Predictive Modeling for Complex Care Management

Georgios Mavroudeas[†], Nafis Neehal[†], Xiao Shou[*], Malik Magdon-Ismail[†],
Jason N. Kuruzovich[‡], Kristin P. Bennett[*†]
[*]Mathematics Department, Rensselaer Polytechnic Institute, Troy, USA
[†]Computer Science Department, Rensselaer Polytechnic Institute, Troy, USA
[‡]Lally School of Management, Rensselaer Polytechnic Institute, Troy, USA

2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 978-1-6654-0126-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/BIBM52615.2021.9669714

***Abstract*—Complex care management (CCM) or hot spotting programs identify and manage high-need/high-cost patients, improving long-term health quality and medical costs. Typically, physicians refer patients to CCM. Despite strict guidelines to ensure that eligible patients are placed in appropriate programs, such a provider-based approach is limited by provider-capacity and the narrow view of a patient that a provider sees. We propose an ML workflow to augment the provider-based approach, that can flag patients who are suited to CCM. Our predictor uses a global view of a patient's entire history across multiple providers and time to identify high-risk individuals from among all the individuals in a matter of seconds. On a monthly basis, we evaluate our predictions against physician referrals. In the test dataset, 41% of the top-500 highest risk individuals found by our model were referred to CCM by a physician at some point in the 6-month window following our prediction (top-500 is a parameter that can be set to match the CCM program's capacity). Of those who were not referred in the 6-month window, 30% were referred at some time in their trajectory. The remaining false positives had a greater than 95% similarity when compared to true positive physician referrals in terms of cost profiles (both prior to referral and after referral) and patient profile. This remarkable similarity suggests that our machine learning predictor can identify new candidates for complex care management and/or predict referrals before a physician has an opportunity to do so.**

I. Introduction

We propose a machine learning (ML) approach for referring patients to complex care management (CCM) and test our approach at a health plan provider in a small metropolitan area. Health care costs in the US are highly concentrated, with 5 percent of patients consuming over 50 percent of resources [1]. One way to improve patient health and address this concentration of expenses is "hot spotting" to identify the highest-needs/highest-cost patients for programs like CCM [2]–[4]. While the benefits of hot spotting and interventions like CCM can be mixed [2]–[4], the sheer magnitude of costs and complexity involved in serving the sickest patients motivates health insurance companies to improve outcomes while controlling spending. Predicting the at-risk patients is a challenging but key enabler of programs such as CCM [5].

Our approach is to predict CCM referral using prior physician referral-patterns and claims data provided by the insurance provider. The goal is to predict the high-needs/high-cost individuals that would benefit from CCM, in particular those not already identified at a point of care. Finding such patients earlier in their health trajectory can significantly improve health outcomes and costs.

A. Decision Support for Complex Care Management

Either the physician or the health plan can enroll a patient in CCM. Even though experts are responsible for CCM-referral, qualified individuals can be missed: the physician/provider network has limited capacity; physicians can only refer patients they see; a single physician may not see the entire trajectory of a patient - a more global view of the patient may be needed to identify them as in need of CCM. Our goal is to augment physician-based CCM referral with ML based predictive referrals. Our use-case is a non-profit insurance company. Our approach is to learn from historical data on physician referrals. This data poses challenges:

1) Non-stationarity: The healthcare landscape and physician referral process are evolving.
2) Unbalanced data: Only 0.5% of people get referred.
3) Sparse data: The majority of patients (including referrals) have no or very few events in their historical profile.
4) Incompletely labeled data: Physicians may only refer the sickest patients and only when they see them. Several eligible individuals may not get referred to CCM.

The general task is time-series event prediction. For a review of event prediction algorithms, see [6]. Before predicting, one must model the problem, and little work has addressed modeling and evaluation. For CCM, we found that several nuanced modeling choices to address specific challenges posed by the CCM training data had big impact on performance across all algorithmic approaches. We present effective ways to deal with sparsity. We find that both **windowing** and **down-sampling** of non-referral data have significant impact on performance (see Section IV). Our methods have implications for CCM and hot spotting, and more generally time series event prediction when the event is rare.

