

Clinical Utility of Automatable Prediction Models for Improving Palliative and End-Of-Life Care Outcomes: Towards Routine Decision Analysis Before Implementation

Supplementary materials

Literature review.....	2
Table S1. Validation studies of automatable or automated prediction models intended to prompt a palliative or end-of-life care intervention	4
Methods.....	7
Table S2. Admission diagnosis–mapped variables included in RF-AdminDemoDx model	12
Table S3. Comorbidity diagnosis–mapped variables included in the RF-AdminDemoDx model	17
Table S4. Characteristics of hospitalizations between July 2011 and July 2016 (Training cohort).....	19
Table S5. Characteristics of hospitalizations between July 2017 and July 2018 (Testing cohort)	21
Table S6. Clinical utility of prediction models with a random hospitalization per patient.....	23
Table S7. Clinical utility of prediction models with the first hospitalization per patient	24
Table S8. Clinical utility of prediction models with the last hospitalization per patient	25
Figure S1. Calibration plot of prediction models in testing cohort.....	26
Figure S2. Forest plot of clinical utility in relevant subgroups of hospitalized patients (RF-AdminDemoDx).....	27
Figure S3. Forest plot of clinical utility in relevant subgroups of hospitalized patients (mHOMR)	28
Figure S4. Forest plot of clinical utility in relevant subgroups of hospitalized patients (RF-AdminDemo)	29
Figure S5. Forest plot of clinical utility in relevant subgroups of hospitalized patients (RF-Minimal).....	30
Figure S6. Forest plot of clinical utility in relevant subgroups of hospitalized patients (Random)	31
Figure S7. Decision curve analysis with a random hospitalization per patient.....	32
Figure S8. Decision curve analysis with the first hospitalization per patient	33
Figure S9. Decision curve analysis with the last hospitalization per patient	34
References	35

Literature review

Context

Automated predictive analytics in health care is a rapidly growing field. We performed an up-to-date systematized review to find studies that evaluated the clinical utility of an automatable prediction model for prompting a palliative or end-of-life care (PEOLC) intervention. Our objective was to describe if and how such studies had evaluated clinical utility, as recommended for clinical prediction models.^{1,2}

Methods

We searched the PubMed database for studies published in any language from inception to March 1st 2021 using the query:

```
(
  (machine learning)
  OR
  (
    (automated OR (real time) OR computerized OR trigger)
    AND
    ((prediction model) OR electronic health records[mesh] OR (decision support))
  )
)
AND
(
  (Advance Care Planning) OR "code?status" OR "resuscitation order" OR "palliative care" OR "end?of?life"
  OR "goal#?of?care" OR "serious?illness"
)
```

Our inclusion criteria were prediction model validation studies where the use-case for the model was implementation in a clinical decision support system (CDSS) that automatically identifies patients for a PEOLC intervention. We included prospective and post-implementation studies as long as they included model validation methods (including internal validation, external validation, or an evaluation of clinical utility). To assess the quality of evidence, we appraised the applicability of validated models for real-time application at the point of care using the *participant selection* and *predictor* domains of PROBAST.³

Study selection

The query returned 334 results, of which two were relevant systematic reviews,^{4,5} and after screening abstracts, 33 were initially selected for satisfying the inclusion criteria. The following studies were then excluded:

- Six studies pertaining to query-based algorithms,⁶⁻¹¹ which are decreasingly used in favour of more flexible prediction models.
- Three studies using purchased US Medicare claims data,¹²⁻¹⁴ where there was no description of how this data source was accessible for automated predictions at the point of care.
- Two studies of models predicting in-hospital mortality where it was unclear if predictions were intended for earlier palliative care or earlier intensive care.^{15,16}
- Two studies where the primary objective was to investigate prognostic factors rather than validate models.^{17,18}

Study characteristics

We included a final of 20 original research articles and provide a summary in Table S1. There is a climbing interest in the use automated prediction models that use routinely collected data for identifying patients that might benefit from PEOLC interventions. Three studies were published before 2018, three in 2018, six in 2019, five in 2020, and two in the first two months of 2021. Applicable settings were sometimes mixed and included general inpatient populations (9), patients with a specific comorbidity (8; including cancer [5], dementia [1], end-stage liver disease [1], and COVID-19 [1]), general outpatient populations (4), and patients discharged from the ED (2). Five studies reported data after model implementation or application: three pilot studies,¹⁹⁻²¹ and two prospective validation studies,^{22,23}. Four of these had been considered quality improvement projects by their respective IRBs.²⁰⁻²³ The actions that models were intended to prompt were sometimes mixed and included palliative care referral (11), goals-of-care discussion (7), outpatient follow-up with ACP (3) and hospice referral (1). There were seven studies that had used a type of external validation (temporal [1], geographic [3], or prospective [3]). Among the six studies in the

general inpatient setting that predicted mortality in a timeframe of 6-12 months, four had a C-statistic above 0.8 and two had a C-statistic above 0.9.

Risk of bias and concerns for applicability in real-time

Seven of the 15 retrospective validation studies did not adequately describe or correlate retrospective conditions with those an automated CDSS generating predictions at the time and place they would be actionable.²⁴⁻³⁰ Eight studies reported model validation metrics in a population that included subgroups where a PEOLC intervention was inappropriate or unnecessary, and could be excluded in real-time without affecting utility. This including patients in obstetrics^{19-22,25,30}, children,^{25,31} or those already receiving palliative care.^{19,25,28} While training a model on these cases might prove useful for algorithmic learning, validity should be reported in a subgroup that excludes them, especially when the subsequent application cohort is different.^{19,20}

Evaluation of clinical utility

None of the studies had assessed the clinical utility of models using decision-analytic methods, including decision curve analysis. One study used decision analysis after model validation, but for quantifying expected financial value from the perspective of reducing system cost, not its clinical value from the perspective of improving patient outcomes.³²

One study compared model-based predictions with clinical outcomes in usual care: in a clinical utility-related assessment, Parchure et al³³ described rates of inappropriately missing GOC documentation and unmet palliative needs among those with both in-hospital mortality for COVID-19 (outcome-positive) and a predicted risk above a 0.5 threshold for this outcome (high-risk). The study included an assessment of expected benefit among true positives: the benefit of high-risk predictions among outcome-positive cases (9/146 patients with in-hospital mortality could have benefited if every high-risk prediction had led to an appropriate GOC discussion or palliative care referral). However, it did not assess the expected harm of false positives: the harm of high-risk predictions for those that survived to discharge.

All prospective application studies reported the C-statistic as the primary metric to justify the chosen model.¹⁹⁻²¹ None reported a decision analysis quantifying the benefits and harms of prediction-based actions relative to usual care before they had proceeded with implementation.

Conclusion

Automatable prediction models for improving PEOLC outcomes were not assessed for their clinical utility using decision-analysis. Studies were rarely well coordinated between clinical requirements (e.g., ensuring both expected benefit and expected harms are considered), informatics requirements (e.g., ensuring all required data is accessible in real-time) and methodological requirements (e.g., ensuring models were validated in the target population for a PEOLC intervention, which should exclude pediatric visits at the ED or admissions in obstetrics). Most implemented models were part of quality improvement projects, highlighting the need to consider feasibility and local factors in the translational process of developing, validating, implementing, and evaluating the impact of automatable prediction models.

Table S1. Validation studies of automatable or automated prediction models intended to prompt a palliative or end-of-life care intervention

Reference; Year; Country	Design; Population ^a ; Sample size	Model(s) evaluated; Predicted outcome; Validation method	Evaluation metrics ^b (AUROC ^c of best model)	Model-based action; Clinical utility of model ^d	Concerns for applicability at point-of-care ^e
Escobar et al., ³⁴ 2015; USA	Retrospective validation; Inpatients surviving to discharge (excluding those for uncomplicated childbirth); n=360036	4 LR models; 7/30-day nonelective rehospitalization or mortality; Random split (50/50)	AUROC (0.76), Pseudo-R ²	Outpatient follow-up (implied to include ACP); not reported	..
Adelson et al., ³² 2017; USA	Retrospective validation; Inpatients with advanced cancer; n=669	Cox proportional hazards model with proprietary Rothman Index (PeraHealth); 30/60/90/180-day mortality; Random split (70/30)	AUROC (0.74), Sensitivity, PPV, Survival curves	Hospice referral; Potential cost savings if every prompt led to hospice within 48h of admission, but no reporting of patient-centered benefit, or harm of hospice referral for false positives	..
Uneno et al., ²⁴ 2017; Japan	Retrospective validation; Patients with cancer receiving chemotherapy n=2693 internal validation, n=367 geographic validation	LR model; 1/2/3/4/5/6-month mortality; Geographic validation	AUROC (0.85), Sensitivity, Specificity, PPV, NPV, Accuracy	GOC discussion; Explicitly mention clinical utility not assessed, planned for future study	Exclusion of patients without death records; impossible to reproduce exclusion criteria in real-time.
Avati et al., ²⁵ 2018; USA	Retrospective validation; All patients and inpatients only (including children); n=221284 total and n=10779 inpatients only	DNN model; 3-to-12-month mortality; Random split (80/10/10)	AUROC (0.93 for all patients and 0.87 for inpatients only), AUPRC, Calibration plot, Brier score	Palliative care referral; not reported	Exclusion of patients that died within 3 months of prediction time; impossible to reproduce exclusion criteria in real-time. Did not report validity in target sample excluding unapplicable cases such as those already in obstetrics, palliative care, or children
Gensheimer et al., ³⁵ 2018; USA	Retrospective validation; Patients treated for metastatic cancer; n=12588	Cox proportional hazards model; Survival curve and mortality within time points from 0.5 to 5 years; Random split (80/20)	AUROC (0.785), Survival curves	GOC discussion; not reported	..
Sahni et al., ²⁶ 2018; USA	Retrospective validation; Inpatients (excluding non-emergent admissions); n=59848	4 models: LR and 3 RF models; 1-year mortality; Random split (80/20)	AUROC (0.86), Calibration plot	ACP (unclear in what setting, presumed to be outpatient); not reported	Uses post-discharge data to evaluate model performance among inpatients; no description of applicability in outpatient setting
Beeksmas et al., ²⁷ 2019; The Netherlands	Retrospective validation; Outpatients; n=1234	LSTM model; life expectancy; Random split (90/10)	Accuracy	Outpatient ACP; not reported	Excludes patients known to have survived more than 5 years. Impossible to reproduce exclusion criteria in real-time
Blom et al., ³¹ 2019; Sweden	Retrospective validation; Patients discharged from ED (including children); n=65776 internal validation, n=55164 geographic validation	6 models: LR, RF, SVM, KNN, AB, MLP; 30-day mortality; Geographic split	AUROC (0.95; n.b., cohort included children), Sensitivity, Specificity	Outpatient GOC discussion; not reported, but includes value-focused description of intended use-case from multiple stakeholder perspectives	Did not report validity in target sample excluding unapplicable cases such as children

Reference; Year; Country	Design; Population ^a ; Sample size	Model(s) evaluated; Predicted outcome; Validation method	Evaluation metrics ^b (AUROC ^c of best model)	Model-based action; Clinical utility of model ^d	Concerns for applicability at point-of-care ^e
Courtright et al.; ²⁰ 2019; USA ^f	Retrospective validation, CDSS feasibility, and pre-post evaluation; Inpatients (excluding obstetrics, hospice, rehabilitation); n=46305	LR model; 6-month mortality; Random split (85/15%)	AUROC (0.86 for less selective test cohort), Calibration plot, Sensitivity, Specificity, PPV, NPV, F1-score	Palliative care referral; not reported before application	Did not report validity in target sample excluding unapplicable cases such as in obstetrics
Parikh et al.; ³⁶ 2019; USA	Retrospective validation; Outpatients with cancer; n=26525	3 models: LR, GBM and RF; 180/500-day mortality; Random split (70/30)	AUROC (0.88), PPV, Calibration plot, Accuracy, Specificity	GOC discussion; not reported	..
Wang et al.; ²⁸ 2019; USA	Retrospective validation; Patients with dementia; n=26921	LSTM model; 6/12/24 month mortality; Random split (90/10)	AUROC (0.978)	Palliative care referral; not reported	Unclear if ICD codes used for patient inclusion are accessible in real-time. Did not report validity in target sample excluding unapplicable cases such as those already receiving palliative care
Wegier et al.; ¹⁹ 2019; Canada ^f	Retrospective validation, CDSS feasibility, and pre-post evaluation; Inpatients (excluding psychiatry); n=640022 (based on previously used dataset ³⁷)	LR model; 1-year mortality; Bootstrap validation	AUROC (0.89 for less selective test cohort), Pseudo-R ² , Calibration plot, Sensitivity, Specificity, PPV, NPV, Positive LR, Negative LR	GOC discussion or palliative care referral; not reported, but includes value-focused description of intended use-case from multiple stakeholder perspectives	Did not report validity in target sample excluding unapplicable cases such as in obstetrics
Barash et al.; ³⁸ 2020; Israel	Retrospective validation; Patients discharged from ED; n=363635	GBM; 30-day mortality; Temporal split	AUROC (0.97), Sensitivity, FPR	Palliative care referral; not reported ^g	Could not be assessed ^g
Lin et al.; ²⁹ 2020; Taiwan	Retrospective validation; Inpatients with ESLD; n=907 internal validation, n=643 geographic validation	RF and AB models; 30-day and 1-to-9-month mortality; Geographic split	AUROC (0.852)	Palliative care referral; not reported	Unclear if ICD codes used for patient selection are accessible in real-time
Major and Aphinyanaphongs; ²² 2020; USA	Retrospective validation, prospective validation, and CDSS feasibility; Inpatients (excluding hospice care); n=128941	RF model; 60-day mortality; Prospective validation	AUROC (0.87), AUPRC, Calibration plot, PPV	Palliative care referral; harm-to-benefit ratio considered before implementation (preference of 1 TP:3 FP among clinicians) but not applied to evaluate clinical utility	Did not report validity in target sample excluding unapplicable cases such as in obstetrics
Manz et al.; ²³ 2020; USA	Prospective validation; Patients with cancer; n=24582	GBM; 180-day mortality; Prospective validation	AUROC (0.89), PPV, NRI compared to standard prognostic indices (ECOG and Elixhauser)	GOC discussions; not reported before application. Compared performance to standard decision-making policies, but did not consider harm-to-benefit ratio when using NRI ³⁹	..
Parchure et al.; ³³ ; 2020; USA	Retrospective validation; Inpatients with COVID-19; n=567	RF model; 20-to-84-hour mortality; Random split (70/30)	AUROC (0.855), Sensitivity, Specificity, Accuracy, AUPRC, Rate of unmet GOC/palliative needs	GOC discussion or palliative care referral; reporting of clinical benefit among cases with in-hospital mortality, but	..

Reference; Year; Country	Design; Population ^a ; Sample size	Model(s) evaluated; Predicted outcome; Validation method	Evaluation metrics ^b (AUROC ^c of best model)	Model-based action; Clinical utility of model ^d	Concerns for applicability at point-of-care ^e
				no reporting or discussion of potential harms	
Blanes-Selva et al., ⁴⁰ 2021, Spain	Retrospective validation; Inpatients (excluding psychiatry and obstetrics); n=65279	5 models: GBM, RF, KNN, SVM, MLP; 1-year mortality; Random split (80/20)	AUROC (0.91), Accuracy, Sensitivity, Specificity, Balanced Error Rate	Palliative care referral; not reported	Unclear if validation sample includes palliative care admissions
Guo et al., ³⁰ 2021; USA	Retrospective validation; Inpatients and outpatients; n=59639	4 models: LSTM, RF, DNN, LR; 1-year mortality; Random split (80/10/10)	AUROC (0.92 for outpatients, 0.97 for inpatients), AUPRC, Accuracy, Sensitivity, Specificity, PPV, F1-score	Palliative care referral; not reported	Describes delay of 3-6 months for administrative claims data, but not described if and how this was handled (intended focus of future study). Did not report validity in target sample excluding unapplicable cases such as in obstetrics
Murphree et al., ²¹ 2021; USA ^f	Retrospective validation, prospective validation, and CDSS feasibility; Inpatients; n=50143	GBM model; Historic palliative care referral; Random split (80/20) then prospective validation	AUROC (0.91), PPV, Calibration plot	Palliative care referral; not reported before application	Did not report validity in target sample excluding unapplicable cases such as in obstetrics

Abbreviations: AB, adaptive boosting; ACP, advance care planning; AUPRC, area under the precision recall curve; AUROC, area under the receiver operating characteristic; CDSS, clinical decision support system; DNN, deep neural network; ESLD, end-stage liver disease; GBM, gradient boosting machine; GOC, goals-of-care; ICD, international classification of disease; KNN, k-nearest neighbours; LR, logistic regression; LSTM, long short-term memory; MLP, multilayer perceptron; NPV, negative predictive value; NRI, net reclassification index; PPV, positive predictive value; RF, random forest; SVM, support vector machine; ^a Excluding children (age < 18 years) unless otherwise indicated.

