Using a Constraint-Based Method to Identify Chronic Disease Patients Who are Apt to Obtain Care Mostly within a Given Health Care System: Retrospective Cohort Study

Yao Tong[1], MS; **Zachary C. Liao**[1], MD, MPH; **Peter Tarczy-Hornoch**[1,2,3], MD; **Gang Luo**[1], PhD
[1]Department of Biomedical Informatics and Medical Education, University of Washington, UW Medicine South Lake Union, 850 Republican Street, Building C, Box 358047, Seattle, WA 98195, USA
[2]Department of Pediatrics, Division of Neonatology, University of Washington, 1959 NE Pacific St, Seattle, WA 98195, USA
[3]Department of Computer Science and Engineering, University of Washington, 185 Stevens Way, Seattle, WA 98195, USA
yaotong@uw.edu, zachcl@uw.edu, pth@uw.edu, luogang@uw.edu

Corresponding author:
Gang Luo, PhD
Department of Biomedical Informatics and Medical Education, University of Washington, UW Medicine South Lake Union, 850 Republican Street, Building C, Box 358047, Seattle, WA 98195, USA
Phone: 1-206-221-4596
Fax: 1-206-221-2671
Email: luogang@uw.edu

Abstract

Background: For several major chronic diseases including asthma, chronic obstructive pulmonary disease (COPD), chronic kidney disease (CKD), and diabetes, a state-of-the-art way to avert poor outcomes is to use predictive models to identify future high-cost patients for preemptive care management interventions. Frequently, an American patient obtains care from multiple healthcare systems, each managed by a distinct institution. As the patient's medical data are spread across these healthcare systems, none of them has complete medical data for the patient. The task of building models to predict an individual patient's cost is currently thought to be impractical on incomplete data, limiting the use of care management to improve outcomes. Recently, we developed a constraint-based method to pinpoint patients apt to obtain care mostly within a given healthcare system. Our method was shown to work well for the cohort of all adult patients at the University of Washington Medicine (UWM) for a 6-month follow-up period. It is unknown how our method performs on patients with various chronic diseases and over follow-up periods of different lengths, and subsequently whether it is reasonable to perform this predictive modeling task on the subset of patients pinpointed by our method.

Objective: To understand our method's potential to enable this predictive modeling task on incomplete medical data, this study assesses our method's performance at the UWM on 5 subgroups of adult patients with major chronic diseases and over follow-up periods of 2 different lengths.

Methods: We used UWM data for all adult patients who obtained care at the UWM in 2018 and PreManage data containing usage information of all hospitals in Washington state in 2019. We evaluated our method's performance over the follow-up periods of 6 months and 12 months on 5 patient subgroups separately, one subgroup for each of 5 diseases: asthma, CKD, type 1 diabetes (T1D), type 2 diabetes (T2D), and COPD.

Results: Our method identified 21.81% (3,194/14,644) of UWM adult patients with asthma. About 66.75% (797/1,194) and 67.13% (1,997/2,975) of their emergency department visits and inpatient stays took place within the UWM in the subsequent 6 months and in the subsequent 12 months, respectively, roughly double the corresponding percentage for all UWM adult patients with asthma. The performance for adult patients with CKD, adult patients with COPD, adult patients with T1D, and adult patients with T2D was reasonably similar to that for adult patients with asthma.

Conclusions: For each of the 5 chronic diseases most relevant to care management, our method can pinpoint a reasonably large subset of patients apt to obtain care mostly within the UWM. This opens the door to building models to predict an individual patient's cost on incomplete data, which was formerly deemed impractical.

International Registered Report Identifier (IRRID): PRR2-10.2196/13783

Keywords: asthma; chronic kidney disease; chronic obstructive pulmonary disease; data analysis; diabetes mellitus; emergency department; health care system; inpatients; patient care management

Introduction

Background

Care management is widely used to improve the outcomes of patients with chronic diseases [1]. Typically, we build a model to predict an individual patient's cost [1-5]. For a patient predicted to incur high cost in the future, we enroll the patient in a care management program for preemptive interventions. Then a care manager will call the patient regularly to check the patient's status and help arrange health and related services. Proper use of care management can cut cost by up to 15%, reduce hospital visits (emergency department visits and inpatient stays) by up to 40%, and bring many other benefits [4,6-13]. Care management is typically used for managing the chronic diseases of asthma, chronic obstructive pulmonary disease (COPD), chronic kidney disease (CKD), and diabetes, as these diseases fulfill 3 conditions making a care management program economically feasible for implementation: 1) the disease has a high prevalence rate; 2) if not treated appropriately, the disease can result in expensive acute exacerbations; and 3) relatively low-cost and effective interventions within the patient's control are available for the disease [6,14].