II. Formulation of the CCM Prediction Problem

Historical data on referrals provided by the domain experts is central in our methods. Let P (for physician) denote the existing referral process into CCM. The input to P is a patient's profile up to the time the decision is being made. The profile is $x \in X$. One main design choice is how to construct a patient's profile x (see Section III and IV). The process P either labels the profile x as a referral to CCM, in which case x is labeled $y = 1$, or not, in which case x is unlabeled. Non-referred patients are unlabelled, since they are not specifically labeled as non-referrals by a domain expert.

$$P : X \rightarrow \{\text{refer, no label}\}$$

One of the design choices in the ML workflow is how to deal with the unlabeled data. The simplest approach treat unlabeled data as non-referrals, $y = 0$. Alternatively, one can *relabel* some unlabeled data if they were referred at a future time.

We obtain training data $(x_1, y_1), \ldots, (x_N, y_N)$, where N is the number of training data points, x_i are patient profiles, and $y_i = 1$ if the patient was referred at that time or relabeled as a referral, and $y_i = 0$ otherwise. The task is to learn a predictor $\hat{P}$ that imitates the physician process P, possibly referring additional non-referred CCM-patients or referring referred patients, but perhaps earlier. We emphasize that our goal here is not to produce the optimal candidates for CCM, which would require data on who would benefit from CCM. We are simply learning to mimic physician referral.

There challenges we address are: heavy class imbalance (0.5% positive labels); partial labeling of just the minority positive class; non-stationarity; and, sparse patient profiles. Further, the target process P is not one process but the sum of all the provider activity. This results in significant inhomogeneity and differing referral practices among physicians and for different diagnoses, leading to bias in the labels.

III. Data

Proprietary de-identified CCM data was provided by an insurance company. The data is primarily derived from claims but also has features from electronic medical records. The data has two main parts: (1) the part used for obtaining the patient profiles to construct the feature vector x for each patient on each month, and (2) the part containing physician referrals including the reasons for the referral.

Data Description. The data has about 22.5M patient records for about 700K patients over 56 months from Oct. 2015 to Jun. 2020. Each record is patient-month containing the patient profile on that month: 69 diagnosis codes, 14 expenditure categories, and demographic data such as gender and age. The expenditures include total cost, prescription/pharmacy, and in-patient care. The referral data contains information on physician referrals to CCM at the patient-month level. The label at each patient-month (referred or unlabeled) is obtained from the referral records. We have 78,492 referral records for 42,050 patients, giving referral date and referral reason.

Training and Test Data. We use Jan. 2017 to Dec. 2018 to create training data $(x_i(t), y_i(t))$ of profile-label pairs for every month t. Every patient could in principle contribute 24 training examples. The resulting training dataset contains approximately 6.9 million patient records for 374,106 patients. We had 13,383 training referral records from 11,219 patients for the 2017-2018 time period.

We evaluate the learned predictor on 12 months after the training period, Jan. 2019 to Dec. 2019. As with training data, we extract profile-label pairs for every patient-month in the test period, giving about 4.5M test examples on about 430K patients with about 10K referral records on 8,303 patients.

Features. As already mentioned, we have 87 raw features of 3 types: diagnosis codes and disease counts, (e.g., diabetes); expenditures in various categories, (e.g., in-patient, pharmacy); demographics, (e.g., age and gender). Categorical variables were mapped to a discrete numeric values. The majority of such variables are diseases diagnoses which are present or absent, and hence map to binary variables. As it is common practice in healthcare, disease diagnosis codes are "sticky" and persist for at least 12 months once given. The stickiness has been determined by the domain experts at the health plan. Despite this persistence in features, the data is still sparse. We give a summary of the data statistics in table I

IV. Methods

Our methodology has three stages.