^b Evaluation of model performance or utility. Not included are descriptive metrics (like variable importance measures) or metrics of construct validity (like physician opinion of appropriateness)

^c Also known as the C-statistic.

^d Quantitative assessment of benefits and harms of model-based actions relative to a standard policy.

^e Concerns that validated models do not use data of the target population that is accessible at the intended time and setting of decision support.

^f Post-application study that included a validation cohort and application cohort. All sample sizes and metrics correspond to assessment in the latest validation cohort.

^g Only the abstract could be reviewed for this study that was not yet in print at the time of review; full text could not be accessed via institutional library, via request to authors, nor via attempted purchase from journal website, which was impeded by internal server errors over a week, and not fixed before review completion.

Methods

Setting

The study took place at the *Centre Hospitalier Universitaire de Sherbrooke* (CHUS), a public university hospital network in the city of Sherbrooke, Quebec, Canada. It is the only institution offering acute hospitalisation services in the city and is the only tertiary care centre for approximately 1.2 million people in a large catchment area (Eastern Townships),⁴¹ resulting in relatively complete longitudinal data. It is partitioned into two sites, *CHUS-Fleurimont* (site A) and *CHUS-Hôtel-Dieu* (site B). Site A generally offers more specialized services than site B, although the staff for a given speciality cover both.

Data collection and analysis period

The start date of July 1st, 2011 coincides with the availability of all necessary predictors in the data warehouse and a large window to use as a training period (five years). The end date of June 30th, 2018 reflects the last date of hospital admissions where the outcome of one-year mortality could be ascertained before the start of analysis in July 2019.

Mortality outcome

The binary outcome predicted by all models was the occurrence of death less than 365 days after admission. We considered this timeframe acceptable for approximating EOL status and potential eligibility for palliative interventions.^{19,25} Mortality outcome data was collected from the provincial vital statistics registry (*Régie de l'assurance maladie de Québec*) for all patients in the data warehouse with a provincial health number, i.e., official death certificates for all residents of Quebec in the overall cohort (99.5% of hospitalizations). This was cross-referenced with in-hospital mortality records, which includes those for patients with a residence outside of Quebec. Since it was possible for mortality records up until June 30th, 2019 to be missing for a subsequent data extraction in July 2019, death certificates up until that date were re-extracted in October 2020. This increased the prevalence of one-year mortality measured in the testing cohort by 8.2% (308 of 3773 hospitalizations), presumably due to administrative delays. We used the latest mortality data extraction for all analyses and consider it having complete case outcomes.

mHOMR model

Wegier et al¹⁹ demonstrated the feasibility of integrating the Modified Hospital One-year Mortality Risk (mHOMR) model into a computerized decision support system (CDSS) for the purpose of motivating discussions about goals of care (GOC). This logistic regression model had been derived from the HOMR model.^{37,42} The rationale for modifying the original model was that HOMR included predictors that were coded after hospital discharge and not available for an automated prediction model at admission, including the admission diagnosis, chronic comorbidities and the need for home oxygen. After excluding these predictors, it was feasible to apply the modified model in a CDSS at their institution. The relatively few types of predictors required made it easy to reproduce the mHOMR model at our institution. It was the only mortality prediction model that was both generalizable to our institution and had been tested in a prospective setting for the purpose of prompting a PEOLC-related intervention.

The mHOMR model was reproduced as closely as possible to the published specification. Two categorical variables, the admitting service and living status, were relevelled to align more closely with local practices and to accommodate real time-available data, respectively. All variable transformations (e.g., inverse of the square root of the number of ED visits plus one) and interactions (e.g., between living status and number of ambulance admissions in the last 12 months), were kept as is.

The living status variable was based on the last known value of this variable coded from previous hospitalizations, if present (after simulating a delay, as described later). About half of values were missing and these were imputed as “Unknown”, with the assumption that the pattern of missingness – whatever it may be – was transportable to future application at our institution.⁴³ The living status variable was considered too important for outright exclusion, and its inclusion, even with missingness, improved performance in the training cohort.

All other variable definitions were identical to the original specification. The model was fitted on 1 000 two-stage bootstrap⁴⁴ samples of unique patients in the training cohort, where the first stage is the selection of a random patient, with replacement, and the second stage is the selection of a random visit of that patient, with replacement. A simpler technique for fitting a logistic regression model on a large amount of longitudinal data is the selection of a single encounter per patient (e.g., randomly,⁴² or the last observed⁴⁵), and discarding the data from all other encounters. In contrast, the two-stage bootstrap procedure allowed considering all hospitalizations for training the model while satisfying the assumption of I.I.D. observations. As was done for the published mHOMR model, the final coefficients were the average bootstrapped coefficients for each variable.

Machine learning model

The predictors used by the ML model reflect the same type of information used in the original HOMR model, although sourced and pre-processed differently. In mHOMR, predictors that required diagnostic codes from discharge abstracts had been excluded. While the information from the discharge abstract of an index hospitalization is not available at the time of admission, it may be available for subsequent hospitalizations after some delay due to administrative processing. We surmised that this information, particularly about chronic comorbidities, was important for risk prediction, even if delayed. This was our main reason for developing a new prediction model. Modifications to the original set of variables in HOMR were guided by (1) operational constraints, such as the availability of electronic data in real-time, (2) clinical judgement, such as the grouping of diagnostic codes for personalized risk prediction, and (3) model assumptions, such as the increased flexibility of machine learning algorithms relative to logistic regression.

A total of 244 predictors were included in the final ML model. Among these variables, 147 represented medical terms mapped from a free-text admission diagnosis accessible in real-time from the EHR (Table S2) and 84 represented diagnostic groups mapped from comorbidity information accessible in real-time from administrative databases (Table S3). One variable flagged the absence of accessible comorbidity information. Among the remainder were variables related to demographics (age, sex and living status), previous care utilization (number of weeks hospitalized in the last 90 days, number of admissions by ambulance in the last 12 months and number of ED visits in the last 12 months) and characteristics of the current admission (admission during flu season, admission type, urgent 30-day readmission, admission via ambulance, ICU admission and admitting service). Unlike with the mHOMR model, continuous variables were not transformed before inclusion and no interactions were specified.

Two of these variables were not included in the original HOMR model: if the admission is during the peak months of the local influenza flu season (which we defined as from December to February, inclusively) and the number of full weeks of hospitalisation in the last 90 days. The first variable was motivated by the overall increased prevalence of mortality observed in these months. We reasoned, prior to the COVID-19 pandemic, that including this information could improve the calibration of a machine learning model. The second variable was to add more information on recent care utilization, which we considered important for predicting mortality.

The admission diagnosis was sourced from the free-text transcription of this information in the EHR at the time of admission. A terminology mapping was created manually between this free text variable and a set of medical terms. This was an iterative process conducted by a physician (RT) where the goal was to group common or prognostically relevant expressions without necessarily being exhaustive. Observations in the testing cohort were not included in this process. The resulting mapping is available in Table S2. Each term was included in the model as a binary variable and each hospitalization could have zero or more positive terms. At least one term could be mapped for 90% of admissions in the training cohort. There was no imputation done for admissions without any mapped terms.

We opted to provide the model with relevant diagnostic information from previous visits after simulating an appropriate delay, rather than exclude this information entirely. To inform this delay, the administrative processes that led to electronic data generation were investigated for variables that were not expected to be available at the time of hospital admission. We delayed the visibility of diagnoses from the discharge abstract of a hospitalization by six months after discharge (the usual delay for codification is around three months) and the visibility of the diagnosis from a pre-hospitalization ED stay by two weeks after ED departure (the usual delay for codification is less than one week). Simulating a longer delay than necessary would facilitate the reproducibility of training conditions when the model would be initially applied. Diagnoses coded for a hospitalization that ended less than five years before an index hospitalization were accessible to the index hospitalization (a “lookback” period of five years), and discharge abstracts since January 1st, 2007 were available for such a lookback.

The Charlson Comorbidity Index (CCI) was used to group comorbidities in the original HOMR model. The CCI and similar diagnostic grouping algorithms (e.g., Elixhauser comorbidities) have been used in several prediction models for personalized prediction.^{22,36,46} While empirically accurate when averaging risk in large cohorts,⁴⁵ we did not consider its grouping appropriate for our clinical application. For example, the Charlson comorbidity “Any malignancy, including lymphoma and leukemia, except malignant neoplasm of skin” (Elixhauser comorbidity: “Solid tumor without metastasis”) includes both patients with pancreatic cancer (C25) and patients with thyroid cancer (C73), while “Chronic pulmonary disease” (identical Elixhauser comorbidity) includes both patients with chronic obstructive pulmonary disease (J44) and patients with asthma (J45).⁴⁷ The diseases in these groups may have similar diagnostic codes, but they do not have similar prognoses.⁴⁸ Alternatively, a machine learning model

could include diagnoses with little restriction,^{25,49} resulting in thousands of distinct diagnostic codes. This strategy might increase predictive performance but the increase in dimensions increases susceptibility to overfitting⁵⁰ and requires extra effort to maintain interpretability for our use case.^{51,52} We decided to take an intermediate approach, where diagnostic codes were regrouped by a physician (RT) into more prognostically similar subsets of disease: taking advantage of the flexibility of machine learning algorithms but using clinical judgement for every variable added. We surmised that this strategy would yield an acceptable balance between predictive performance and model interpretability. The source of data was considered in the grouping, with a distinction made between (1) diagnoses coded from a pre-hospitalization ED stay (diagnostic codes from ED stays without subsequent hospitalizations could not be extracted), (2) diagnoses coded from the “Principal diagnosis explaining admission” field of discharge abstracts and (3) all diagnoses coded from discharge abstracts (including the principal diagnosis). As an illustration, the prognostic significance of a pulmonary embolism is different if “pulmonary embolism” was the reason for a hospitalization via the ED one month ago, versus if “pulmonary embolism” was mentioned as a past event on a discharge abstract one year ago. ED-associated codes originated from the local instance of the *Système d'information de gestion des urgences* database⁵³ while discharge abstract-related codes originated from the local instance of the *Maintenance et exploitation des données pour l'étude de la clientèle hospitalière* database⁵⁴ (both of which were accessible via the institutional data warehouse). The resulting groupings for each source are available in Table S3.

The codification of the living status variable followed the same timeline as the discharge abstract codification and its visibility was also delayed for six months after discharge. Like for the mHOMR model, the value was imputed as “Unknown” when there was no living status visible for a given hospitalization.

Real-time access from provincial databases was not deemed feasible for the home oxygen variable, and so it was inferred from the ICD-10 code for dependence on supplemental oxygen (Z99.8) from discharge abstracts, subject to the same six-month delay. Archive personnel confirmed the specificity of this code for indicating the mention of home oxygen on the discharge abstracts of our institution.

The ML model was trained with the Random Forest (RF) algorithm.⁵⁵ This method was chosen as we found more complex machine learning algorithms (e.g., gradient tree boosting) had modest performance gain, a finding similar to that of others³⁶, but required more effort to minimize overfitting. Trees were grown using the probability forest method.⁵⁶ The number of trees was set to 512 and the number of variables to try per split was set at the default setting of 15 (the floor of the square root of the number of predictors; this was done after ensuring neighbouring values were not more appropriate using a random split in the training cohort). Factor variables were kept as is (rather than one-hot encoded) and partitioned after ordering by outcome.^{57(p310)} The sampling algorithm was modified such that the in-bag sample of each tree was selected using the two-stage bootstrap procedure described above.

Model validation

We assessed the predictive validity of the models for the outcome of one-year mortality by analyzing usual metrics: discrimination using the C-statistic and calibration with using a calibration plot. We also report the Brier score as an overall metric of validity, although note its limitations.⁵⁸ Internal validity is reported to provide an indication of model reproducibility had we not employed temporally external validation, which is more representative of actual performance at the point of care.⁵⁹ Note that we did not use internal validation to test optimal variable selection; all variables were either prespecified (originating from HOMR or mHOMR) or based on clinical judgement with reproducible definitions.

K-fold cross-validation ($k = 10$) was used to estimate the internal validity of the models.⁵⁰ Folds were constructed by randomly splitting the set of unique patients into ten groups, then rejoining all hospitalizations to their corresponding patient. This prevented the possibility of leaking a given patient's information between training and testing sets for the k^{th} model. The reported performance measures were obtained by calculating a given statistic in each fold after selecting one random hospitalization per patient.

External validity was estimated in 1 000 two-stage bootstrap resamples of the testing cohort. Note that a given model was trained with the training cohort once and applied to the testing cohort once; each iteration of the bootstrap resampled from the predictions fixed to each observation in the testing cohort. The median value was used as the point estimate and the bias-corrected and accelerated (BCa) method⁶⁰ was used to construct 95% confidence intervals.

Valuation of outcomes

There are various benefits that can be attributed to the discussion of goals-of-care (GOC) among patients at the end-of-life (EOL), including an improved quality of life, reduced use of aggressive therapy near death, improved adjustment to loss by family, and care that is generally more consistent with patient preferences.⁶¹ Conversely, the absence of GOC discussion in the same population can have the opposite, harmful effects,⁶¹ and could even be considered a medical error of omission.⁶² The main cost of discussing GOC is clinician time,⁶¹ a constrained resource.⁶³

We used the documentation of code status orders (CSO), encoding the basic preferences for life sustaining therapy, as a process measure of GOC discussions. When a CSO had been documented, it was assumed that a minimal discussion about GOC had occurred. If no CSO had been documented during a hospitalization, then a GOC discussion had not been documented as per hospital policy (CSO documented in the EHR and all other details documented in the paper chart).

A useful prediction model would maximize the benefit of discussing and documenting GOC for patients at the EOL (true positive CSOs) while minimizing the cost of relatively inefficient resource allocation, i.e., when this action occurs for patients not at the EOL (false positive CSOs). We used the recommended⁶⁴⁻⁶⁷ methods described by Vickers⁶⁸ to value these outcomes on the same scale, relative to the benefit of true positives. The value of a true positive is one beneficial outcome, representing the various benefits of a GOC discussion for a patient at the EOL. The value of a false positive is one harmful outcome that is equal in magnitude to a beneficial outcome multiplied by a TP:FP exchange rate.^{69,70}

Clinical prediction models are usually intended to be used by clinicians, who incorporate risk and patient preferences to advise patients about proceeding or not with a given intervention (like biopsy or not biopsy).⁷⁰ In our context of quality improvement, however, the intention of a prediction model is not to ultimately advise *patients* on some course of action – e.g., between a full code and a DNR: we leave that decision-making between clinicians and patients – but to advise *clinicians* on some course of action: initiate a GOC discussion for a patient at risk of being at the EOL. In this scenario, the exchange rate represents the acceptable number of patients who are at the EOL with whom to discuss GOC (nTP) in exchange for doing so with patients who are not at the EOL (nFP); the exchange rate is equal to nTP:nFP. An equivalent interpretation is the acceptable odds, $p/(1-p)$, of a patient being at the EOL (p = probability of being at the EOL) in order to act in favor of discussing GOC; $p = \text{nTP}/(\text{nTP}+\text{nFP})$. For example, if it were considered acceptable to have ten GOC discussions if at least one would be for a patient at the EOL, then the exchange rate is 1:9. Equivalently stated, it would be acceptable to spend the time having a GOC discussion with a patient who had at least a 10% chance ($p = 1/10$) of being at the EOL. In expected utility theory,⁶³ this would imply that the harm of inaction is nine times worse than the harm of action: a rational decision-maker would be expected to act if the chance that inaction leads to harm was greater than 10%.

