278,590	1872	76,077	1237	193,786	1030	268,393
Age, years [mean (SD)]	56.6 (5.8)	55.6 (6.7)	70.6 (14.5)	66.2 (13.6)	82.0 (7.8)	77.2 (7.7)	69.1 (12.1)	64.4 (12.1)
Age, years [median (IQR)]	57.5 (53.0–61.1)	56.6 (50.7–61.2)	71.0 (57.8–85.0)	65.5 (55.2–77.7)	82.3 (76.3–87.9)	76.4 (70.9–82.8)	72.1 (59.0–79.5)	64.5 (54.9–74.3)
Female	1031 (46.9)	142,044 (51.0)	1205 (64.4)	51,186 (67.3)	700 (56.6)	103,729 (53.5)	583 (56.6)	138,003 (51.4)
Number of visits past 180 days [median (IQR)]	45.0 (28.0–67.0)	9.0 (5.0–17.0)	26.0 (12.0–52.0)	15.0 (7.0–40.0)	21.0 (12.0–35.0)	11.0 (6.0–20.0)	39.0 (21.0–63.0)	12.0 (6.0–23.0)
Charlson index [mean (SD)]	9.4 (3.5)	4.0 (2.7)	7.1 (4.1)	5.5 (3.3)	7.1 (3.6)	5.3 (3.2)	8.6 (3.8)	4.9 (3.2)
Cancer	1908 (86.8)	84,318 (30.3)	631 (33.7)	8380 (11.0)	426 (34.4)	50,481 (26.0)	572 (55.5)	72,706 (27.1)
Chronic renal failure	358 (16.3)	16,517 (5.9)	441 (23.6)	9048 (11.9)	245 (19.8)	19,331 (10.0)	229 (22.2)	24,190 (9.0)
Dementia	85 (3.9)	2198 (0.8)	549 (29.3)	6700 (8.8)	471 (38.1)	21,870 (11.3)	313 (30.4)	17,214 (6.4)
Heart failure	226 (10.3)	19,966 (7.2)	498 (26.6)	10,689 (14.1)	402 (32.5)	35,179 (18.2)	288 (28.0)	34,311 (12.8)
Liver failure	98 (4.5)	2034 (0.7)	80 (4.3)	732 (1.0)	7 (0.6)	563 (0.3)	23 (2.2)	1815 (0.7)
Parkinsons disease	17 (0.8)	1685 (0.6)	40 (2.1)	639 (0.8)	64 (5.2)	4037 (2.1)	35 (3.4)	3278 (1.2)
Respiratory failure	349 (15.9)	8164 (2.9)	301 (16.1)	4432 (5.8)	153 (12.4)	9867 (5.1)	133 (12.9)	12,239 (4.6)

Data are expressed as n (%) unless otherwise specified

CCAЕ IBM® MarketScan® Commercial Claims and Encounters, MDCD IBM® MarketScan® Multi-State Medicaid, MDCR IBM® MarketScan® Medicare Supplemental Beneficiaries, Optum Optum® De-Identified Clinformatics® Data Mart Database, SD standard deviation, IQR interquartile range

Table 2 Internal validation of models on the 25% subset of the dataset held back during model training for each dataset

Dataset	AUC (95% CI)	Calibration intercept	Calibration slope
CCAЕ	0.983 (0.978–0.987)	0.000	1.05
MDCD	0.947 (0.937–0.957)	–0.001	1.06
MDCR	0.918 (0.905–0.930)	0.000	1.08
Optum	0.945 (0.932–0.958)	0.000	1.07

AUC area under receiver operating characteristic curve, CI confidence interval, CCAЕ IBM® MarketScan® Commercial Claims and Encounters, MDCD IBM® MarketScan® Multi-State Medicaid, MDCR IBM® MarketScan® Medicare Supplemental Beneficiaries, Optum Optum® De-Identified Clinformatics® Data Mart Database

Our models also performed well on external validation (Table 3). After applying a model learned on one dataset to the other datasets, the AUCs ranged from 0.840 for the CCAЕ-learned model applied to MDCD, to 0.956 for the

MDCR-learned model applied to CCAЕ and for the MDCD-learned model applied to CCAЕ. The graphical representations of the AUCs for the external validations are shown in Fig. 2.