1) Pre-processing for: a) Converting time-series data to supervised data. b) Feature engineering, including centering, scaling and possibly reducing dimension. c) Relabeling data to augment the physician referrals, and possibly down-sampling to address class imbalance.
2) Predictive Modeling to score each patient in a month. The top-500 scores are the model positives. We used many ML methods to model the physician process P.
3) A suite of tools to evaluate model predictions on test data, assessing efficacy of rare event prediction in time series.

A. Pre-Processing

We address several challenges in the temporal data. First, it is relatively straightforward to convert the time series data into standard supervised data by treating each patient-month as a distinct training example. Each patient profile gets its training label from the referral data.

Next we address the sparsity of positive labels by relabeling unlabeled examples using the fact that a patient's state is sticky. Hence, we relabel a month as a physician referral (positive data point) if a physician referred the patient at some point in the forward 6-month window (6-months is a parameter, which for simplicity we keep fixed in this study). This propagates physician referrals 6-months into the past.

To address sparsity of the feature matrix which has only about 5.5% non-zero entries, we used PCA. which densifies the feature-matrix. A side benefit is we are able to significantly lower the feature-dimension without much loss in information.

Lastly, to address the extreme class imbalance under-sample negative examples. For every positive example, we randomly choose a corresponding negative example, producing a training set with the same number of positive and negative examples.

B. Modeling

We deployed several standard ML methods to learn the physician process P. These methods are available in Python's

TABLE I
DATA SUMMARY

	Time period	Months	Patient Records	Patients	Referrals	Data Matrix	Sparsity
Train	01/2017–12/2018	24	6.85M	374,106	13,383	6.85M × 87	5.46% non-zero
Test	01/2019–12/2019	12	4.50M	426,954	9,765	4.50M × 87	5.56% non-zero

scikit-learn library [7]: (i) Logistic regression with l_2 regularization (ii) Naïve Bayes classifier (iii) Gradient boosting classifier (iv) Neural networks. To select hyper-parameters, we used cross validation. Neural Networks struggled with the sparse data and consistently underperformed in accuracy and runtime. Hence, our results focus on the three other classifiers.

C. Evaluation

The trained models generate monthly predictions moving forward, for each patient-month in the test set. We evaluated the predictions using confusion matrices, precision, recall, accuracy, F1 score, balanced accuracy score, and ROC-AUC scores. To align with the priorities of managed healthcare, the metric we focused on is top-k precision, that is the accuracy of the ML's positive predictions. The reason is that physician referrals cannot be undone, that is, false negatives are irrelevant. The potential benefits come from identifying new candidates for CCM. Specifically, do the model's false positives have quantifiable value? We perform in-depth analysis of the model false positives to provide evidence that these false positives are good candidates for CCM.

Figure 1 shows how we compute precision. For each patient-month, the model scores patients and the top k are model positives, for $k = 500$ (to match the available capacity at the insurance company). All other patients in the month are model negatives. We consider three possibilities: (i) A patient in the top k has a referral in the next g months, a true positive (TP). (ii) A patient in the top k does not get a referral in the next g months, a false positive (FP) (iii) A patient is referred but has no model positive in the prior g months, a false negative (FN). The referral time horizon g is the evaluation window.

V. RESULTS

We now evaluate the machine learning workflow against physician referrals, interpret the final predictor, and compare the design choices available. Our evidence shows:

1) Predictive performance depends significantly on the machine learning model. Robust boosted models worked best. The top performer was gradient-boosted trees.
2) Preprocessing (scaling, down-sampling, dimension reduction (PCA) and relabeling to enforce referral-stickiness) can significantly impact on predictive performance. However, for gradient-boosted trees which are automatically regularized, the impact was minimal on accuracy. However, pre-processing significantly improved the efficiency of our top performing model without loss in accuracy. This often happens with high-dimensional data.
3) The model positives are corroborated by the physician referrals and in-depth analysis of the patient profiles. Our model can augment the provider based referral system by identifying additional candidates for CCM.
   - Both physician and model refer the sickest patients.
   - The model false positives appear to deteriorate in health, an argument in favor of placing them in CCM.
   - Model false positives have patient profiles that are remarkably similar to physician referrals.