Decision-analytic assumptions

To simplify our analysis, we assumed that alerts would deterministically cause the action of CSO documentation (after an appropriate GOC discussion). In reality, it is unlikely that all alerts would be heeded, but this assumption allows us to estimate the maximum benefit attributable to alerts.

While we simulated counterfactual changes in CSO status based on simulated alerts, we did not alter any association between cases and factual EOL outcomes. Communication about GOC and EOL preferences can be associated with increases in life expectancy,⁶¹ but we assumed our definition of EOL status, one-year mortality, to be relatively robust to any such effects. We did not expect that alerts leading to CSO documentation would shorten life to less than a year for those that would have otherwise lived longer or prolong life to more than a year for those that would have otherwise died before.

Confidence intervals of clinical utility estimates

A non-parametric bootstrap with 1000 resamples of the CDSS-eligible cohort was used to generate the 95% confidence intervals for all clinical utility estimates (adjusted with the BCa procedure). All intermediate calculation steps were repeated in each bootstrap, including the threshold that resulted from a desired P(Alert) of 10% and the observed exchange rate using Equation 4.

Statistical software extensions

The *tidyverse* collection of packages⁷¹ was used for data inspection and preparation. The *ranger* package⁷² was used to train the Random Forest models. The two-stage bootstrap described previously was implemented in C++ and

linked to R using the *Rcpp* package.⁷³ The BCa bootstrap procedure was implemented using the source code of the *mediate* function of the *mediation* package.⁷⁴ The *ggplot2*, *cowplot* and *forestplot* packages⁷⁵⁻⁷⁷ were used for data visualization and generating plots.

Table S2. Admission diagnosis–mapped variables included in RF-AdminDemoDx model^a

Variable	Free-text expression ^b	ICD-9 expression ^c
Abcess	abces abcede collection	
Abdominal pain	(douleur dlr)(?:(.+)abdo epigast)	789[.]0
Acute coronary syndrome	(syn(?:(.+)bcoro) ((angor angine)(?:(.+)instable) (coronarien) (stemi) (s(W)?()?t(\W)?()?e(W)?()?m(W)?()?) (\bsca\b) (\infarct(?:(.+)myocard) infarctus (hemo dynamic) (\bptca\b) coronaro	410[.]411[.](9[^(0)] [^(9)])
Acute leukemia	leucemie \blma\b \blla\b	20[458][.][09]
Alcohol	\broh\b alcool ethyl delirium tremen	
Anemia	anemie anémique	285[.]9 280[.]9
Aneurism	anevrism anevrism	
Angina	(angine) (angor) \bmcas\b maladie(?:(.+)cardiaque	411[.]90 413[.]429[.]21 414[.]0
Aortic aneurism	\baaa\b anevrism(?:(.+)aorte	441[.]4
Appendicitis	app[ea]ndicit appendicect appendiculaire	54[013][.]9
ARDS	\bards\b	
Arrhythmia	arythmi arrythmi (\bcve\b) (cardioversion)	427[.]479 426[.]7
Arthropathy	arthro tendon articula coxalgie ankylose	715[.]
Ascites	\bascit	789[.]5
Aspiration pneumonia	(aspiration)	507[.]0
Asthma	asthm[ea] broncho(W)?spas[tm]	493[.]519.1
Atrial fibrillation	(\bfa(p)?\b) (\bflutter\b) (\bfibri(?:(.+)aur\w+)	427[.]3
Bariatric	obesite(?:(.+)severe morbid)	
Benign tumor		2[12][0-9][.]
Biliary colic	colique(?:(.+)bilair cholelithias cholestase	574[.]2
Bladder cancer	((cancer) carcinome (\bneo) (masse) (\bca\b))(?:(.+)vessie vesical urothelial ureteral)	
Brain cancer	(\bgbm\b) (\glioblasto) astrocytom \bglome	191[.]
Brain hemorrhage	(\bhip\b) (\bhematom hemor)(?:(.+)bft\b occip cerebral front parietal temporal cerebell fosse parench thalam cortical (intra(.)?(ventric cranien)))(\bavc\b \bacv\b accident cerebro)(?:(.+)hemor)	431[.]
Brain injury	(\bhsd\b) (\bhematom hemor)(?:(.+)dural) btcc\b (trauma contusion)(?:(.+)bcranie n cerebra)	432[.]1 85[14][.]0
Brain lesion	(masse lesion tumeur)(?:(.+)cerebral frontal cerebell parietal occipital temporal)	
Breast cancer	((cancer) (\bcarcinome) (\bdenoca) (\bdenok) (\bneo) (ad(n)?k) (\bca\b))(?:(.+)sein\b mammaire canalaire lobulaire)	174[.]9
Bronchiectasis	bronchiect	
Bronchitis	bronchite bronchiolite surinfection bronchiq	
C. Difficile	\bc(?:(.+)diff	
Cancer	(cancer) (\bdenoca) (\bdenok) (\bneo) (\bad(n)?k\b) neoplasi bcarcino anaplasia sarcome (\bca\b) hodgkin (\blnh\b) (\blymphome) \blma\b melanome leucemie	
Carotid stenosis	st(?:(.+)carotid bci[gd]\b) endarterect	433[.]12
Cellulitis	\bcellulit	682[.]67
Chemotherapy	(\bcotx\b) (\chimio) (\bcia\b)	
Chest pain	\bdrs\b douleur(?:(.+)thora retro(W)?stern)	786[.]5[01]
Cholangitis	cholangit	576[.]
Cholecystitis	cholecystit cholecystect	575[.]
Choledocholithiasis	choledocho \bcpre\b (calcul lithias)(?:(.+)choledo)	574[.]5
Chronic leukemia	(synd\w+(?:(.+)myelo lympho) (\bwaldenstrom) \bllc\b	20[45][.]1

Variable	Free-text expression ^b	ICD-9 expression ^c
Cirrhosis	\bcirrho varice(?:.+oeso hemorrr(?:.+(varic oeso))\bencephalop(?:.+hepat	571[.][5]572[.]21456[.]1
Colitis	\bcolite pancolite	
Colorectal cancer	((cancer)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(tumeur)(masse)(\bca\b))(?:.+)(coliq \bcaec sigmoid \bcolo rect appendice \bgrele\b)	15[34][.]
Conduction abnormality	bloc(?:.+)[23]e complet\b[a[. -]?v\b auricul)\bb[.]?a[.]?v\b mobitz bradycard \bb[.]?b[.]?g\b(\bsick sinus\b)(\bsss\b)maladie(?:.+sinus(\bpause(s)?\b)(tachy(?:.*brady)	426[.]0-689
COPD	(mpoc\b)(malad\w+ pulm\w+ obstr\w+ chron\w+)(m(\W)?()?p(\W)?()?o(\W)?()?c) emphysem	
Cytopenia	pancytopeni bicytopeni	
Delirium	(\bdelire\b)(delirium) confusion((alteration perte etat)(?:.+conscience))\b agitati on \baec\b	293[.]307[.]9348[.]3331[.]0
Dementia	\bdemenc alzheimer(\bt(?:.+(cognitif comportement))\btnc\b berrance\b	29[04][.]310[.]1
Diabetes	diabet(\bd[bm]\b) hyperglycemie hypoglycemie etat hyper[]?(gly osm)	250[.]09251[.]2790[.]6250[.]
Dialysis	irc irc terminal \bdialyse	
Diarrhea	(diarrh)	
Disk disorder	hernie(?:.+)(disc dls d) dissectomie discoidect discopathie discarthro radiculopathe lombo-sciatalgie	722[.]724[.]3
Diverticular disease	diverticul(it os ai)	562[.]
DVT	emboli(?:.+) pulm \bt[vp]p\b (h)?rombophleb thrombose veineuse phlebite embolie massive ep massive	415[.]451[.]28
Dysphagia	\bachalasiae \bdysphagi	787[.]2
Dyspnea	\bdyspne	799[.]0786[.]03
Ear disorder	mastoidit vertige \boto labyrinth cholesteatom neuron(?:.+) vestibul \bsurdite\b \bmeniere\b	38[0-9][.]780[.]42
Electrolytes	(hypo hyper)(na ca k mg mag) deshydrat	276[.]1276[.]0276[.]8276[.]7275[.]42
Endocarditis	endocardit	424[.]9421[.]01
ENT cancer	((cancer)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(\bca\b))(?:.+)(\borl\b larynx pharynx langue)	161[.]
Enteritis	enterit \bileite	
EOL care	(palliati) soins confort phase terminal	
Electrophysiology study	(e p s)(electrophysio)(\beps\b) \bdefib\b pace(?:.+) defib pace(?:.+) maker \bpac\b (\bpmp\b)	
Esophageal cancer	((cancer)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(\bca\b))(?:.+)(oesophag)	150[.]9
Esophageal varices	varice(?:.+oeso hemorrr(?:.+(varic oeso)	456[.]1
Eye	oeil \bretin ophthalm	
Falls	\bchute(s)?\bt(?:.+) marche\b	781[.]2
Febrile neutropenia	(neutro pancytopenie)(?:.+) febr fièvre(?:.+) neutro	288[.]02
Fertility	fertilite	
Fracture	(fractur)(\bfx\b)(\b[#]\b) rofi	
Gastric cancer	((cancer)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(\bca\b))(?:.+)(gastr estomac)	151[.]9
Gastritis	gastrit	535[.]5
GI bleed	\bhd[hb]\b rectorragie melena hemorrr(?:.+) digest	578[.]569[.]31
Guillain Barré syndrome	guil(?:.+) bar mill(?:.+) fis \bci dp\b	357[.]0
Gynecological disorder	\badenomyos ligature \bendometriosis endometrite \bhysterect	62[56][.]620[.]2617[.]0621[.]3
Heart failure	(\bins\w* cardia)(def\w+ cardiaq)(dysf\w+(?:.+)(\bv g\b ventricul diastol systol) (\bcmp\b)(ic(?:.+) decompens diastol systol)) surcharge (oedem(?:.+) pulm poumon) (\boap\b) anasarque	414[.]89[.]428[.]19[.]518[.]4812[.]425[.]4429[.]9276[.]60
Hemoptysis	hemoptys	786[.]3

Variable	Free-text expression ^b	ICD-9 expression ^c
Hepatic failure	\bins(?:+)hepatiq	573[.]81
Hepatitis	hepatit	573[.]3 571[.]1
Hip fracture	(([#] fx fracture rofi)(?:+)(hanche femur femoral troch intertro (bas)(?:+)(cervic peri(-?)pro (sous)(?:+)(capi)	
Hypertension	hypertensi\bhta\b	
Infection	(\b(sur)?infect)\bfievre febrile hypertherm temperature	
Inguinal hernia	hernie(s)? ing	550[.]
Intestinal ischemia	((enterite colite)(?:+)ischemique)((ischemie angine)(?:+)(mesenter intestin))	557[.] 09
Intestinal polyp	polyp(?:+)(coliq\bcaec sigmoid\bcolo rect appendice\bgrele\b intestin\bileo)	211[.]3
Intoxication	(poly)?intox sevraige (dro[gu]e medicam) toxicom abus drogue	977[.] 291[.] 30 345[.] 292[.]0
Joint prosthesis	\bpth\b prot\w+ tot\w+	
Liver cancer	cholangiocarc carcinome hepat hepato(-?)carcinome hepatome\bchc\b	
Loss of autonomy	(pert\w+ ((d'autonomie)((dauton\w+ auton\w+))) deconditionnement epuisement	
Lower leg fracture	(([#] fx fracture rofi)(?:+)(cheville tibia rotule mal(l)?(col[.]) perone jambe genou ped patella maisonneuve)	
Lumbar pelvis fracture	(([#] fx fracture burst)(?:+)(pelvi pubi\bacetab bassin\b (\W)?[1-5] lombaire sacrum)	80 58[.] .
Lung cancer	((cancer)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(?:+)(poumon pulmon\b si gd\b lobe (inf sup moy))	16 23[.] .
Lung mass	(masse nodule lesion tumeur)(?:+)(poumon pulmon\b si gd\b lobe (inf sup moy))\bbtt\b biopsie(?:+)(trans(?:+)(thorac	
Lymphoma	lymphome\b lnh\b	202[.]
Melanoma	\bmelanom	172[.]
Meningitis	meningit encephalit	32 0-4[.] .
Metastasis	metasta[st] carcinomatos epanchement(?:+)(neo bst(?:+)(b(4 iv)(.)? mesothel iom	19 789[.] .
Inflammatory bowel disease	(crohn)(colite ulc)	555[.] 556[.]19
Multiple myeloma	(myelom\w+ multip)	203[.]
Osteomyelitis	osteomyelit (plaie ped)(?:+)(infect diabet)	730[.]2
Osteoporosis	osteoporo	733[.] 01
Other hernia	hernie (inc omb)	55 1-3[.] .
Pancreatic cancer	((cancer)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(\bca\b))(?:+)(pancrea)	157[.]9
Pancreatic mass	(masse tumeur lesion)(?:+)(pancrea(s)tique)	
Pancreatitis	pancreatite	577[.]0
Parkinson's	parkinson	332[.]0
Perforation	perforation(?:+)(intestin grele colon colique caec jeju duoden ilea)(ulcere ulcus divertic chole appendic)\bperfore pneumoperit	
Pericardial effusion	epanchement(?:+)(pericard	
Pericarditis	(myo peri)(?:+)(cardite	423[.]91
Pleural effusion	\bepanchement(?:+)(pleural pleureux)	511[.] 0-9
Pneumonia	(pneumonie) empyeme	48 0123568[.] .
Pneumothorax	pneumothora\b bptx\b	512[.]
Pregnancy	travail (cesarien)(grossess)(\d+ semaine)(rupt\w+ prematu\w+)(\brppm\b)(rciu) (placenta)(accouche)(post([-])?partum)(gestation) fertilit	6 3-7 0-9 [vV]22[.] 2
Prolapsus	\bprolap(a)?sus (relachement relaxation)(?:+)(pelvien)(cysto recto colpo entero)(?:+)(cele\b procidence\b colpo	618[.] 788[.]3
Prostate cancer	((cancer)(carcinome)(\badenoca)(\badenok)(\bneo)(ad(n?)k)(\bca\b))(?:+)(prostata prostatact	185[.]