In the United States, a patient often obtains care from several healthcare systems such as academic medical centers and private physician groups. As the patient's medical data are spread across these healthcare systems, none of them has complete medical data for the patient. Our prior work showed that <1/3 of hospital visits by University of Washington Medicine (UWM) adult patients took place within the UWM in a 6-month follow-up period from April to October 2017 [15]. Other researchers showed similar evidence of care fragmentation for adult hospital visits in Massachusetts [16] and for emergency department visits in Indiana [17]. Typical models for forecasting an individual patient's cost presume complete historical data [14,18,19]. A healthcare system with incomplete data for its patients does not use these models, resulting in many predictably costly patients being missed by care management interventions and having poor outcomes.

Recently, we developed the first constraint-based method to pinpoint a reasonably large subset of patients apt to obtain care mostly within a given healthcare system [15]. For a 6-month follow-up period from April to October 2017, we showed that our method worked well for the cohort of all adult patients at the UWM [15]. It is unknown how our method performs on patients with various chronic diseases and over follow-up periods of different lengths. If our method performs well in these cases, for the subset of patients with chronic diseases that is pinpointed by our method and for which the healthcare system has more complete data, we could build a model to predict an individual patient's cost. This would be better than the current practice of not using any cost prediction model to facilitate care management for this healthcare system at all.

Objectives

To understand our method's potential to enable building models to predict an individual patient's cost on incomplete medical data, this study assesses our method's performance at the UWM on 5 subgroups of adult patients and over follow-up periods of 2 different lengths. Each subgroup corresponds to a separate one of the 5 major chronic diseases: asthma, CKD, COPD, type 1 diabetes (T1D), and type 2 diabetes (T2D), for which care management is used.

Methods

Patient population

As the largest academic healthcare system in Washington state, the UWM provides both clinic-based and hospital-based care for adults. As shown in Figure 1, our patient cohort covered all adult patients (age≥18) who visited the UWM during 2018 and had information kept in the UWM's enterprise data warehouse. Unless explicitly specified as a particular type of visit, a visit can be of any type in this paper. Patients who died during 2018 were excluded from our cohort.

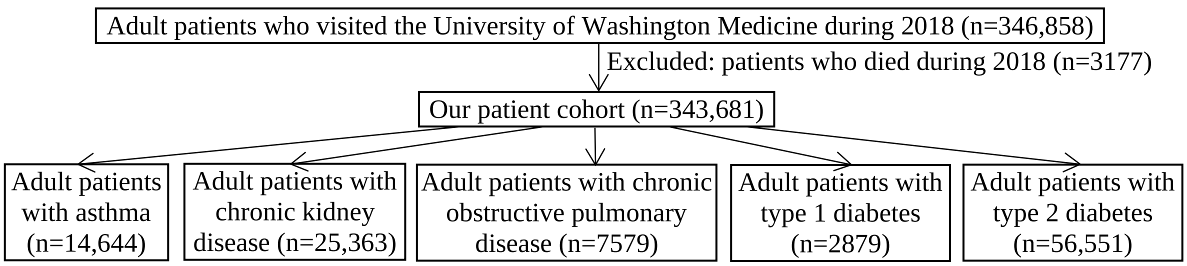

Figure 1. The patient cohort and the 5 patient subgroups.

Data set

We used clinical and administrative data in the UWM's enterprise data warehouse during 2011-2018. The data set included information on demographics, visits, diagnoses, laboratory tests, medications, and primary care physicians (PCPs) for patients in our cohort. We also used 2019 PreManage data of UWM patients. As a commercial product of Collective Medical Technologies Inc., PreManage provides diagnosis and visit data of hospital visits (emergency department visits and inpatient stays) at all hospitals in Washington state as well as many hospitals in other U.S. states [20]. As shown in Figure 2, we used January 1, 2019 as the index date to separate the subsequent and prior periods for our analysis task.