A. Machine Learning Model, Pre-processing and Relabeling

Tables II and III compare three representative machine learning models: Naive Bayes, regularized logistic regression, and our top performer, gradient-boosted trees. These models are available in Python's scikit-learn package [7].

We compare models with and without downsampling, top-5 PCA, and relabeling. For our top model, to benefit from the drastic efficiency gains of PCA and downsampling, relabeling is important and gives a 4% boost in precision. The relabeling compensates for the slight information loss with PCA and downsampling. The efficiency gain from downsampling (close to 200×) is due to the extreme class imbalance. Note that when relabeling, to enforce stickiness of referrals, we used a 6-month window to match the 6-month evaluation window. We did different relabeling windows. There is a general near-monotonic trend with more relabeling leading to better performance. We focus on 6-month relabeling.

TABLE II
EFFECT OF PCA (12-MONTH TEST PRECISION).

Model	**PCA**		**No PCA**	
	No relabeling	Relabeling	No relabeling	Relabeling
Gradient Boosting	**0.371**	**0.411**	**0.401**	**0.419**
Logistic Regression	0.164	0.159	0.288	0.298
Naïve Bayes	0.103	0.104	0.111	0.119

TABLE III
EFFECT OF DOWNSAMPLING (12-MONTH TEST PRECISION).

Model	**Downsample**		**No Downsample**	
	No Relabeling	Relabeling	No Relabeling	Relabeling
Gradient Boosting	**0.371**	**0.411**	**0.410**	**0.418**
Logistic Regression	0.164	0.159	0.166	0.157
Naïve Bayes	0.103	0.104	0.099	0.097

B. Interpreting the Model: Feature Analysis

We analyze the important features for our top-performing gradient-boosted predictor. Such feature analysis gives insight

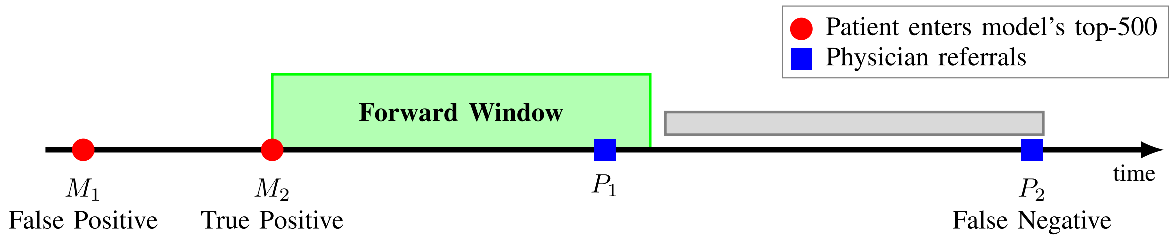

Fig. 1. Evaluation Metric. The patient enters the top-500 for model referral at M_1 and M_2. Physician referrals occur at P_1 and P_2. M_1 is a false positive because there is no physician referral in M_1's forward window. M_2 is a true positive because there is a physician referral in M_2's forward window. The physician referral P_2 is not in the forward window of any prediction model prediction in the gray box, and hence is a false negative.

TABLE IV
TOP 10 IMPORTANT FEATURES FOR PREDICTION ON CCM DATASET

Rank	Feature Name	Category	Type	Importance
1	Total Cost	Cost	Dollar	0.532
2	Chronic Disease Count	Dignosis	Number	0.218
3	Age at Midmonth	Demographic	Float	0.049
4	Cost for Radiology	Cost	Dollar	0.042
5	Tobacco Use	Diagnosis	Boolean	0.036
6	Cost for Primary Care	Cost	Dollar	0.021
7	Cost for Outpatient	Cost	Dollar	0.021
8	Cost for Inpatient	Cost	Dollar	0.019
9	Depression	Diagnosis	Boolean	0.012
10	Bipolar Disorder	Diagnosis	Boolean	0.010

into the model, identifying the parts of a patients' profile most indicative of CCM-eligibility. The top-10 features are in Table IV. The importance roughly captures how often the feature is a determining factor in the final decision [7]. From a healthcare perspective, the model has honed in on sensible features. For example the top three features are the historical total expenditures, a proxy for a patients' health (higher cost means more sick), number of chronic diseases and age.