Variable	Free-text expression ^b	ICD-9 expression ^c
Pulmonary fibrosis	(fibrose)(?:(pulm))\bfp\b	515[.]
Pulmonary hypertension	((hypertension htn) pulm)\bhpt\b	
PVD	(ische(?:(\bgr)(membre)(pied)(mollet)))(claudic)((\bins(?:(isch(?:(art)(vascul)(mif dg))))(amputat)\bfrein\b(\b(arterio angio)(\b graph plast))	443[.]9 440[.]9 459[.]932 785[.]4
PVD and gangrene		785[.]4
PVD and arterial insufficiency		443[.]9 440[.]9
PVD and ischemia		459[.]932
Reanimation	\bacr\b arret cardi reanimation	427[.]5 798[.]1
Renal failure	(\bi(\W)?r(\W)?a\b)(\bi(\W)?r(\W)?c\b)(\bins(?:(*) renal)	58[1-3][.]584[.]90 586[.]9
Respiratory failure	((\bins detresse difficulte)(?:(respirat))\bdesatur hypoxemi	786[.]01 786[.]0\b
Seizures	status(?:(epile convuls partiel))(epilep(s)t)i	780[.]3 345[.]9
Sepsis	(sepsis)(\bseptiqu)	038[.]0-489 790[.]7
Severe	\bsevere\b	
Shock	\bchoc\b(?:(ana)(?! hem)(?! hypo)	785[.]5
Spondylopathy	stenose(?:(foram spinal lombaire cervical)[cdt s d])myelopathie laminect compression medul(l)?aire foraminect spondylolisthesi	
Stroke	(\bavc\b)(a v c)(acc(?:(cerebr(?:(\bavc\b)(a(\W)?v(\W)?c) \bavc\b acc(?:(+))\basc(?:(cere(?:(+))	436[.]9 784[.]3 342[.]
Subarachnoid hemorrhage	(hematom hemor)(?:(arachnoid)\bhsa\b	430[.]90
Syncope or orthostatic hypotension	\bsyncop tension(?:(orthosta lipoth[ymi]	780[.]2 458[.]0 458[.]9
Tachycardia	(\bt(\W)?s(\W)?v)(tachycardie) \bt(\W)?v\b arythmie(?:(ventric(\bt(\W)?a(\W)?p)	427[.]0-2
Tamponade	tamponnade tamponade	
Thyroid cancer	((cancer carcinome)(\bneo)(masse) nodule tumeur(\bca\b)(?:(thyroid folliculaire papillaire))thyroïdect	193[.]
Transient ischemic attack	(\bi(\W)?c(\W)?t\b)((isch paresie)(?:(transitoir) amauros(?:(fuga	435[.]
Tonsillitis	amygdalite amygdalect	
Trauma	trauma laceration plaie (luxation) (rupture)	95[89][.]86 0-9[.]
Trigeminal neuralgia	nevralgie(?:(trijum	
Tumor	(tumeur)(nodule)(masse)(lesion)(adenome) meningiome adenopathie	
Urinary lithiasis	calcul lithias colique neph coraliform nephrostomie \bj\b	
Urinary retention	retention	788[.]2
Urological procedure	\brtuv\b \brtuv\b bhp \brtup\b \bhp\b \baps\b prostate \bhbp\b hématurie	
Urinary tract infection	(\bpna\b)(pyelonephrit)(p n a)(urosepsis)(\bcystite)(infec\w+ urinaire)	590[.]18 599[.]0 595[.]9 601
Valve prosthesis	bentall tiron(?:(david)(\brvm\b)(\brva\b)	
Valve regurgitation	(\bins(?:(regurg(?:(mitral aort)	424[.]01 1
Valve stenosis	(\bstenos(?:(mitral aort \bao\b))\bsa\b	424[.]13 39 45[.]10
Virus	\bsag\b (synd(?:(gripp viral))grippe grippal virus rhume influenz \bi(\W)?v(\W)?r(\W)?s\b	465[.]91 487[.]1 079[.]9
Weight loss or fatigue	perte de poids(\bdeg\b)((dim\w+)(det\w+)(?:(etat g))faiblesse etat(?:(ge neral)\basthenie	783[.]2 780[.]7

Abbreviations: ARDS, acute respiratory distress syndrome; COPD, chronic obstructive pulmonary disease; DVT, deep venous thrombosis; ED, emergency department; ENT, ear, nose, throat; EOL, end-of-life; GI, gastro-intestinal; ICU, intensive care unit; PVD, peripheral vascular disease.

^a Admission diagnoses are transcribed from the physician-written order form into the EHR by administrative clerks at the time of admission. The EHR has a list of standard ICD-9 codes that may be attached to the admission diagnosis. When they are selected, both the code and its standard label are concatenated with the free-text admission diagnosis by the EHR. The resulting, real-time-accessible variable is an unstandardized string of characters that was

preprocessed to make all letters lowercase and to strip accents. Each row defines a binary variable included as a predictor in the RF-AdminDemoDx model. For each variable, if either expression matches the preprocessed string, the variable value is true (or 1), otherwise it is false (or 0).

^b Regular expressions are intended for use with admission diagnoses in french, including terms and abbreviations that are specific to medical language in Quebec.

^c Regular expressions that match ICD-9 codes embedded in admission diagnosis string.

Table S3. Comorbidity diagnosis–mapped variables included in the RF-AdminDemoDx model^a

Variable ^b	ED diagnosis ^c	Summary diagnosis ^d	Principal diagnosis ^c
Atrial fibrillation	I48	I48	
Acute respiratory failure	J960	J960	
Alcohol-related diagnosis	F10 E52 G621 I426 K292 Z721	F10 E52 G621 I426 K292 Z721	
Anasarca	E877 R60	E877 R60	
Anemia		D5[1-3] D649 D638	
Recent angina	I20		
Anticoagulation		Z921	
Aortic stenosis		I350	
Ascites	R18	R18	
Asthma	J45	J45	
Breast cancer		C50 Z853	
Bronchiectasis	J47	J47	
Cachexia	R64 R53	R64 R54	
CAD		I20 I25 Z955	
Chemotherapy-related diagnosis 1		Z926	
Chemotherapy-related diagnosis 2	Z511 D700	Z511	
Chest cancer 1		Z85[12] C3[0-39]	
Chest cancer 2	C34	C3[4-8]	
CHF		I099 I11 I13 I255 I42[05679] I43 I50 P290	
CHF-related admission	I099 I11 I13 I255 I42[05679] I43 I50 P290		I099 I11 I13 I255 I42[05679] I43 I50 P290
Chronic respiratory failure	J96[19]	J96[19]	
CNS cancer	C71	C71	
COPD/emphysema	J44[01] J43	J44 J43	
CVD	G4[56] H340 I6[0-9]	G4[56] H340 I6[0-9]	
Dementia		F0[0-3] F051 G30 G31	
Malnutrition		E4[36]	
Depression		F204 F31[345] F32 F33 F341 F412 F432	
Dialysis		Z992 Z49	
Diabetes-related diagnosis 1		E1[0-4][09]	
Diabetes-related diagnosis 2	E1[0-4] G632 H360 N083	E1[0-4][1-8] G632 H360 N083	
Endocrine cancer		C7[3-5]	
Falls	R296 W06 W1[89] R268	R296 W06 W1[89]	
Frailty	R64 R53 R296 W06 W1[89] R268	E4[36] R64 R54 R296 W06 W1[89]	
GI cancer 1		Z850 C1[789] C20	
GI cancer 2	C18	C1[56] C2[4]	
GI cancer 3		C2[12356]	
GU cancer 1		C5[124] C560 C6[1467] Z85[45]	
GU cancer 2	C57 C61	C5[3] C6[02358]	
GU cancer 3		C5[57] C56[1-9]	
Hematological cancer 1		C8[12358] Z85[67] C921 C9[16]	
Hematological cancer 2		C8[46] C9[04]	
Hematological cancer 3	C8[1-9] C9[0-6]	C9[35] C92[04589]	
Home oxygen		Z998	
Crohn's disease	K50		K50
Interstitial lung disease		J84	
Hepatic disease 1		B18 K70[01239] K71[3457] K7[34] K759 K76[2-489] Z944 K86[01]	
Hepatic disease 2	K7[1-6] Z944	K704 K711 K72 K758 K76[57]	
Hepatic disease risk factors		I85[09] I864 I98[23] K766	
Metastatic solid cancer		C7[789] C80	
Past MI		I21 I22 I252	
Recent MI	I21 I22 I252		
Musculoskeletal cancer		C4[015-9]	
Obesity-hypoventilation syndrome		E662	
Obstetrics-related diagnosis	O Z37		O Z37
ENT cancer		C0[1-9] C1[0-4]	
Palliative		Z515	
Paralysis		G041 G114 G80[12] G8[12] G83[0-49]	

Variable ^b	ED diagnosis ^c	Summary diagnosis ^d	Principal diagnosis ^e
Past pulmonary embolism			I26
Respiratory-related admission			J4[0-467] J684 J70[13]
Pneumonia admission			J1[2-8]
Pseudomonas-related diagnosis		B965 J151	
Psychiatric disorder		F2[02-589] F302 F31[25]	
Pulmonary hypertension	I27	I27	
Peripheral vascular disease	I7[01] I73[189] I771 I79[02] K55[189] Z95[89]	I7[01] I73[189] I771 I79[02] K55[189] Z95[89]	
Recent abdominal pain	R104		
Recent anemia	D649		
Recent back pain	M545		
Recent cancer-related diagnosis	C50 Z853 C767		
Recent chest pain	R074		
Recent colitis	A099		
Recent complication	T8		
Recent GI bleed	K922		
Recent hip fracture	S72		
Recent interstitial nephritis	N10		
Recent intestinal occlusion	K56		
Recent pulmonary embolism	I26		
Recent perforation	K578		
Recent pneumonia	J1[2-8]		
Recent urinary tract infection	N390		
Renal disease 1	N189	I12 I131 N03[2-7] N05[2-7] N18[1238] N19 Z940	
Renal disease 2		N18[045] N250	
Skin cancer	C44	C4[34]	
Transplantation-related diagnosis		Z94[0-4]	
Valvular disease		I0[5-8] I3[4-9] Z95[2-4]	

Abbreviations: CAD, coronary artery disease; CHF, congestive heart failure; CNS, central nervous system; COPD, chronic pulmonary obstructive disease, CVD, cerebrovascular disease; ED, emergency department; ENT, ear, nose, throat; GI, gastro-intestinal; GU, genito-urinary; MI, myocardial infarction.

^a Diagnoses given as space-delimited regular expressions intended to be applied to period-stripped ICD-10 codes. Each row defines a binary variable included as a predictor in the RF-AdminDemoDx model. For each row, if any of the expressions match any of their respectively sourced ICD-10 codes (ED, summary or principal), the variable value is true (or 1), otherwise it is false (or 0).

^b Numbered variables are categorized by prognostic similarity; higher numbers generally indicate diagnoses with worse prognosis.

^c Diagnoses associated with an ED visit that led to hospitalization between six months and two weeks before index admission.

^d Any diagnosis on a hospital discharge abstract between five years and six months before index admission.

^e Principal diagnosis on a hospital discharge abstract between five years and six months before index admission.

Table S4. Characteristics of hospitalizations between July 2011 and July 2016 (Training cohort)^a

	Overall (n=122 860)	Vital status one year after admission	
		Dead (n=18 541)	Alive (n=104 319)
Age, median (IQR), y	64 (45-76)	75 (64-84)	62 (40-74)
Sex			
Male	55 695 (45)	9 981 (54)	45 714 (44)
Female	67 165 (55)	8 560 (46)	58 605 (56)
Hospital site			
A	79 868 (65)	11 260 (61)	68 608 (66)
B	42 992 (35)	7 281 (39)	35 711 (34)
Service type			
Medical	59 763 (49)	13 469 (73)	46 294 (44)
Surgical	45 085 (37)	4 561 (25)	40 524 (39)
Critical care ^b	1 479 (1)	500 (3)	979 (1)
Obstetrics	16 533 (13)	11 (0)	16 522 (16)
Admission type			
Non-elective	83 125 (68)	16 771 (90)	66 354 (64)
Elective	24 136 (20)	1 764 (10)	22 372 (21)
Obstetrics	15 599 (13)	6 (0)	15 593 (15)
Real-time-accessible living status ^c			
Home	58 214 (47)	11 046 (60)	47 168 (45)
Nursing home	1 938 (2)	688 (4)	1 250 (1)
Chronic care hospital	952 (1)	414 (2)	538 (1)
Unknown	61 756 (50)	6 393 (34)	55 363 (53)
Living status at discharge			
Home	61 055 (50)	5 401 (29)	55 654 (53)
Home with CLSC liaison	42 590 (35)	4 699 (25)	37 891 (36)
Short term transitional care	6 574 (5)	1 439 (8)	5 135 (5)
Nursing home	4 589 (4)	1 087 (6)	3 502 (3)
Chronic care hospital	2 226 (2)	857 (5)	1 369 (1)
Other ^d	1 161 (1)	393 (2)	768 (1)
Death in hospital	4 665 (4)	4 665 (25)	0 (0)
ED visits ^e			
0	73 817 (60)	7 796 (42)	66 021 (63)
1-2	35 847 (29)	6 842 (37)	29 005 (28)
3 or more	13 196 (11)	3 903 (21)	9 293 (9)
Admissions by ambulance ^e			
0	105 595 (86)	12 979 (70)	92 616 (89)
1-2	14 439 (12)	4 432 (24)	10 007 (10)
3 or more	2 826 (2)	1 130 (6)	1 696 (2)
Weeks recently hospitalized ^f			
0	105 957 (86)	12 690 (68)	93 267 (89)
1-2	12 700 (10)	4 157 (22)	8 543 (8)

	Overall (n=122 860)	Vital status one year after admission	
		Dead (n=18 541)	Alive (n=104 319)
3 or more	4 203 (3)	1 694 (9)	2 509 (2)
Visible comorbidities ^c	67 865 (55)	13 602 (73)	54 263 (52)
Admission during flu season ^c	30 280 (25)	4 736 (26)	25 544 (24)
ED admission	62 448 (51)	12 956 (70)	49 492 (47)
Ambulance admission	36 333 (30)	9 368 (51)	26 965 (26)
Urgent 30-day readmission	11 493 (9)	3 810 (21)	7 683 (7)
ICU admission	3 173 (3)	811 (4)	2 362 (2)
ICU stay during hospitalization	17 119 (14)	3 368 (18)	13 751 (13)
Length of stay, median (IQR), d	3 (2-7)	7 (3-14)	3 (2-6)
Major comorbidities ^g			
Congestive heart failure	12 202 (10)	3 769 (20)	8 433 (8)
Chronic pulmonary disease	22 181 (18)	5 709 (31)	16 472 (16)
Dementia	7 324 (6)	2 459 (13)	4 865 (5)
Metastatic cancer	8 618 (7)	5 192 (28)	3 426 (3)

Abbreviations: CLSC, *Centre local de services communautaires*; DNI, do-not-intubate; DNR, do-not-resuscitate; ED, emergency department; EOL, end of life.

^a Data given as number (percentage) of hospitalizations unless otherwise indicated. Percentages may not add to 100 due to rounding.

^b Represent direct admissions to the ICU before a primary non-critical care service could be specified (i.e., the responsible service upon ICU discharge). ICU exposure is more precisely measured with the variables “ICU admission” and “ICU stay during hospitalization”.

^c See description in Table 1 (main text).

^d Includes transfer to another hospital, to a rehabilitation center, to a palliative care center, or discharge against medical advice.

^e In the year before admission.

^f In the 90 days before admission.

^g Charlson comorbidities using ICD-10 codes by Quan et al⁴⁷ and ascertained using the discharge abstract of index hospitalization and of those in the year before discharge.

Table S5. Characteristics of hospitalizations between July 2017 and July 2018 (Testing cohort)^a

	Overall (n=26 291)	Vital status one year after admission	
		Dead (n=3 968)	Alive (n=22 323)
Age, median (IQR), y	65 (48-76)	75 (65-85)	63 (43-74)
Sex			
Male	12 176 (46)	2 170 (55)	10 006 (45)
Female	14 115 (54)	1 798 (45)	12 317 (55)
Hospital site			
A	17 054 (65)	2 441 (62)	14 613 (65)
B	9 237 (35)	1 527 (38)	7 710 (35)
Service type			
Medical	12 721 (48)	2 952 (74)	9 769 (44)
Surgical	9 773 (37)	811 (20)	8 962 (40)
Critical care ^b	621 (2)	204 (5)	417 (2)
Obstetrics	3 176 (12)	1 (0)	3 175 (14)
Admission type			
Non-elective	17 733 (67)	3 613 (91)	14 120 (63)
Elective	5 607 (21)	354 (9)	5 253 (24)
Obstetrics	2 951 (11)	1 (0)	2 950 (13)
Real-time-accessible living status ^c			
Home	12 852 (49)	2 300 (58)	10 552 (47)
Nursing home	1 293 (5)	451 (11)	842 (4)
Chronic care hospital	297 (1)	111 (3)	186 (1)
Unknown	11 849 (45)	1 106 (28)	10 743 (48)
Living status at discharge			
Home	12 452 (47)	1 005 (25)	11 447 (51)
Home with CLSC liaison	8 878 (34)	899 (23)	7 979 (36)
Short term transitional care	1 410 (5)	306 (8)	1 104 (5)
Nursing home	1 622 (6)	355 (9)	1 267 (6)
Chronic care hospital	535 (2)	227 (6)	308 (1)
Other ^d	339 (1)	121 (3)	218 (1)
Death	1 055 (4)	1 055 (27)	0 (0)
ED visits ^e			
0	15 827 (60)	1 646 (41)	14 181 (64)
1-2	7 471 (28)	1 396 (35)	6 075 (27)
3 or more	2 993 (11)	926 (23)	2 067 (9)
Admissions by ambulance ^e			
0	22 695 (86)	2 759 (70)	19 936 (89)
1-2	2 986 (11)	955 (24)	2 031 (9)
3 or more	610 (2)	254 (6)	356 (2)
Weeks recently hospitalized ^f			
0	22 664 (86)	2 720 (69)	19 944 (89)
1-2	2 775 (11)	896 (23)	1 879 (8)

	Overall (n=26 291)	Vital status one year after admission	
		Dead (n=3 968)	Alive (n=22 323)
3 or more	852 (3)	352 (9)	500 (2)
Past discharge abstract accessible ^c	14 410 (55)	2 932 (74)	11 478 (51)
Admission during flu season ^c	6 517 (25)	1 061 (27)	5 456 (24)
ED admission	13 268 (50)	2 814 (71)	10 454 (47)
Ambulance admission	7 592 (29)	2 005 (51)	5 587 (25)
Urgent 30-day readmission	2 521 (10)	851 (21)	1 670 (7)
ICU admission	1 082 (4)	245 (6)	837 (4)
ICU stay during hospitalization	3 588 (14)	719 (18)	2 869 (13)
Length of stay, median (IQR), d	3 (2-7)	7 (3-14)	3 (2-6)
Code status preference ^g			
Full code	2 331 (9)	285 (7)	2 046 (9)
DNR/Intubation-OK	937 (4)	263 (7)	674 (3)
DNR/DNI	4 505 (17)	2 159 (54)	2 346 (11)
Not documented	18 518 (70)	1 261 (32)	17 257 (77)
Major comorbidity ^h			
Congestive heart failure	2 826 (11)	960 (24)	1 866 (8)
Chronic pulmonary disease	4 834 (18)	1 235 (31)	3 599 (16)
Dementia	1 836 (7)	660 (17)	1 176 (5)
Metastatic cancer	1 891 (7)	1 020 (26)	871 (4)

Abbreviations: CLSC, *Centre local de services communautaires*; DNI, do-not-intubate; DNR, do-not-resuscitate; ED, emergency department; EOL, end of life.