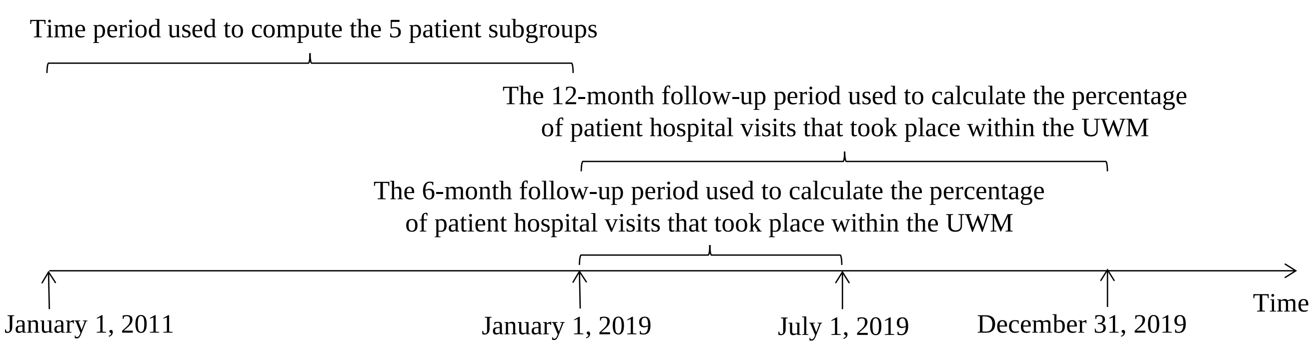

Figure 2. The time periods used to compute the patient subgroups and the percentages of patient hospital visits that took place within the UWM. UWM: University of Washington Medicine.

Patient subgroups

We considered 5 patient subgroups comprised of patients with a specific major chronic disease in our patient cohort in 2018. One subgroup was created for each of 5 major chronic diseases: asthma, CKD, COPD, T1D, and T2D.

Asthma case definition

A patient was deemed to have asthma in 2018 if the patient had ≥1 asthma diagnosis code (International Classification of Diseases, Ninth Revision [ICD-9]: 493.0x, 493.8x, 493.1x, 493.9x; International Classification of Diseases, Tenth Revision [ICD-10]: J45.x) in 2018 [21-23].

CKD case definition

A patient was deemed to have CKD if the patient had an estimated glomerular filtration rate (eGFR) < 60 mL/min/1.73m^2 or proteinuria in 2 measurements that were ≥3 months apart [24,25]. The UWM computed eGFR using the MDRD (Modification of Diet in Renal Disease) equation: eGFR (mL/min/1.73m^2) = 175 × age$^{-0.203}$ × serum creatinine$^{-1.154}$ × 0.742 (if female) × 1.212 (if black or African American) [26]. Proteinuria was detected when the urine dipstick test result for protein was ≥1+ (30 mg/dL) [24].

COPD case definition

By adjusting the criteria adopted by the National Quality Forum and the Centers for Medicare and Medicaid Services [27-29], we encompassed emergency department and outpatient visit data [30] to identify COPD patients. A patient was deemed to have COPD if the patient was ≥40 and fulfilled any of these 4 conditions:

1) an outpatient visit diagnosis code of COPD (ICD-9: 491.22, 491.21, 491.9, 491.8, 493.2x, 492.8, 496; ICD-10: J42, J41.8, J44.*, J43.*) followed by ≥1 prescription of long-acting muscarinic antagonist (aclidinium, glycopyrrolate, tiotropium, and umeclidinium) within 6 months,
2) ≥1 emergency department or ≥2 outpatient visit diagnosis codes of COPD (ICD-9: 491.22, 491.21, 491.9, 491.8, 493.2x, 492.8, 496; ICD-10: J42, J41.8, J44.*, J43.*),
3) ≥1 inpatient stay discharge having a principal diagnosis code of COPD (ICD-9: 491.22, 491.21, 491.9, 491.8, 493.2x, 492.8, 496; ICD-10: J42, J41.8, J44.*, J43.*), and
4) ≥1 inpatient stay discharge having a principal diagnosis code of respiratory failure (ICD-9: 518.82, 518.81, 799.1, 518.84; ICD-10: J96.0*, J80, J96.9*, J96.2*, R09.2) and a secondary diagnosis code of acute COPD exacerbation (ICD-9: 491.22, 491.21, 493.22, 493.21; ICD-10: J44.1, J44.0).