C. Comparing Model Predictions to Physician Referrals

We now compare the details of the machine learning predictions with the physician referrals. We analyze: True positives (TP), the model's top-500 riskiest that were also physician referrals; False positives (FP), the model's top-500 riskiest that didn't get a physician referral; False negatives (FN), physician referals that didn't make into model's top-500 riskiest duing the prior evaluation window. We already saw that the precision of the model-positives is a staggering 41%. We now delve into more details on the false positives and false negatives.

On average we find that around **30%** of our FP individuals were referred to CCM at some point in their medical trajectory (either before or after the model captured them). So in a sense, these are only mild false positives. We also examine the FP and TP subgroups using three additional approaches: patient similarity to physician referrals, referral reasons, and cost/disease similarity to physician referrals. These experimental results corroborate that the FP and TP subgroups, which our ML models predict every month, are highly similar to one other with respect to their corresponding patient profiles. We conclude that in each month, most patients belonging to the FP subgroup would have been appropriate for referral to CCM.

Patient Similarity. In table V, we present statistics that yield important insights on how much similarity there is on average between the patients from TP and FP subgroups. We examine the closest member in the TP group for each member of the FP group. The closest patient is defined using Cosine Similarity [8]. To quantify the significance of the similarity between FP and TP, in each month, we sample a random cohort of 5,000 patients and compute the fraction of this random cohort whose similarity to its closest TP is below the similarity of the FP cohort. The results are in Table V. We find that on average, there is more than **95%** similarity between the FP and TP cohort and the significance is on average larger than 85% of random patients.

Referral Reasons. Tables VI and VII provide insights on the distribution of the referral reasons among the people referred by our model. Specifically table VI, showcases the distribution of referral reasons among the TP and FN groups. Table VII shows the distribution of the top 10 referral reasons by rate in our dataset, and provides the FP and TP rates within each group. This suggests that our model captures a large portion of the most prominent referral reasons, obtaining good precision on these referral codes. A notable exception, high risk pregnancy, is not well modeled because prenatal costs tend to be fixed costs that are uncorrelated with severity.

D. Cost and Chronic Disease Count Similarity

Lastly, we analyze the two important features cost and disease count in the 12 months prior to and after referral. We compare TP, FP, FN and the rest of the patients not referred by our model or physicians. Figure 2 shows the results. In graphs 2(a-c), we find an astounding similarity between the FP and TP groups suggesting that the group corresponding to FP might include patients who should be in CCM. The group indicated by the "REST" included the bulk of the patients and maintains a low cost and disease profile. The FN group, even though it is higher in ranking than the "REST" group, is substantially lower than the TP and FP groups. Figure 2(c) shows that the patients that are on poor chronic health trajectories are found by both our model and the physicians.

TABLE V
AVERAGE COSINE SIMILARITY BETWEEN FALSE POSITIVES AND CLOSEST TRUE POSITIVE. WE ALSO SHOW THE FRACTION OF RANDOM PATIENTS WITH LOWER COSINE SIMILARITY THAN FALSE POSITIVES. OUR FALSE POSITIVES ARE MUCH MORE SIMILAR TO TRUE POSITIVES THAN RANDOM PATIENTS.