The prevalence of subjects in end-of-life care is quite low in our datasets, i.e. < 1% in CCAЕ, MDCR, and Optum, and 2.5% in MDCD. At low outcome prevalence, examination of the precision-recall curves is important. Figure 3 shows the precision-recall curves for each of the four datasets. As expected, precision was highest in the dataset with the highest prevalence, i.e. MDCD. The curve indicates that at a recall of 0.5, the precision was about 22%, which is close to 10 times the baseline rate. This indicates that, while precision is low, the model was able to effectively discriminate between those in end-of-life care from those not in end-of-life care.

The primary objective of these models was to improve the ability to predict which subjects are in end-of-life care. It is, however, possible to examine the covariates in the model

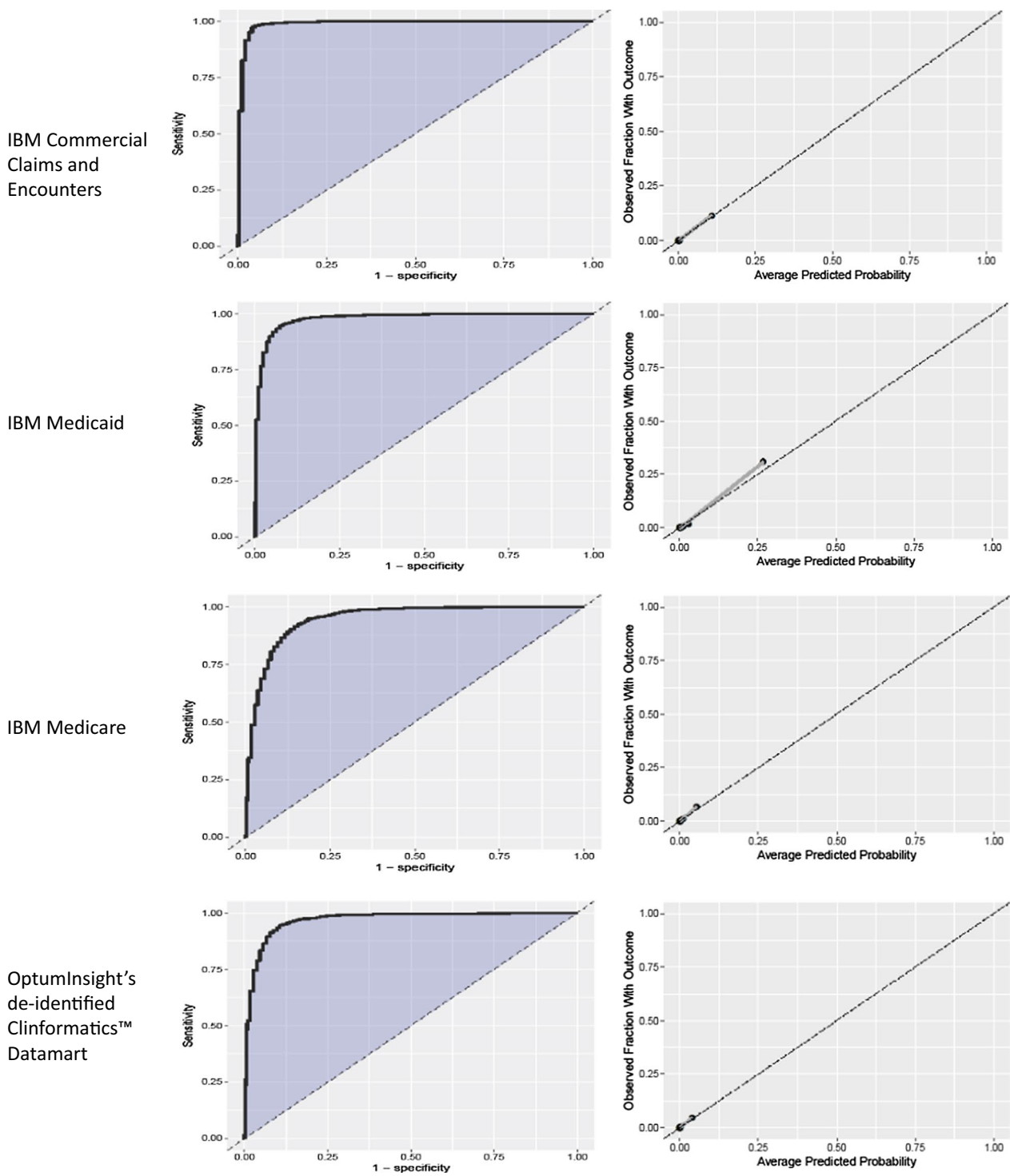


Fig. 1 Area under the receiver operator characteristic curve and calibration curves for the internal validation of the models from the four datasets used in this study

to understand important predictors and to understand the similarities and differences between the models. The number of observations in the prior 180 days, the presence of malignant neoplastic disease, and dementia were among

the common model-determined factors between the four datasets. Many differences existed between the models. For example, a condition of ‘failure to thrive’ was an important factor in the models for MDCD, MDCR, and Optum, which

Table 3 External validation of each model when applying to the other three datasets

Train dataset	Test dataset	AUC (95% CI)
CCAE	MDCD	0.840 (0.834–0.846)