^a Data given as number (percentage) of hospitalizations unless otherwise indicated. Percentages may not add to 100 due to rounding.

^b Represent direct admissions to the ICU before a primary non-critical care service could be specified (i.e., the responsible service upon ICU discharge). ICU exposure is more precisely measured with the variables “ICU admission” and “ICU stay during hospitalization”.

^c See description in Table 1 (main text).

^d Includes transfer to another hospital, to a rehabilitation center, to a palliative care center, or discharge against medical advice.

^e In the year before admission.

^f In the 90 days before admission.

^g Last preference documented during hospitalization if one was documented.

^h Charlson comorbidities using ICD-10 codes by Quan et al.⁴⁷ and ascertained using discharge abstract of index hospitalization and of those in the year before discharge.

Table S6. Clinical utility of prediction models with a random hospitalization per patient

	RF-AdminDemoDx	mHOMR	RF-AdminDemo	RF-Minimal	Random
Threshold	0.423 (0.417-0.431)	0.414 (0.407-0.420)	0.413 (0.405-0.418)	0.398 (0.394-0.411)	0.899 (0.892-0.903)
No. Alerts	1649 (1637-1664)	1649 (1638-1664)	1649 (1637-1664)	1651 (1641-1675)	1649 (1638-1665)
eRD, %	9.1 (8.0-10.3)	5.7 (4.7-6.5)	6.3 (5.3-7.4)	5.4 (4.5-6.3)	3.0 (2.3-3.7)
eRR	1.13 (1.11-1.14)	1.08 (1.06-1.09)	1.09 (1.07-1.10)	1.07 (1.06-1.09)	1.04 (1.03-1.05)
Benefit	0.1111 (0.1062-0.1155)	0.1064 (0.1017-0.1105)	0.1073 (0.1026-0.1115)	0.1061 (0.1012-0.1104)	0.1027 (0.0979-0.1071)
Harm	0.1064 (0.1016-0.1111)	0.1092 (0.1041-0.1139)	0.1076 (0.1026-0.1124)	0.1099 (0.1047-0.1147)	0.1286 (0.1225-0.1344)
Net benefit	0.0047 (0.0028-0.0067)	-0.0027 (-0.0048 to -0.0012)	-0.0004 (-0.0023-0.0016)	-0.0038 (-0.0058 to -0.0021)	-0.0258 (-0.0284 to -0.0236)
Standardized net benefit ^a	0.0347 (0.0210-0.0492)	-0.0202 (-0.0356 to -0.0083)	-0.0027 (-0.0164-0.0113)	-0.0282 (-0.0423 to -0.0155)	-0.1895 (-0.2061 to -0.1738)
NNB, Alerts	8.0 (7.1-9.3)	12.9 (11.3-16.0)	11.6 (10.1-14.0)	13.5 (11.9-16.8)	24.6 (20.0-33.1)
PPV, %	56.9 (54.4-59.3)	40.9 (38.4-43.0)	44.2 (41.8-46.6)	39.4 (37.1-41.6)	13.9 (12.3-15.7)
NPV, %	91.2 (90.7-91.7)	89.4 (88.9-89.9)	89.8 (89.3-90.3)	89.2 (88.8-89.8)	86.4 (85.9-87.0)
Sensitivity, %	41.8 (40.1-43.3)	30.0 (28.4-31.6)	32.4 (30.8-34.2)	28.9 (27.4-30.6)	10.2 (9.0-11.4)
Specificity, %	95.0 (94.7-95.3)	93.2 (92.9-93.4)	93.5 (93.3-93.8)	93.0 (92.7-93.2)	90.0 (89.8-90.2)
C-statistic ^b	0.86 (0.85-0.87)	0.80 (0.79-0.81)	0.81 (0.80-0.82)	0.79 (0.78-0.80)	0.50 (0.49-0.51)

Abbreviations: eRR, expected relative risk; eRD, expected risk difference; NNB, number needed to benefit; NPV, negative predictive value; PPV, positive predictive value.

Data presented as point estimate (95% CI). For all models: Sample size = 16490 patients; Threshold set such that $P(\text{Alert}) = 10\%$; TP CSOs in usual care = 1627 (1546-1700); FP CSOs in usual care = 3614 (3507-3723); Observed exchange rate = 0.450 (0.424-0.474); Prevalence of cases at the EOL = 13.6% (13.1%-14.1%); Random alert rule included as a point of reference, not as a strategy under consideration.

^a Calculated as the net benefit divided by the prevalence of EOL status (1-year mortality).

^b Threshold independent.

Table S7. Clinical utility of prediction models with the first hospitalization per patient

	RF-AdminDemoDx	mHOMR	RF-AdminDemo	RF-Minimal	Random
Threshold	0.403 (0.395-0.410)	0.391 (0.385-0.398)	0.385 (0.379-0.391)	0.394 (0.391-0.404)	0.899 (0.894-0.905)
No. Alerts	1649 (1637-1665)	1649 (1637-1664)	1649 (1637-1665)	1661 (1642-1676)	1649 (1638-1664)
eRD, %	10.3 (9.1-11.5)	6.2 (5.2-7.2)	6.4 (5.3-7.4)	6.3 (5.3-7.3)	3.3 (2.7-4.1)
eRR	1.15 (1.13-1.17)	1.09 (1.08-1.11)	1.09 (1.08-1.11)	1.09 (1.08-1.11)	1.05 (1.04-1.06)
Benefit	0.1045 (0.1002-0.1090)	0.0992 (0.0949-0.1036)	0.0995 (0.0949-0.1037)	0.0993 (0.0949-0.1036)	0.0954 (0.0913-0.0999)
Harm	0.0987 (0.0945-0.1036)	0.1006 (0.0964-0.1058)	0.0991 (0.0950-0.1042)	0.1011 (0.0967-0.1061)	0.1188 (0.1134-0.1247)
Net benefit	0.0059 (0.0038-0.0079)	-0.0014 (-0.0034-0.0002)	0.0003 (-0.0016-0.0019)	-0.0018 (-0.0038 to -0.0001)	-0.0234 (-0.0257 to -0.0212)
Standardized net benefit ^a	0.0445 (0.0288-0.0592)	-0.0109 (-0.0255-0.0014)	0.0024 (-0.0120-0.0143)	-0.0135 (-0.0291 to -0.0008)	-0.1766 (-0.1926 to -0.1615)
NNB, Alerts	7.4 (6.6-8.3)	12.2 (10.5-14.5)	11.8 (10.3-14.2)	12.0 (10.5-14.6)	22.6 (18.3-28.8)
PPV, %	53.1 (50.5-55.6)	39.1 (36.6-41.5)	40.8 (38.1-43.0)	38.2 (35.8-40.4)	12.8 (11.2-14.6)
NPV, %	91.2 (90.7-91.6)	89.6 (89.1-90.1)	89.8 (89.3-90.3)	89.5 (89.0-90.0)	86.7 (86.2-87.3)
Sensitivity, %	40.1 (38.4-41.8)	29.5 (27.8-31.1)	30.8 (29.0-32.4)	29.0 (27.2-30.5)	9.7 (8.5-10.9)
Specificity, %	94.6 (94.3-94.9)	93.0 (92.7-93.2)	93.2 (92.9-93.4)	92.8 (92.6-93.1)	89.9 (89.8-90.1)
C-statistic ^b	0.85 (0.84-0.86)	0.79 (0.78-0.80)	0.80 (0.79-0.81)	0.79 (0.78-0.80)	0.49 (0.48-0.51)

Abbreviations: eRR, expected relative risk, eRD, expected risk difference; NNB, number needed to benefit; NPV, negative predictive value; PPV, positive predictive value.

Data presented as point estimate (95% CI). For all models: Sample size = 16490 patients; Threshold set such that $P(\text{Alert}) = 10\%$; TP CSOs in usual care = 1500 (1436-1574); FP CSOs in usual care = 3653 (3539-3765); Observed exchange rate = 0.411 (0.388-0.437); Prevalence of cases at the EOL = 13.2% (12.8%-13.8%); Random alert rule included as a point of reference, not as a strategy under consideration.

^a Calculated as the net benefit divided by the prevalence of EOL status (1-year mortality).

^b Threshold independent.

Table S8. Clinical utility of prediction models with the last hospitalization per patient

	RF-AdminDemoDx	mHOMR	RF-AdminDemo	RF-Minimal	Random
Threshold	0.439 (0.432-0.446)	0.432 (0.425-0.438)	0.433 (0.427-0.439)	0.399 (0.394-0.410)	0.899 (0.893-0.904)
No. Alerts	1649 (1637-1665)	1649 (1638-1665)	1649 (1638-1665)	1651 (1639-1672)	1649 (1637-1664)
eRD, %	7.4 (6.4-8.5)	4.7 (3.8-5.6)	5.3 (4.5-6.4)	4.6 (3.8-5.4)	2.4 (1.8-3.0)
eRR	1.10 (1.08-1.11)	1.06 (1.05-1.07)	1.07 (1.06-1.08)	1.06 (1.05-1.07)	1.03 (1.02-1.04)
Benefit	0.1178 (0.1135-0.1230)	0.1140 (0.1094-0.1187)	0.1149 (0.1106-0.1197)	0.1139 (0.1093-0.1190)	0.1108 (0.1062-0.1156)
Harm	0.1151 (0.1102-0.1203)	0.1180 (0.1132-0.1235)	0.1169 (0.1119-0.1222)	0.1189 (0.1138-0.1241)	0.1394 (0.1331-0.1460)
Net benefit	0.0027 (0.0008-0.0047)	-0.0040 (-0.0059 to -0.0021)	-0.0019 (-0.0036 to -0.0001)	-0.0049 (-0.0068 to -0.0031)	-0.0286 (-0.0312 to -0.0260)
Standardized net benefit ^a	0.0194 (0.0056-0.0327)	-0.0287 (-0.0416 to -0.0150)	-0.0138 (-0.0258 to -0.0007)	-0.0350 (-0.0483 to -0.0223)	-0.2037 (-0.2190 to -0.1879)
NNB, Alerts	9.6 (8.5-11.4)	15.3 (13.0-19.1)	13.4 (11.3-16.1)	15.4 (13.2-19.1)	30.0 (24.3-42.3)
PPV, %	61.1 (58.6-63.7)	44.5 (42.2-47.0)	47.8 (45.6-50.4)	42.3 (39.7-44.5)	13.9 (12.3-15.7)
NPV, %	91.2 (90.7-91.6)	89.3 (88.8-89.8)	89.7 (89.2-90.2)	89.1 (88.6-89.5)	85.9 (85.3-86.5)
Sensitivity, %	43.5 (41.9-45.1)	31.7 (30.2-33.1)	34.0 (32.6-35.6)	30.1 (28.5-31.5)	9.9 (8.8-11.1)
Specificity, %	95.5 (95.2-95.8)	93.5 (93.3-93.8)	93.9 (93.7-94.2)	93.3 (93.0-93.5)	90.0 (89.8-90.2)
C-statistic ^b	0.87 (0.86-0.88)	0.81 (0.80-0.82)	0.82 (0.82-0.83)	0.80 (0.79-0.81)	0.50 (0.48-0.51)

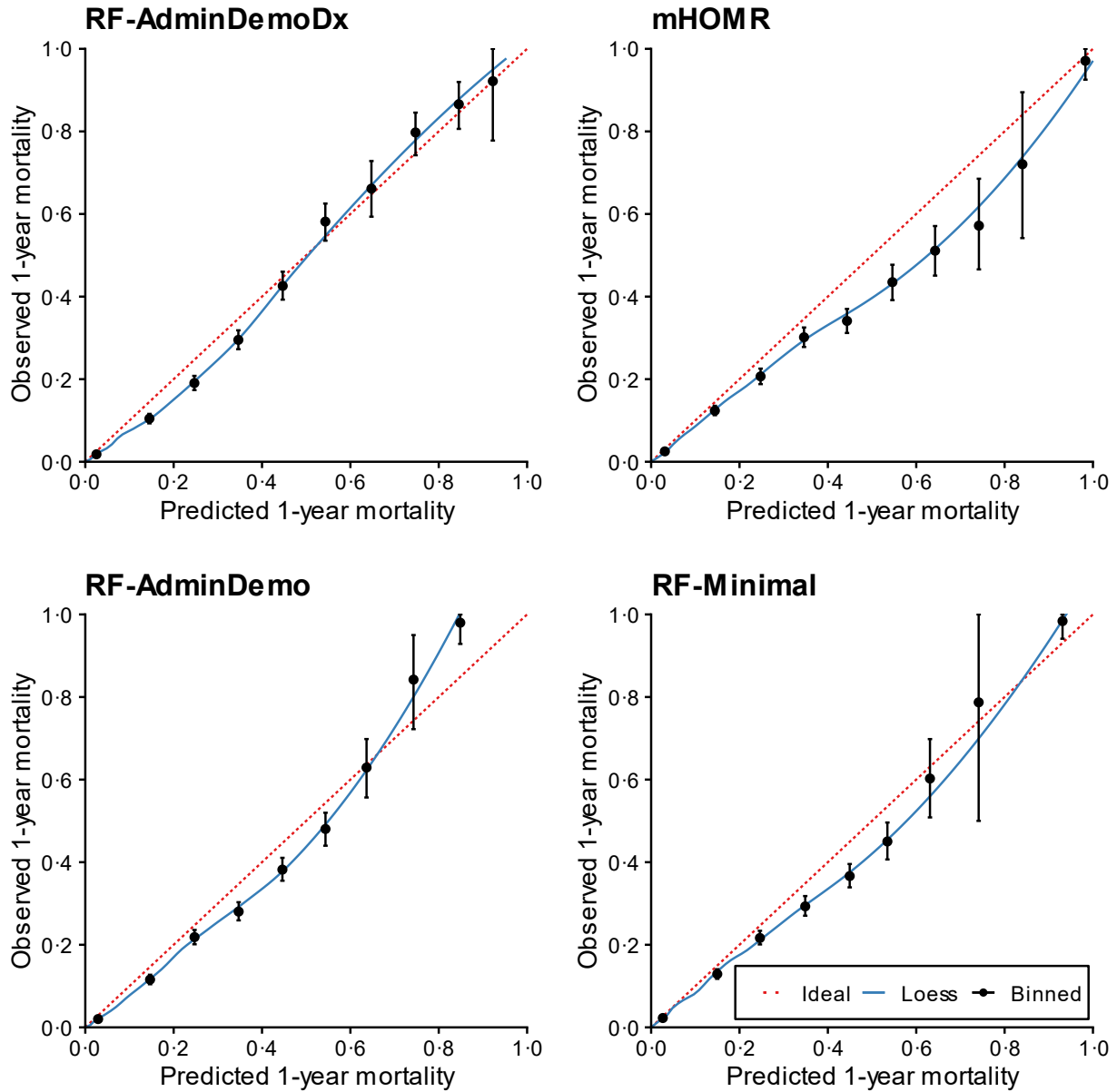
Abbreviations: eRR, expected relative risk, eRD, expected risk difference; NNB, number needed to benefit; NPV, negative predictive value; PPV, positive predictive value.