MYR	201901	201902	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912
Avg. Cos Similarity	0.976	0.966	0.978	0.977	0.973	0.968	0.979	0.974	0.963	0.975	0.970	0.961
Avg. Cos Percentile	87.463	83.756	86.400	75.716	88.287	86.347	80.709	84.994	82.322	83.545	84.777	83.747

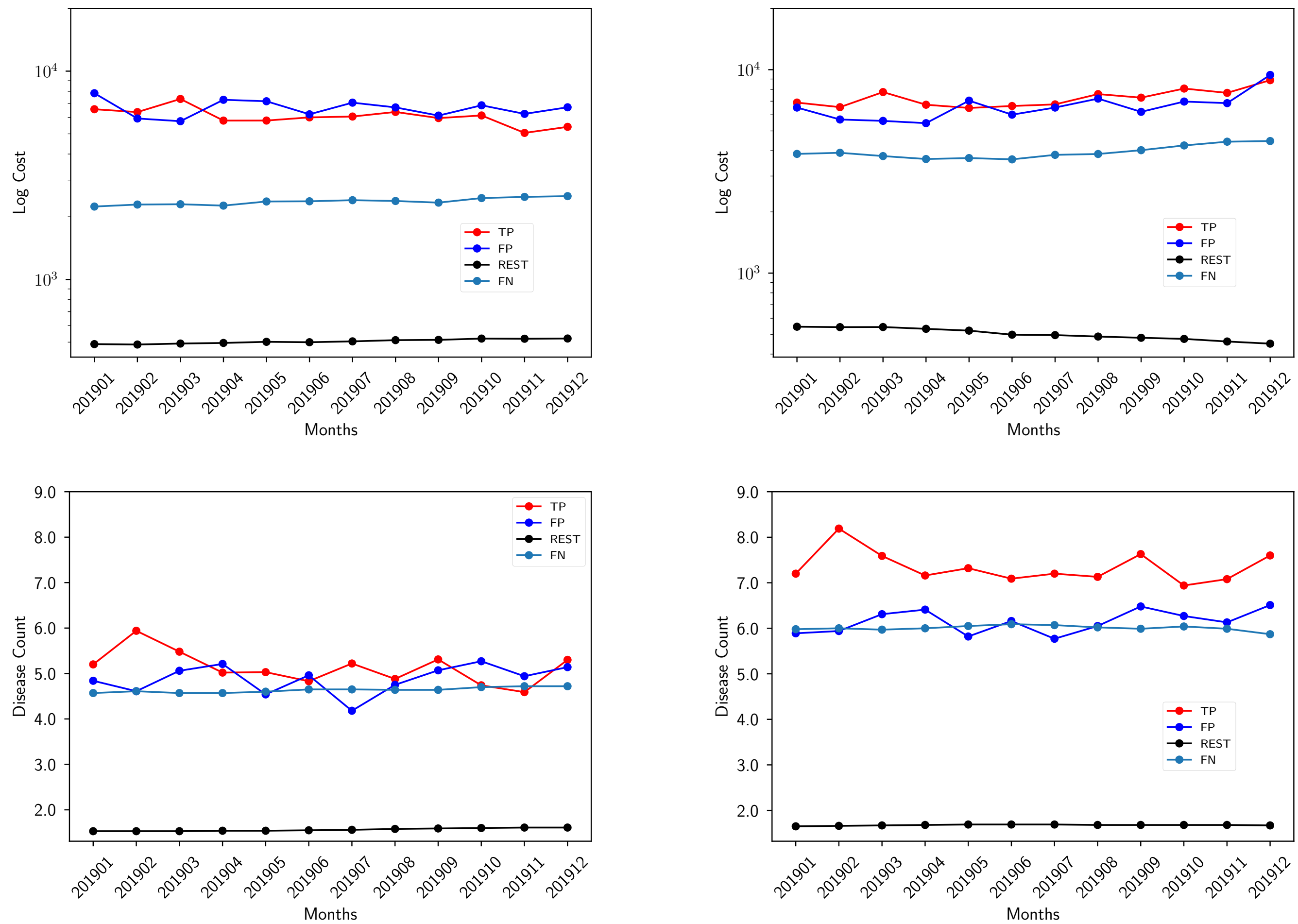

Fig. 2. Cost and Chronic Disease Counts prior to and after model referral. Average patient profile for: (Left) the year prior to the referral month, the past. (Right) the year after the referral month, the future. Prior profile of true and false positives are comparable. The future disease profile is more severe for true positives: patients identified by both the model and physicians are the sickest. The false positives and false negatives have comparable disease profiles with false positives having higher future cost. The false positives appear sicker than other other physician-referrals (false negatives), and might benefit from CCM.