	MDCR	0.883 (0.879–0.887)
	Optum	0.935 (0.931–0.939)
MDCD	CCAE	0.956 (0.952–0.960)
	MDCR	0.879 (0.875–0.883)
	Optum	0.909 (0.906–0.912)
MDCR	CCAE	0.956 (0.952–0.960)
	MDCD	0.888 (0.882–0.894)
	Optum	0.936 (0.932–0.940)
Optum	CCAE	0.963 (0.959–0.967)
	MDCD	0.868 (0.862–0.874)
	MDCR	0.893 (0.889–0.897)

AUC area under receiver operating characteristic curve, CI confidence interval, CCAE IBM® MarketScan® Commercial Claims and Encounters, MDCD IBM® MarketScan® Multi-State Medicaid, MDCR IBM® MarketScan® Medicare Supplemental Beneficiaries, Optum Optum® De-Identified Clinformatics® Data Mart Database

included many patients older than 65 years of age. It was not a factor in the CCAE model, which is composed of primarily younger individuals still in the work force. ‘Asthenia’ was also an important predictor in the models for MDCD, MDCR, and Optum, but not in CCAE. The number of clinical visits in the 180 days prior to the index date was a primary factor in the CCAE dataset, but not in MDCD, MDCR, or Optum. Similarly, ‘motor neuron disease’ was also an important predictor in CCAE, but not in the other datasets. The full prediction models for each of the datasets are shown in the Supplementary Appendix 2 spreadsheet.

4 Discussion

The results from this study show that developing diagnostic predictive models for determining subjects in end-of-life care at the time of a particular medication treatment is possible. This type of model may prove useful for improving estimates for mortality endpoints in pharmacoepidemiological studies. To our knowledge, this is the first study to develop a model for assessing subjects in end-of-life care using large-scale datasets.