Data presented as point estimate (95% CI). For all models: Sample size = 16490 patients; Threshold set such that $P(\text{Alert}) = 10\%$; TP CSOs in usual care = 1772 (1698-1853); FP CSOs in usual care = 3626 (3512-3735); Observed exchange rate = 0.489 (0.461-0.518); Prevalence of cases at the EOL = 14.1% (13.6%-14.6%); Random alert rule included as a point of reference, not as a strategy under consideration.

^a Calculated as the net benefit divided by the prevalence of EOL status (1-year mortality).

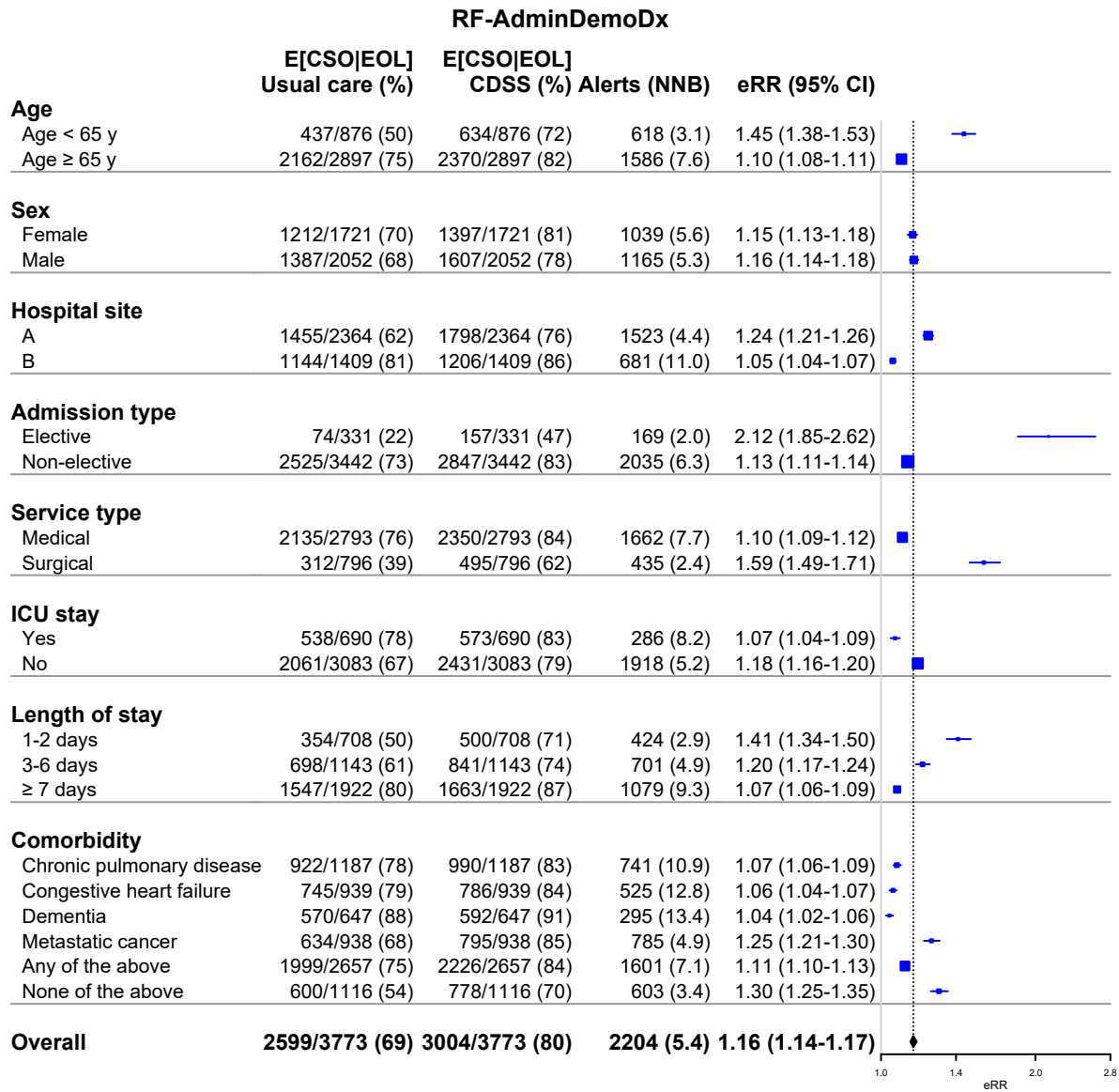
^b Threshold independent.

Figure S1. Calibration plot of prediction models in testing cohort



Proportion with the outcome (y-axis) in ten binned groups (x-axis). Bins with less than five observations removed. Binned estimates generated with 1 000 two-stage bootstrap resamples of the testing cohort; mean predicted probability and mean observed proportion over bootstraps plotted on x- and y-axis, respectively. Error bars (95% CI) plotted using corresponding percentiles for bootstraps in each bin. Non-parametric curve generated by selecting one random hospitalization per patient and applying the loess algorithm (*span* set to 0.4) to individual observed outcome (0 or 1) vs. individual predicted probability.

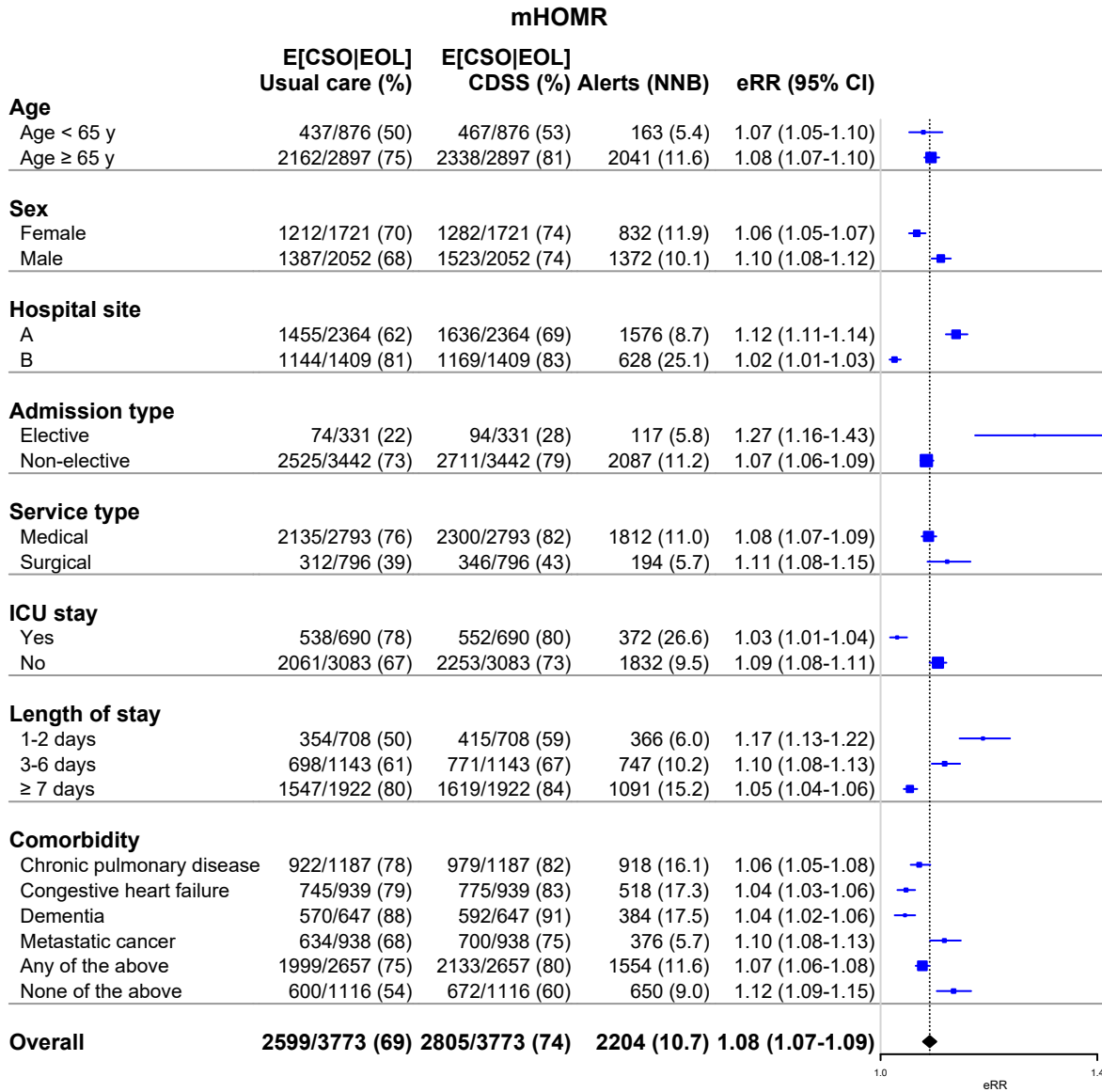
Figure S2. Forest plot of clinical utility in relevant subgroups of hospitalized patients (RF-AdminDemoDx)



Abbreviations: CDSS, clinical decision support system; CSO, code status order; EOL, end of life; NNB, number needed to benefit.

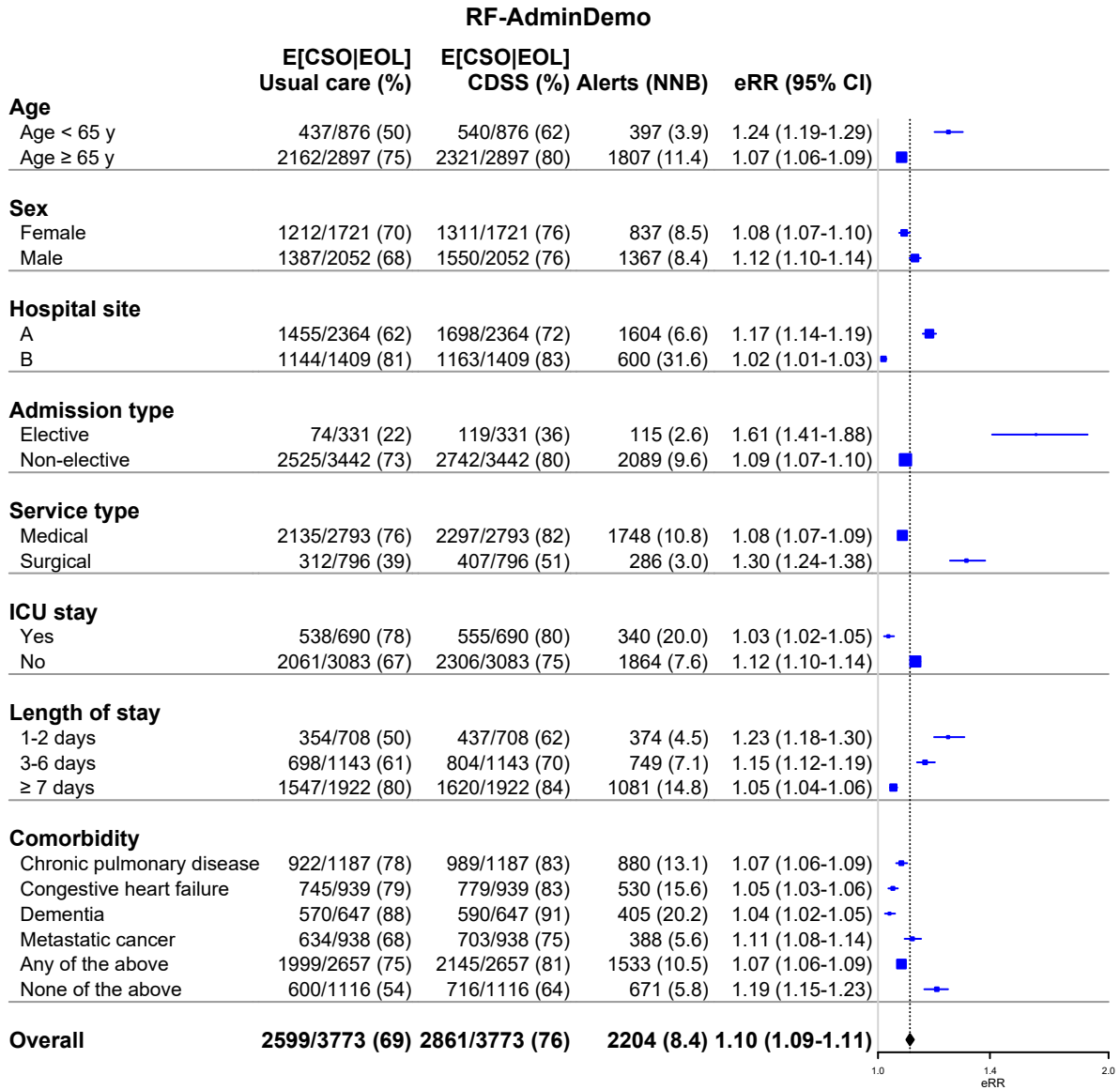
The term E[CSO | EOL] is the expected number of CSOs among patients at the EOL for a given strategy using decision tree 2 and is presented as $n(\text{CSO} \& \text{EOL})/n(\text{EOL})$. The difference in this proportion between “Usual care” and “CDSS” is the eRD. Horizontal error bars (95% CI) generated with 1000 bootstrap resamples of hospitalizations in each subgroup.

Figure S3. Forest plot of clinical utility in relevant subgroups of hospitalized patients (mHOMR)



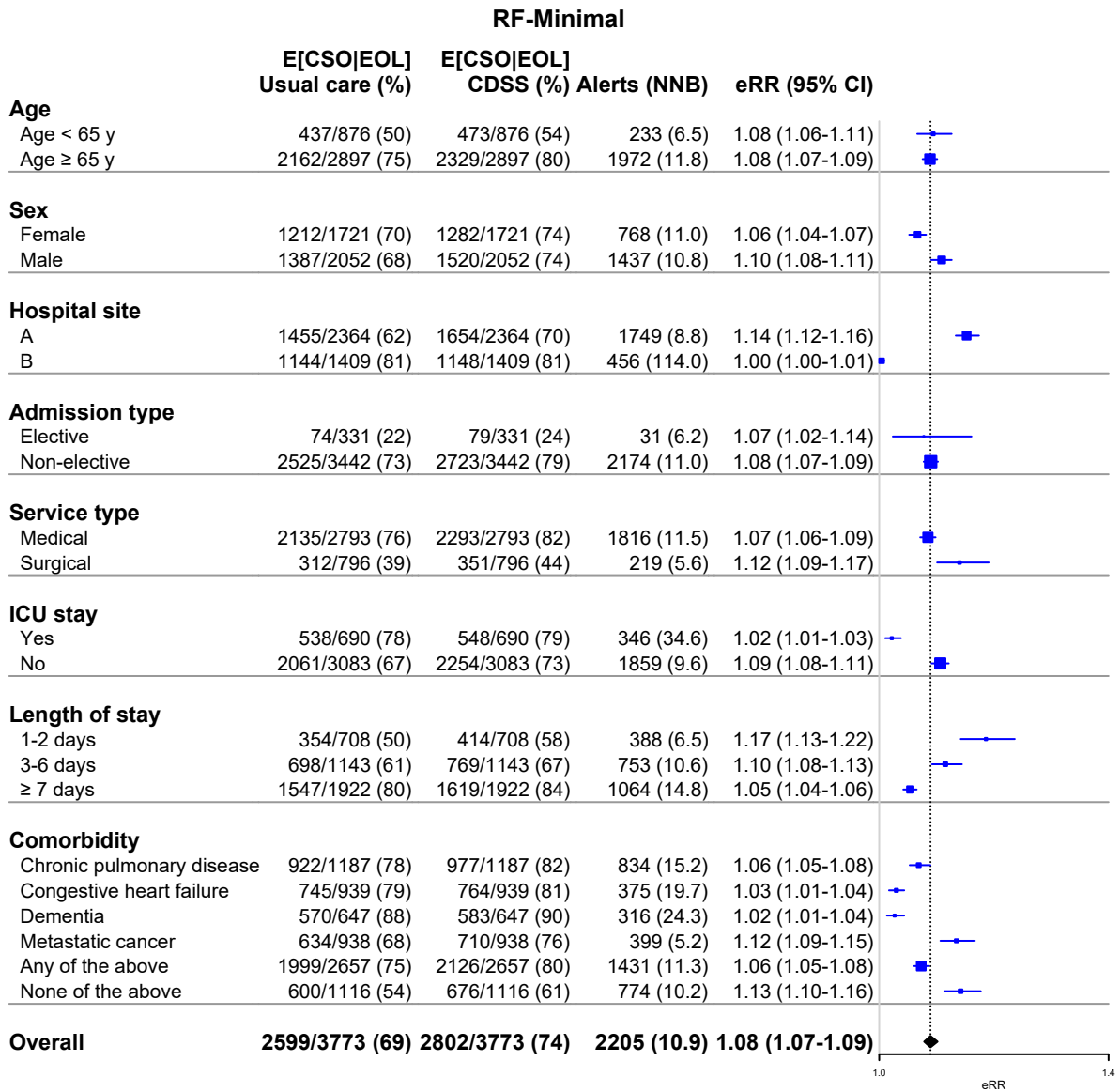
See caption for Figure S2.