VI. CONCLUSION AND FUTURE WORK

A key contribution of this work is to demonstrate the potential of an ML model trained with physician referrals to identify similarly complex patients, offering the potential of scaling the insights of doctors systematically across a population. We developed a pipeline to effectively model CCM using sparse, unbalanced, and inaccurately labeled data, and demonstrated it in a case study. Our results showed that the gradient boosting model has a high degree of precision (41%) across the subset of the top 500 riskiest patients, an evaluation directly aligned with the needs of the health plan provider's use case. To summarize our results, machine learning is able to identify a large fraction of physician referrals. Indeed these patients appear to be the sickest patients whose numbers of chronic disease counts drastically increases in the year following referral. The machine learning also identifies candidates for CCM which are not picked up by the current CCM-referral process. In a detailed analysis of these "false positives", their remarkable similarity to actual physician referrals suggests that our machine learning predictor can identify new candidates for complex care management and/or predict referrals before a physician has an opportunity to do so.

TABLE VI
REFERRAL REASON FOR TRUE POSITIVES AND FALSE POSITIVES (% OF MODEL PREDICTIONS IN A REASON CODE). THE MODEL'S REASON-CODE DISTRIBUTION IS SIMILAR TO THE PHYSICAN DISTRIBUTION IN TABLE VII

False Negatives	**Perc%**	**True Positives**	**Perc%**
Other	31.44	Other	22.1
Behavioral Health	17.25	Behavioral Health	33.44
Coordination of Care	19.4	Coordination of Care	14.81
Substance Use	5.0	Substance Use	9.11
Mental Health Condition	4.67	Mental Health Condition	9.1
High Risk Pregnancy	7.04	Diabetes	2.02
Diabetes	3.99	Oncology	1.57
Oncology	1.63	Psychosocial Concerns	1.14
Psychosocial Concerns	1.73	COPD	0.91
COPD	1.2	Mul Co-Morb or Cplx Cond	1.08

TABLE VII
REASONS FOR PHYSICIAN REFERRALS (FRACTION OF PHYSICIAN REFERRALS IN A REASON-CODE THAT WERE CAPTURED IN THE MODEL'S TRUE POSITIVES). ACCURACY VARIES OVER REASON CODE.

Referral Reasons	**True Positives%**	**False Negatives%**	**Rate%**
Other	17	83	29
Behavioral Health	37	63	21
Coordination of Care	19	81	18
Substance Use	35	65	6
Mental Health Condition	37	63	6
High Risk Pregnancy	3	97	6
Diabetes	13	87	4
Oncology	23	77	2
Psychosocial Concerns	17	83	2
COPD	18	82	1

Despite a greatly unbalanced dataset, we are able to demonstrate effective ways of improving the modeling process through both windowing and downsampling. Specifically, windowing the data seems to marginally improve modeling outcomes. Further, we show that downsampling can be used to make the training process dramatically more efficient with little impact on model performance.

The maini limitation of our research is that our goal was to mimic physicians based on claims data alone so the system is designed for usage by health payers not providers. Additional patients not found by physicians could benefit from CCM and electronic health records data could provide additional useful features. Further, predicting based on physician-referral patterns could propagate biases and disparities in existing CCM referrals. A system which could incorporate additional input from domain experts, e.g. clinicians, and also provide explanations for its predictions could enhance user confidence in the system and increase usage.

Another natural extension is further study to understand the degree to which CCM may be effective in improving patient outcomes. Randomized clinical trials represent the gold standard in such evaluations. For example a recent randomized controlled trial found a hot-spotting program had an insignificant effect in hospital readmissions [3]. This analysis could be extended to an observational study of CCM effectiveness for a variety of health and usage based outcomes. Since CCM-based interventions may actually increase usage and costs in the near term, the long-term effects of CCM on both patient health-outcomes and costs require more study. Our proposed CCM pipeline creates propensity-to-treat models which could be used in propensity-based treatment effect analysis. We leave these extensions and enhancements to future work.