We purposely used administrative codes for terminal or hospice care to develop our gold-standard cases for model development, choosing not to use palliative care codes. The World Health Organization definition of palliative care includes its applicability “early in the course of illness, in conjunction with other therapies that are intended to prolong life, such as chemotherapy or radiation therapy, and

includes those investigations needed to better understand and manage distressing clinical complications” [19]. In recent years, the use of palliative care has expanded to a wide range of illnesses that, while life-threatening, may not explicitly indicate end-of-life care [20]. In this study, we focused on terminal or hospice care which we believe will be more relevant to future data. Our case-defining algorithm also likely had a higher specificity for determining end-of-life care than one incorporating palliative care. The predictive model developed using this initial algorithm will also have higher specificity, which is critical for studies examining mortality endpoints.

There has been little prior research using diagnostic models for determining subjects in observational studies who were in end-of-life care at the time of the exposure of interest. Guadagnolo and colleagues examined radiation treatment in the last 30 days of life to ascertain utilization of this therapy during end-of-life care [21]. Those authors acknowledged that using the last 30 days of life as a criterion for end-of-life care at the time of treatment may not have accurately discriminated between the use of radiation for palliation versus curative purposes. For their study examining end-of-life cancer treatment, Huo et al. used hospice care codes to determine those in end-of-life care [22]. Those authors cited prior studies examining the many factors involved in a patient’s use of hospice care, including sex, marital status, hospice availability, and changes in attitudes toward death over time [23–25]. Following the publication of that study, several other studies have found that factors such as cancer type, physician preference, and geography also affect the decision to enter hospice care [26–28]. These limitations may have significantly reduced the sensitivity of this method for ascertainment of those in end-of-life care. The current study complements the works of Avati and colleagues and Jung and colleagues, who developed prognostic predictive models for proactively determining patients who should be referred for palliative care [29, 30]. Our use of a predictive model to improve the performance of a heuristic model increases our capability to determine those in end-of-life care.

The value of using a diagnostic predictive model for determining the probability of subjects in end-of-life care in addition to a heuristic model using administrative codes is for improving the overall sensitivity of finding those subjects in end-of-life care. Our modeling process found subjects who did not have codes in their patient record for end-of-life care but had high probabilities for being in end-of-life care. It is not known why these subjects did not have the codes for end-of-life care. Regardless of the rationale for not including the codes in the patient record, the model provides the capability to reduce systematic bias due to misclassification, in this case classifying those in end-of-life care as not being in end-of-life care. A link to the R package used to create