Figure S4. Forest plot of clinical utility in relevant subgroups of hospitalized patients (RF-AdminDemo)



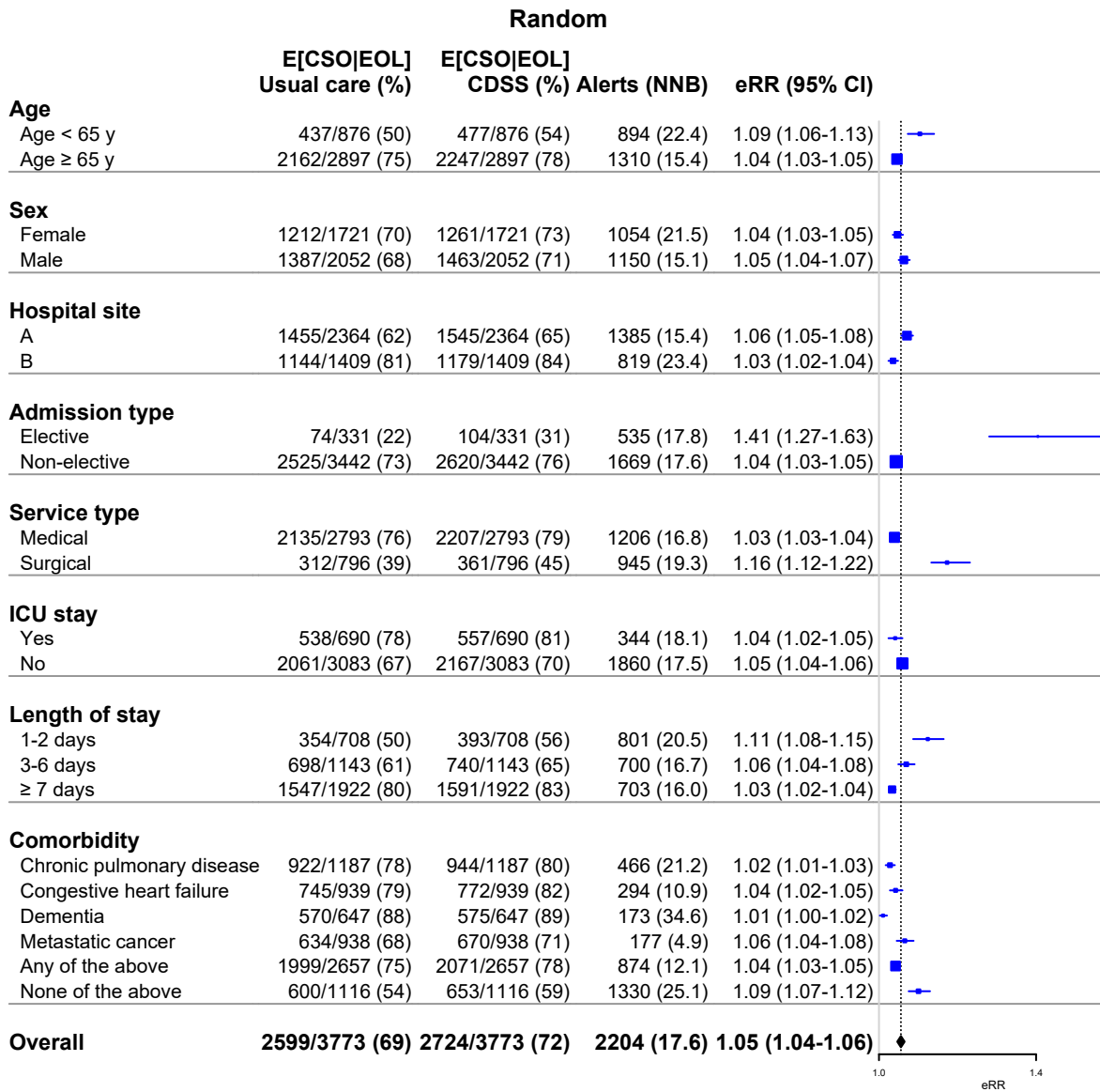
See caption for Figure S2.

Figure S5. Forest plot of clinical utility in relevant subgroups of hospitalized patients (RF-Minimal)



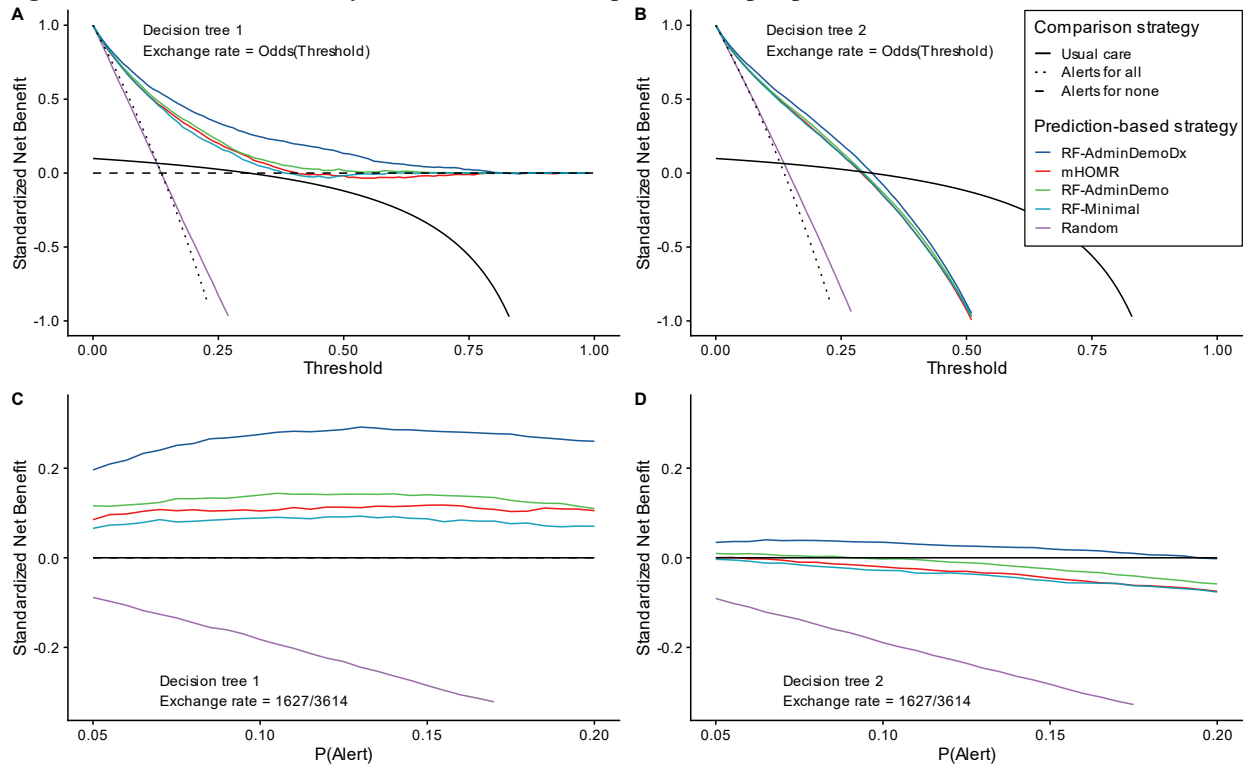
See caption for Figure S2.

Figure S6. Forest plot of clinical utility in relevant subgroups of hospitalized patients (Random)



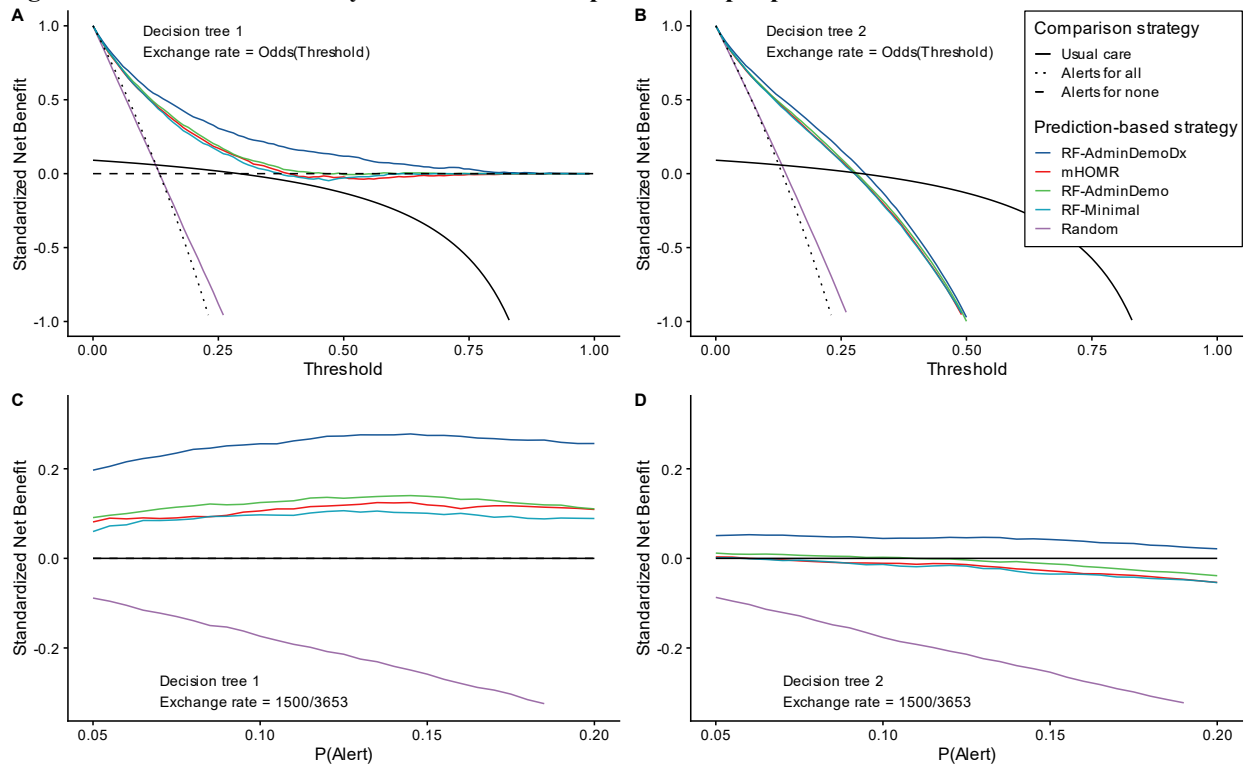
See caption for Figure S2.

Figure S7. Decision curve analysis with a random hospitalization per patient



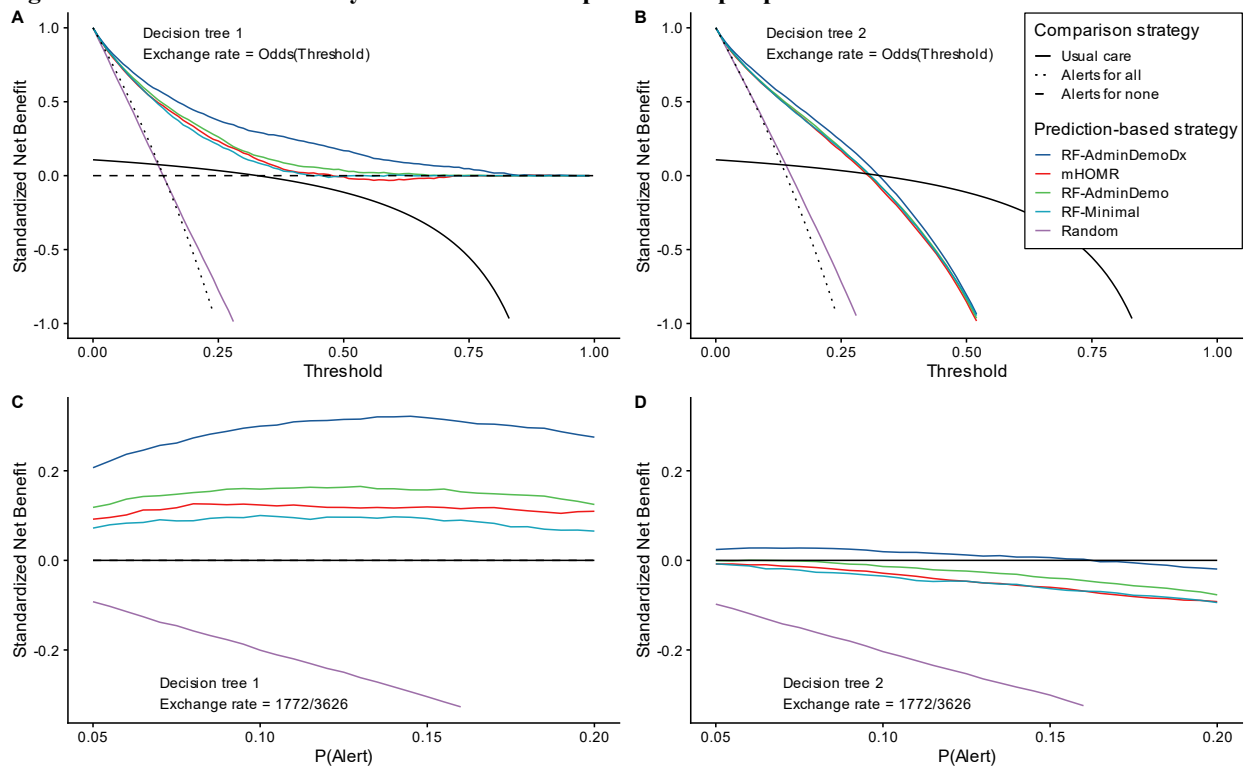
Exchange rate in bottom panels calculated using observed TP and FP rates in included sample (n=16940).

Figure S8. Decision curve analysis with the first hospitalization per patient



Exchange rate in bottom panels calculated using observed TP and FP rates in included sample (n=16940).

Figure S9. Decision curve analysis with the last hospitalization per patient



Exchange rate in bottom panels calculated using observed TP and FP rates in included sample (n=16940).