ACKNOWLEDGEMENT

Supported by a grant from Capital District's Physicians' Health Plan. We thank M. Seridi many useful discussions.

REFERENCES

[1] S. B. Cohen, "Statistical brief 455 the concentration of health care expenditures and related expenses for costly medical conditions, 2012," *Medical Expenditure Panel Survey: Agency for Healthcare Research and Quality (US)*, 2014.

[2] L. M. Marcotte, A. Reddy, and J. Liao, "Addressing avoidable healthcare costs: Time to cool off on hotspotting in primary care?" *J. General Internal Medicine*, vol. 34, no. 11, pp. 2634–2636, 2019.

[3] A. Finkelstein, A. Zhou, S. Taubman, and J. Doyle, "Health care hotspotting—a randomized, controlled trial," *NEJM*, vol. 382, no. 2, pp. 152–162, 2020.

[4] D. M. Zulman, C. P. Chee, S. C. Ezeji-Okoye *et al.*, "Effect of an intensive outpatient program to augment primary care for high-need veterans affairs patients: a randomized clinical trial," *JAMA internal medicine*, vol. 177, no. 2, pp. 166–175, 2017.

[5] J. M. McWilliams and A. L. Schwartz, "Focusing on high-cost patients: the key to addressing high costs," *NEJM*, vol. 376, no. 9, p. 807, 2017.

[6] S. M. Molaei and M. R. Keyvanpour, "An analytical review for event prediction system on time series," in *IEEE IPRIA*, 2015, pp. 1–6.

[7] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel *et al.*, "Scikit-learn: Machine learning in Python," *JMLR*, vol. 12, pp. 2825–2830, 2011.

[8] C. D. Manning, P. Raghavan, and H. Schutze, "Introduction to information retrieval," *CUP*, vol. 20, pp. 405–416, 2008.

[9] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proc. ICML*, 2009, pp. 41–48.

[10] K. J. Bowers, S. D. Johnson, and K. Pease, "Prospective hot-spotting: the future of crime mapping?" *British Journal of Criminology*, vol. 44, no. 5, pp. 641–658, 2004.

[11] S. D. Johnson and K. J. Bowers, "The burglary as clue to the future: The beginnings of prospective hot-spotting," *European J. of Criminology*, vol. 1, no. 2, pp. 237–255, 2004.

[12] H. Mattie, P. Reidy, P. Bachtiger, E. Lindemer, N. Nikolaev, M. Jouni, J. Schaefer, M. Sherman, and T. Panch, "A framework for predicting impactability of digital care management using machine learning methods," *Population health management*, vol. 23, no. 4, pp. 319–325, 2020.

[13] Z. Obermeyer and E. J. Emanuel, "Predicting the future—big data, machine learning, and clinical medicine," *NEJM*, vol. 375, no. 13, p. 1216, 2016.

[14] A. L. Beam and I. S. Kohane, "Big data and machine learning in health care," *JAMA*, vol. 319, no. 13, pp. 1317–1318, 2018.

[15] T. Panch, P. Szolovits, and R. Atun, "Artificial intelligence, machine learning and health systems," *J. Global Health*, vol. 8, no. 2, 2018.

[16] M. Diwakar, A. Tripathi, K. Joshi, M. Memoria, P. Singh *et al.*, "Latest trends on heart disease prediction using machine learning and image fusion," *Materials Today: Proceedings*, vol. 37, pp. 3213–3218, 2021.

[17] U. Evci, F. Pedregosa, A. Gomez, and E. Elsen, "The difficulty of training sparse neural networks," *arXiv preprint:1906.10732*, 2019.

[18] T. Gale, E. Elsen, and S. Hooker, "The state of sparsity in deep neural networks," *arXiv preprint:1902.09574*, 2019.