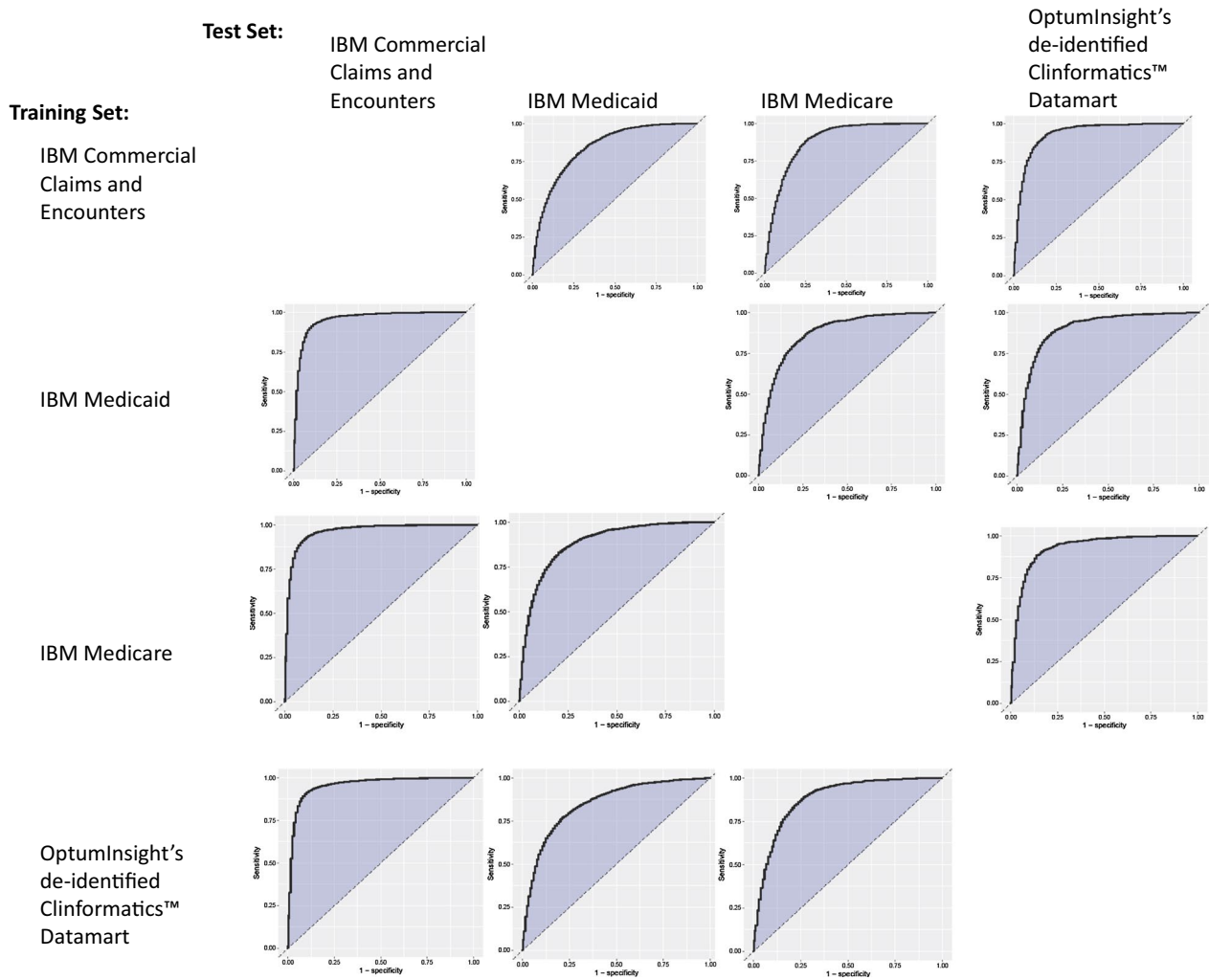


Fig. 2 Area under the receiver operator characteristic curves for the external validation of the models from the four datasets used in this study

these models is included in Supplementary Appendix 3. This package, along with the instructions, also provided in Supplementary Appendix 3, may be used by researchers in their own studies.

Our models performed well at discriminating end-of-life care from those not in end-of-life care, but were not perfect. We do not recommend removing patients predicted by the model as in end-of-life care. However, our models can be used as a sensitivity analysis for causal inference where we recommend applying the model to any patient with one of the conditions that make them high risk for end-of-life care to assign them a risk of being in end-of-life care. For patients who do not have a condition that makes them high risk, their risk will be 0. This risk value can be used as a covariate in a propensity model or in the outcome model. It would be interesting in future work to see the impact of including our model as a covariate into causal inference studies where end-of-life care is a likely confounder.

There are a number of strengths to the present study. We developed our models using four large datasets covering a wide variety of ages and socioeconomic conditions. In using multiple datasets, we were able to both internally and externally validate each model. Our method used large-scale regularized regression for model development, allowing for the use of factors known to be critical in determining end-of-life care as well as those that may not have been previously known. There were also a number of limitations to our study. The use of administrative datasets primarily maintained for insurance billing is well-known to have significant deficits. Our gold standard for case ascertainment was based on administrative codes for terminal or hospice care, which may be subject to misclassification errors. This model was also developed and validated on people who had at least one of several diagnoses, e.g. cancer, Parkinson's disease, and dementia, therefore it may perform differently on people who do not have any of those diagnoses or whose end of life