References

1. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiol Camb Mass*. 2010;21(1):128-138. doi:10.1097/EDE.0b013e3181c30fb2
2. Steyerberg EW. Evaluation of Clinical Usefulness. *Clin Predict Models*. Published online 2019:309-328. doi:10.1007/978-3-030-16399-0_16
3. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med*. 2019;170(1):51-58. doi:10.7326/M18-1376
4. Huber MT, Highland JD, Krishnamoorthi VR, Tang JW-Y. Utilizing the Electronic Health Record to Improve Advance Care Planning: A Systematic Review. *Am J Hosp Palliat Med*. 2018;35(3):532-541. doi:10.1177/1049909117715217
5. Bush RA, Pérez A, Baum T, Etland C, Connelly CD. A systematic review of the use of the electronic health record for patient identification, communication, and clinical support in palliative care. *JAMIA Open*. 2018;1(2):294-303. doi:10.1093/jamiaopen/ooy028
6. Rhodes RL, Kazi S, Xuan L, Amarasingham R, Halm EA. Initial Development of a Computer Algorithm to Identify Patients With Breast and Lung Cancer Having Poor Prognosis in a Safety Net Hospital. *Am J Hosp Palliat Care*. 2016;33(7):678-683. doi:10.1177/1049909115591499
7. Mason B, Boyd K, Murray SA, et al. Developing a computerised search to help UK General Practices identify more patients for palliative care planning: a feasibility study. *BMC Fam Pract*. 2015;16:99. doi:10.1186/s12875-015-0312-z
8. Cox CE, Jones DM, Reagan W, et al. Palliative Care Planner: A Pilot Study to Evaluate Acceptability and Usability of an Electronic Health Records System-integrated, Needs-targeted App Platform. *Ann Am Thorac Soc*. 2018;15(1):59-68. doi:10.1513/AnnalsATS.201706-500OC
9. Ernecoff NC, Wessell KL, Gabriel S, Carey TS, Hanson LC. A Novel Screening Method to Identify Late-Stage Dementia Patients for Palliative Care Research and Practice. *J Pain Symptom Manage*. 2018;55(4):1152-1158.e1. doi:10.1016/j.jpainsymman.2017.12.480
10. Tan A, Durbin M, Chung FR, et al. Design and implementation of a clinical decision support tool for primary palliative Care for Emergency Medicine (PRIM-ER). *BMC Med Inform Decis Mak*. 2020;20(1):13. doi:10.1186/s12911-020-1021-7
11. Hua MS, Ma X, Li G, Wunsch H. Derivation of data-driven triggers for palliative care consultation in critically ill patients. *J Crit Care*. 2018;46:79-83. doi:10.1016/j.jcrc.2018.04.014
12. Makar M, Ghassemi M, Cutler DM, Obermeyer Z. Short-term Mortality Prediction for Elderly Patients Using Medicare Claims Data. *Int J Mach Learn Comput*. 2015;5(3):192-197. doi:10.7763/IJMLC.2015.V5.506
13. Berg GD, Gurley VF. Development and validation of 15-month mortality prediction models: a retrospective observational comparison of machine-learning techniques in a national sample of Medicare recipients. *BMJ Open*. 2019;9(7):e022935. doi:10.1136/bmjopen-2018-022935
14. Cary MP, Zhuang F, Draelos RL, et al. Machine Learning Algorithms to Predict Mortality and Allocate Palliative Care for Older Patients With Hip Fracture. *J Am Med Dir Assoc*. 2021;22(2):291-296. doi:10.1016/j.jamda.2020.09.025

15. Reardon PM, Parimbelli E, Wilk S, et al. Incorporating Laboratory Values Into a Machine Learning Model Improves In-Hospital Mortality Predictions After Rapid Response Team Call. *Crit Care Explor.* 2019;1(7):e0023. doi:10.1097/CCE.0000000000000023
16. Deschepper M, Waegeman W, Vogelaers D, Eeckloo K. Using structured pathology data to predict hospital-wide mortality at admission. *PLOS ONE.* 2020;15(6):e0235117. doi:10.1371/journal.pone.0235117
17. Jain SS, Sarkar IN, Stey PC, Anand RS, Biron DR, Chen ES. Using Demographic Factors and Comorbidities to Develop a Predictive Model for ICU Mortality in Patients with Acute Exacerbation COPD. *AMIA Annu Symp Proc AMIA Symp.* 2018;2018:1319-1328.
18. Cao Y, Bass GA, Ahl R, et al. The statistical importance of P-POSSUM scores for predicting mortality after emergency laparotomy in geriatric patients. *BMC Med Inform Decis Mak.* 2020;20(1):86. doi:10.1186/s12911-020-1100-9
19. Wegier P, Koo E, Ansari S, et al. mHOMR: a feasibility study of an automated system for identifying inpatients having an elevated risk of 1-year mortality. *BMJ Qual Saf.* Published online June 28, 2019:bmjqs-2018-009285. doi:10.1136/bmjqs-2018-009285
20. Courtright KR, Chivers C, Becker M, et al. Electronic Health Record Mortality Prediction Model for Targeted Palliative Care Among Hospitalized Medical Patients: a Pilot Quasi-experimental Study. *J Gen Intern Med.* 2019;34(9):1841-1847. doi:10.1007/s11606-019-05169-2
21. Murphree DH, Wilson PM, Asai SW, et al. Improving the delivery of palliative care through predictive modeling and healthcare informatics. *J Am Med Inform Assoc.* 2021;(ocaa211). doi:10.1093/jamia/ocaa211
22. Major VJ, Aphinyanaphongs Y. Development, implementation, and prospective validation of a model to predict 60-day end-of-life in hospitalized adults upon admission at three sites. *BMC Med Inform Decis Mak.* 2020;20. doi:10.1186/s12911-020-01235-6
23. Manz CR, Chen J, Liu M, et al. Validation of a Machine Learning Algorithm to Predict 180-Day Mortality for Outpatients With Cancer. *JAMA Oncol.* Published online September 24, 2020. doi:10.1001/jamaoncol.2020.4331
24. Uneno Y, Taneishi K, Kanai M, et al. Development and validation of a set of six adaptable prognosis prediction (SAP) models based on time-series real-world big data analysis for patients with cancer receiving chemotherapy: A multicenter case crossover study. *PLoS One.* 2017;12(8):e0183291. doi:10.1371/journal.pone.0183291
25. 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. doi:10.1186/s12911-018-0677-8
26. Sahni N, Simon G, Arora R. Development and Validation of Machine Learning Models for Prediction of 1-Year Mortality Utilizing Electronic Medical Record Data Available at the End of Hospitalization in Multicondition Patients: a Proof-of-Concept Study. *J Gen Intern Med.* 2018;33(6):921-928. doi:10.1007/s11606-018-4316-y
27. Beeksma M, Verberne S, van den Bosch A, Das E, Hendrickx I, Groenewoud S. Predicting life expectancy with a long short-term memory recurrent neural network using electronic medical records. *BMC Med Inform Decis Mak.* 2019;19(1):36. doi:10.1186/s12911-019-0775-2
28. Wang L, Sha L, Lakin JR, et al. Development and Validation of a Deep Learning Algorithm for Mortality Prediction in Selecting Patients With Dementia for Earlier Palliative Care Interventions. *JAMA Netw Open.* 2019;2(7):e196972-e196972. doi:10.1001/jamanetworkopen.2019.6972

29. Lin Y-J, Chen R-J, Tang J-H, et al. Machine-Learning Monitoring System for Predicting Mortality Among Patients With Noncancer End-Stage Liver Disease: Retrospective Study. *JMIR Med Inform.* 2020;8(10):e24305. doi:10.2196/24305
30. Guo A, Foraker R, White P, Chivers C, Courtright K, Moore N. Using electronic health records and claims data to identify high-risk patients likely to benefit from palliative care. *Am J Manag Care.* 2021;27(1):e7-e15. doi:10.37765/ajmc.2021.88578
31. Blom MC, Ashfaq A, Sant'Anna A, Anderson PD, Lingman M. Training machine learning models to predict 30-day mortality in patients discharged from the emergency department: a retrospective, population-based registry study. *BMJ Open.* 2019;9(8):e028015. doi:10.1136/bmjopen-2018-028015
32. Adelson K, Lee DKK, Velji S, et al. Development of Imminent Mortality Predictor for Advanced Cancer (IMPAC), a Tool to Predict Short-Term Mortality in Hospitalized Patients With Advanced Cancer. *J Oncol Pract.* 2018;14(3):e168-e175. doi:10.1200/JOP.2017.023200
33. Parchure P, Joshi H, Dharmarajan K, et al. Development and validation of a machine learning-based prediction model for near-term in-hospital mortality among patients with COVID-19. *BMJ Support Palliat Care.* Published online September 22, 2020. doi:10.1136/bmjspcare-2020-002602
34. Escobar GJ, Ragins A, Scheirer P, Liu V, Robles J, Kipnis P. Nonelective Rehospitalizations and Postdischarge Mortality: Predictive Models Suitable for Use in Real Time. *Med Care.* 2015;53(11):916-923. doi:10.1097/MLR.0000000000000435
35. Gensheimer MF, Henry AS, Wood DJ, et al. Automated Survival Prediction in Metastatic Cancer Patients Using High-Dimensional Electronic Medical Record Data. *JNCI J Natl Cancer Inst.* 2018;111(6):568-574. doi:10.1093/jnci/djy178
36. Parikh RB, Manz C, Chivers C, et al. Machine Learning Approaches to Predict 6-Month Mortality Among Patients With Cancer. *JAMA Netw Open.* 2019;2(10). doi:10.1001/jamanetworkopen.2019.15997
37. Walraven C van. The Hospital-patient One-year Mortality Risk score accurately predicted long-term death risk in hospitalized patients. *J Clin Epidemiol.* 2014;67(9):1025-1034. doi:10.1016/j.jclinepi.2014.05.003
38. Barash Y, Soffer S, Grossman E, et al. Alerting on mortality among patients discharged from the emergency department: a machine learning model. *Postgrad Med J.* Published online December 3, 2020. doi:10.1136/postgradmedj-2020-138899
39. Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014;160(2):122-131. doi:10.7326/M13-1522
40. Blanes-Selva V, Ruiz-García V, Tortajada S, Benedí J-M, Valdivieso B, García-Gómez JM. Design of 1-year mortality forecast at hospital admission: A machine learning approach. *Health Informatics J.* 2021;27(1):1460458220987580. doi:10.1177/1460458220987580
41. Portraits de la population. Accessed January 11, 2021. <https://www.santeestrie.qc.ca/medias-publications/sante-publique/portraits-de-la-population/>
42. Walraven C van, McAlister FA, Bakal JA, Hawken S, Donzé J. External validation of the Hospital-patient One-year Mortality Risk (HOMR) model for predicting death within 1 year after hospital admission. *CMAJ.* 2015;187(10):725-733. doi:10.1503/cmaj.150209
43. Groenwold RHH. Informative missingness in electronic health record systems: the curse of knowing. *Diagn Progn Res.* 2020;4(1):8. doi:10.1186/s41512-020-00077-0

44. Field CA, Welsh AH. Bootstrapping clustered data. *J R Stat Soc Ser B Stat Methodol.* 2007;69(3):369-390. doi:10.1111/j.1467-9868.2007.00593.x
45. Quan H, Li B, Couris CM, et al. Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries. *Am J Epidemiol.* 2011;173(6):676-682. doi:10.1093/aje/kwq433
46. Steyerberg EW. Restrictions on Candidate Predictors. *Clin Predict Models.* Published online 2019:191-206. doi:10.1007/978-3-030-16399-0_10
47. Quan H, Sundararajan V, Halfon P, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. *Med Care.* 2005;43(11):1130-1139.
48. Jameson JL, Fauci A, Kasper D, Hauser S, Longo D, Loscalzo J. *Harrison's Principles of Internal Medicine, Twentieth Edition.* 20th edition. McGraw-Hill Education / Medical; 2018.
49. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *Npj Digit Med.* 2018;1(1):1-10. doi:10.1038/s41746-018-0029-1
50. Hastie T, Tibshirani R, Friedman J. Model Assessment and Selection. In: Hastie T, Tibshirani R, Friedman J, eds. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics. Springer; 2009:219-259. doi:10.1007/978-0-387-84858-7_7
51. Lipton ZC. The Mythos of Model Interpretability. *ArXiv160603490 Cs Stat.* Published online June 10, 2016. Accessed August 26, 2019. <http://arxiv.org/abs/1606.03490>
52. Ahmad MA, Eckert C, Teredesai A, McKelvey G. Interpretable Machine Learning in Healthcare. *IEEE Intell Inform Bull.* 2018;19(1):7.
53. SIGDU - Système d'information de gestion des urgences - Actifs informationnels - Professionnels de la santé - MSSS. Accessed July 12, 2020. <https://www.msss.gouv.qc.ca/professionnels/technologies-de-l-information/actifs-informationnels/sigdu/>
54. MED-ECHO - Sources de données et métadonnées - Professionnels de la santé - MSSS. Accessed July 12, 2020. <https://www.msss.gouv.qc.ca/professionnels/documentation-sources-de-donnees-et-indicateurs/sources-de-donnees-et-metadonnees/med-echo/>
55. Breiman L. Random Forests. *Mach Learn.* 2001;45(1):5-32. doi:10.1023/A:1010933404324
56. Malley JD, Kruppa J, Dasgupta A, Malley KG, Ziegler A. Probability Machines: Consistent Probability Estimation Using Nonparametric Learning Machines. *Methods Inf Med.* 2012;51(1):74-81. doi:10.3414/ME00-01-0052
57. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition.* 2 edition. Springer; 2009.
58. Assel M, Sjoberg DD, Vickers AJ. The Brier score does not evaluate the clinical utility of diagnostic tests or prediction models. *Diagn Progn Res.* 2017;1(1):19. doi:10.1186/s41512-017-0020-3
59. Major VJ, Jethani N, Aphinyanaphongs Y. Estimating real-world performance of a predictive model: a case-study in predicting mortality. *JAMIA Open.* 2020;3(2):243-251. doi:10.1093/jamiaopen/ooaa008
60. DiCiccio TJ, Efron B. Bootstrap confidence intervals. *Stat Sci.* 1996;11(3):189-228. doi:10.1214/ss/1032280214

61. Bernacki RE, Block SD, American College of Physicians High Value Care Task Force. Communication about serious illness care goals: a review and synthesis of best practices. *JAMA Intern Med.* 2014;174(12):1994-2003. doi:10.1001/jamainternmed.2014.5271
62. Allison TA, Sudore RL. Disregard of patients' preferences is a medical error: comment on "Failure to engage hospitalized elderly patients and their families in advance care planning." *JAMA Intern Med.* 2013;173(9):787. doi:10.1001/jamainternmed.2013.203
63. Hunink MGM, Weinstein MC, Wittenberg E, et al. *Decision Making in Health and Medicine: Integrating Evidence and Values.* 2nd ed. Cambridge University Press; 2014. doi:10.1017/CBO9781139506779
64. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. *J Clin Oncol Off J Am Soc Clin Oncol.* 2016;34(21):2534-2540. doi:10.1200/JCO.2015.65.5654
65. Vickers AJ, Calster BV, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ.* 2016;352:i6. doi:10.1136/bmj.i6
66. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA.* 2015;313(4):409-410. doi:10.1001/jama.2015.37
67. Localio AR, Goodman S. Beyond the Usual Prediction Accuracy Metrics: Reporting Results for Clinical Decision Making. *Ann Intern Med.* 2012;157(4):294-295. doi:10.7326/0003-4819-157-4-201208210-00014
68. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak Int J Soc Med Decis Mak.* 2006;26(6):565-574. doi:10.1177/0272989X06295361
69. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol.* 2018;74(6):796-804. doi:10.1016/j.eururo.2018.08.038
70. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res.* 2019;3(1):18. doi:10.1186/s41512-019-0064-7
71. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *J Open Source Softw.* 2019;4(43):1686. doi:10.21105/joss.01686
72. Wright MN, Ziegler A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J Stat Softw.* 2017;77(1):1-17. doi:10.18637/jss.v077.i01
73. Eddelbuettel D, Francois R. Rcpp: Seamless R and C++ Integration. *J Stat Softw.* 2011;40(1):1-18. doi:10.18637/jss.v040.i08
74. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R Package for Causal Mediation Analysis. *J Stat Softw.* 2014;59(1):1-38. doi:10.18637/jss.v059.i05
75. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis.* 2nd ed. Springer International Publishing; 2016. doi:10.1007/978-3-319-24277-4
76. Claus Wilke, Spencer J Fox, Tim Bates, et al. *Wilkelab/Cowplot: 1.1.1.* Zenodo; 2021. doi:10.5281/zenodo.4411966
77. Gordon M, Lumley T. *Forestplot: Advanced Forest Plot Using "grid" Graphics.*; 2020. Accessed January 28, 2021. <https://CRAN.R-project.org/package=forestplot>