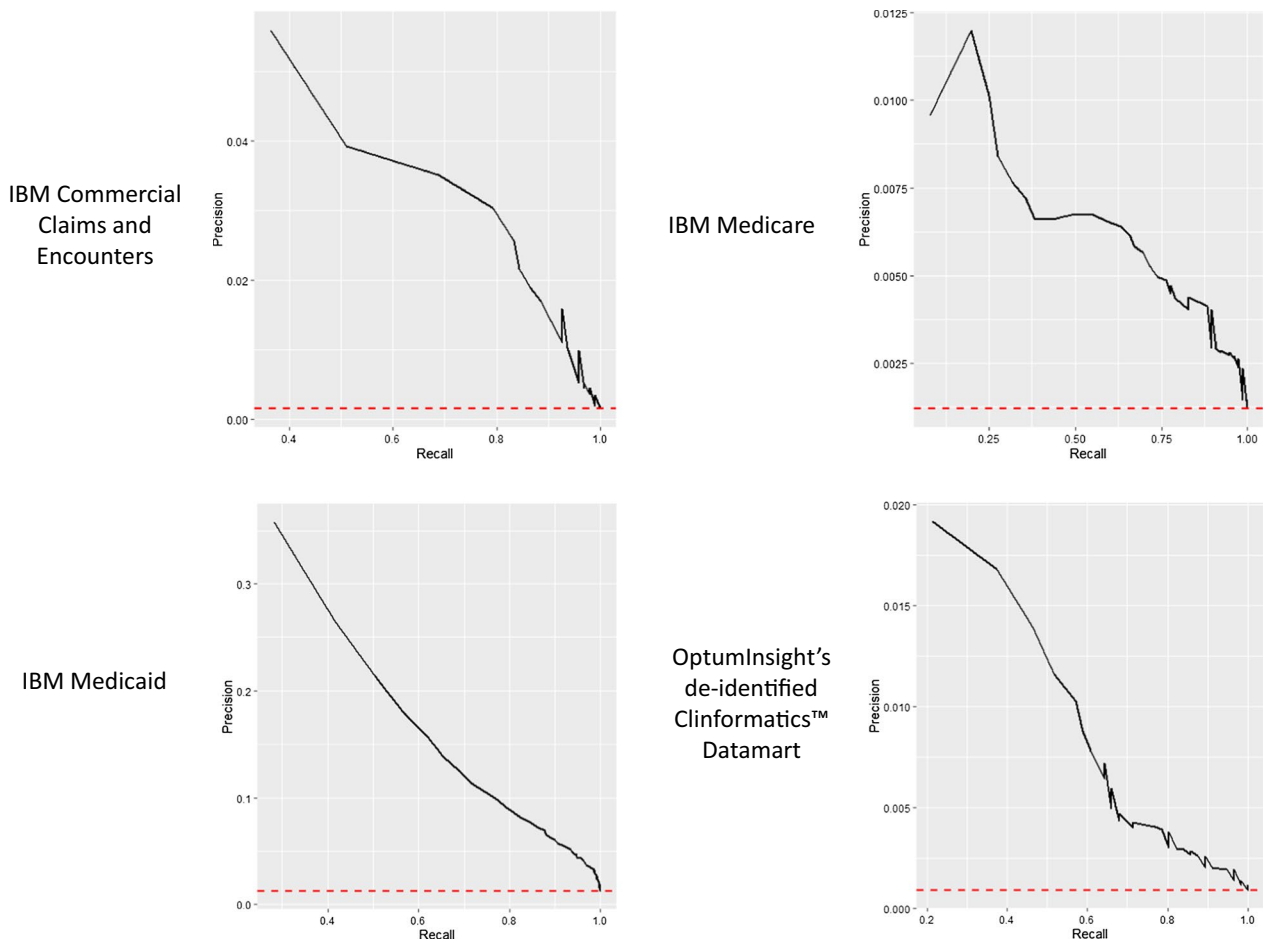


Fig. 3 Area under precision-recall curves for the internal validation of the models from the four datasets used in this study. The black line is the recall (sensitivity) plotted against precision (positive predictive

value) and the red dashed line is the fraction of the target population who have the outcome (population average outcome risk)

is not influenced by a chronic disease but by a more acute disease or event (e.g, trauma).

5 Conclusion

The models developed for determining those subjects in end-of-life care are an important step forward for improving the validity of exposure risks involving mortality endpoints. Possible next steps would include applying the model to epidemiological studies. Accounting for subjects who received an exposure to a drug or procedure at the point in their treatment when they were in end-of-life care will greatly improve the validity of the risk profile for those treatments.

Compliance with Ethical Standards

Conflict of interest Joel Swerdel, Jenna M. Reys, Daniel Fife, and Patrick Ryan were full-time employees of Janssen Research & Develop-

ment LLC at the time the study was conducted. They own stock, stock options, and pension rights from the company.

Funding This work was supported by Janssen Research & Development LLC, Raritan, NJ, USA.

Ethics approval The Optum and IBM MarketScan databases used in this study were reviewed by the New England IRB and were determined to be exempt from broad IRB approval as this research project did not involve human subject research.

Consent to participate Not applicable.

Data sharing The source data for this study were licensed by Johnson & Johnson from IBM MarketScan and Optum, and hence we are not allowed to share the licensed data publicly. However, the same data used in this study are available for purchase by contracting with the database owners, IBM MarketScan Research Databases (contact at: <http://www.ibm.com/us-en/marketplace/marketscan-research-databases>) and Optum (contact at: <http://www.optum.com/solutions/data-analytics/data/real-world-data-analytics-a-cpl/claims-data.html>). The authors did not have any special access privileges that other parties who license the data and contract with IBM MarketScan and Optum would not have.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

- Smith A, Murphy L, Bennett K, Barron TI. Patterns of statin initiation and continuation in patients with breast or colorectal cancer, towards end-of-life. *Support Care Cancer*. 2017;25(5):1629–37.
- Arana A, Johannes CB, McQuay LJ, Varas-Lorenzo C, Fife D, Rothman KJ. Risk of out-of-hospital sudden cardiac death in users of domperidone, proton pump inhibitors, or metoclopramide: a population-based nested case-control study. *Drug Saf*. 2015;38(12):1187–99.
- Varas-Lorenzo C, Arana A, Johannes CB, McQuay LJ, Rothman KJ, Fife D. Improving the identification of out-of-hospital sudden cardiac deaths in a general practice research database. *Drugs Real World Outcomes*. 2016;3(3):353–8.
- Hulshof TA, Zuidema SU, Ostelo RW, Luijendijk HJ. The mortality risk of conventional antipsychotics in elderly patients: a systematic review and meta-analysis of randomized placebo-controlled trials. *J Am Med Dir Assoc*. 2015;16(10):817–24.
- Agarwal V, Podchiyska T, Banda JM, Goel V, Leung TI, Minty EP, et al. Learning statistical models of phenotypes using noisy labeled training data. *J Am Med Inform Assoc*. 2016;23(6):1166–73.
- Halpern Y, Hornig S, Choi Y, Sontag D. Electronic medical record phenotyping using the anchor and learn framework. *J Am Med Inform Assoc*. 2016;23(4):731–40.
- Demirer E, Miller AC, Kunter E, Kartaloglu Z, Barnett SD, Elamin EM. Predictive models for tuberculous pleural effusions in a high tuberculosis prevalence region. *Lung*. 2012;190(2):239–48.
- Huertás-Fernández I, García-Gómez FJ, García-Solis D, Benítez-Rivero S, Marin-Oyaga VA, Jesús S, et al. Machine learning models for the differential diagnosis of vascular parkinsonism and Parkinson's disease using [(123)I]FP-CIT SPECT. *Eur J Nucl Med Mol Imaging*. 2015;42(1):112–9.
- Williams BA, Ladapo JA, Merhige ME. External validation of models for estimating pretest probability of coronary artery disease among individuals undergoing myocardial perfusion imaging. *Int J Cardiol*. 2015;182:534–40.
- Stang PE, Ryan PB, Racoosin JA, Overhage JM, Hartzema AG, Reich C, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med*. 2010;153(9):600–6.
- National Hospice and Palliative Care Organization. NHPCO's facts and figures hospice care in America. 2017. https://legacy.nhpco.org/sites/default/files/public/Statistics_Research/2017_Facts_Figures.pdf.
- 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(5):373–83.
- Young BA, Lin E, Von Korff M, Simon G, Ciechanowski P, Ludman EJ, et al. Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. *Am J Manag Care*. 2008;14(1):15–23.
- Welles CC, Whooley MA, Na B, Ganz P, Schiller NB, Turakhia MP. The CHADS2 score predicts ischemic stroke in the absence of atrial fibrillation among patients with coronary heart disease: data from the Heart and Soul Study. *Am Heart J*. 2011;162(3):555–61.
- Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *J Am Med Inform Assoc*. 2018;25(8):969–75.
- Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B (Methodol)*. 1996;58(1):267–88.
- Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D. Massive parallelization of serial inference algorithms for a complex generalized linear model. *ACM Trans Model Comput Simul*. 2013;23(1):1–23.
- Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. *Med Decis Mak*. 1993;13(1):49–58.
- World Health Organization. WHO Definition of Palliative Care. 2017. <http://www.who.int/cancer/palliative/definition/en/>.
- Hughes MT, Smith TJ. The growth of palliative care in the United States. *Annu Rev Public Health*. 2014;35:459–75.
- Guadagnolo BA, Liao KP, Elting L, Giordano S, Buchholz TA, Shih YC. Use of radiation therapy in the last 30 days of life among a large population-based cohort of elderly patients in the United States. *J Clin Oncol*. 2013;31(1):80–7.
- Huo J, Du XL, Lairson DR, Chan W, Jiang J, Buchholz TA, et al. Utilization of surgery, chemotherapy, radiation therapy, and hospice at the end of life for patients diagnosed with metastatic melanoma. *Am J Clin Oncol*. 2015;38(3):235–41.
- Miesfeldt S, Murray K, Lucas L, Chang CH, Goodman D, Morden NE. Association of age, gender, and race with intensity of end-of-life care for Medicare beneficiaries with cancer. *J Palliat Med*. 2012;15(5):548–54.
- Virnig BA, Kind S, McBean M, Fisher E. Geographic variation in hospice use prior to death. *J Am Geriatr Soc*. 2000;48(9):1117–25.
- Lackan NA, Ostir GV, Freeman JL, Mahnken JD, Goodwin JS. Decreasing variation in the use of hospice among older adults with breast, colorectal, lung, and prostate cancer. *Med Care*. 2004;42(2):116–22.
- Obermeyer Z, Powers BW, Makar M, Keating NL, Cutler DM. Physician characteristics strongly predict patient enrollment in hospice. *Health Aff (Millwood)*. 2015;34(6):993–1000.
- Lindley LC, Edwards SL. Geographic variation in California pediatric hospice care for children and adolescents: 2007–2010. *Am J Hosp Palliat Care*. 2018;35(1):15–20.
- Adersen M, Thygesen LC, Jensen AB, Neergaard MA, Sjogren P, Groenvold M. Is admittance to specialised palliative care among cancer patients related to sex, age and cancer diagnosis? A nationwide study from the Danish Palliative Care Database (DPD). *BMC Palliat Care*. 2017;16(1):21.
- Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. *BMC Med Inform Decis Mak*. 2018;18(Suppl 4):122.
- Jung K, Sudat SEK, Kwon N, Stewart WF, Shah NH. Predicting need for advanced illness or palliative care in a primary care population using electronic health record data. *J Biomed Inform*. 2019;92:103115.