T1D and T2D case definition

We used Nichols *et al.*’s method [31] to identify diabetes patients. A patient was deemed to have diabetes if the patient had ≥1 inpatient stay diagnosis code of diabetes (ICD-9: 250.x, 357.2, 362.0x, 366.41; ICD-10: E10.x, E11.x) or any 2 of the following events happening within 2 years of each other:

1) hemoglobin A1c (HbA1c) ≥6.5%,
2) random plasma glucose ≥200 mg/dL,
3) fasting plasma glucose ≥126 mg/dL,
4) an outpatient visit diagnosis code of diabetes (ICD-9: 250.x, 357.2, 362.0x, 366.41; ICD-10: E10.x, E11.x), and
5) a prescription of anti-hyperglycemic medication (α-glucosidase inhibitor, amylin analogue, biguanide, dipeptidyl peptidase-4 inhibitor, incretin mimetic, insulin, meglitinide, sulfonylurea, and thiazolidinedione).

Two events of the same type, like 2 events of HbA1c ≥6.5%, would qualify if they happened on 2 different days. Since metformin, a biguanide, and thiazolidinedione could be used for other diseases, we did not count 2 prescriptions of metformin or thiazolidinedione with no other manifestation of diabetes. We also excluded events occurring during pregnancy.

We used Klompas *et al.*’s method [32,33] to distinguish T1D and T2D. Using all diagnosis codes, laboratory test results, and medication prescriptions during 2011-2018, we regarded a diabetic patient to have T1D if the patient fulfilled any of the following 4 conditions:

1) the number of T1D diagnosis codes (ICD-9: 250.x3, 250.x1; ICD-10: E10.x) > the number of T2D diagnosis codes (ICD-9: 250.x2, 250.x0; ICD-10: E11.x) and a prescription of glucagon,
2) the number of T1D diagnosis codes (ICD-9: 250.x3, 250.x1; ICD-10: E10.x) > the number of T2D diagnosis codes (ICD-9: 250.x2, 250.x0; ICD-10: E11.x) and no prescription of any oral hypoglycemic medication other than metformin,
3) a negative C-peptide laboratory test result, and
4) a positive diabetes autoantibody laboratory test result.

A diabetic patient was deemed to have T2D if the patient was not deemed to have T1D.

Our recently developed constraint-based method for identifying patients

We looked at 3 UWM hospitals whose clinical and administrative data are kept in the UWM’s enterprise data warehouse: University of Washington Medical Center, Harborview Medical Center, and Northwest Hospital. They are all in Seattle, Washington. To identify patients apt to obtain care mostly within the UWM, we used the parameterized PCP constraint developed in our recent paper [15]: the patient resides within d miles of ≥1 of the 3 UWM hospitals and has a UWM PCP. The distance between a UWM hospital and a patient's home is the ellipsoid great circle distance computed by the distVincentyEllipsoid function contained in R's geosphere package version 1.5-5 [34]. d is a parameter. For all UWM adult patients and the follow-up period of 6 months, we showed that d's optimal value is 5 [15]. The UWM PCPs are inclined to refer within the UWM. Thus, intuitively, patients having a UWM PCP are apt to obtain a larger percentage of their care from the UWM than other patients. All else being equal, the UWM tends to provide a larger portion of a patient's care when the patient resides closer to UWM hospitals. Moreover, the number of UWM patients fulfilling the constraint grows with d. When $d=+\infty$, distance plays no role any more.

Data analysis

As shown in Figure 2, we considered 2 follow-up periods starting from January 1, 2019: the subsequent 6 months (January 1, 2019 - June 30, 2019) and the subsequent 12 months (January 1, 2019 - December 31, 2019). The 6-month follow-up period was chosen to be consistent with the length of the follow-up period used in our prior paper [15]. The 12-month follow-up period was chosen because to facilitate care management, typically at least 1 year of historical data is needed to build models to predict an individual patient's cost [14]. For each of the 5 patient subgroups and each of the 2 follow-up periods, we computed our method's performance on identifying patients apt to obtain care mostly within the UWM. We employed administrative data in the UWM's enterprise data warehouse to assess whether a patient fulfilled the parameterized PCP constraint. For each of the 5 patient subgroups, we calculated the percentage of patients fulfilling the constraint = $n_0/d_0\times100\%$. Here, n_0 is the number of patients in the subgroup fulfilling the constraint. d_0 is the number of patients in the subgroup. For all patients in the subgroup fulfilling the constraint, we used PreManage data to calculate:

(1) the percentage of their hospital visits taking place within the UWM in the subsequent 6 months = $n_1/d_1\times100\%$. Here, n_1 is the number of their hospital visits taking place within the UWM in the subsequent 6 months. d_1 is the number of their hospital visits taking place anywhere in the subsequent 6 months.
(2) the percentage of their hospital visits taking place within the UWM in the subsequent 12 months = $n_2/d_2\times100\%$. Here, n_2 is the number of their hospital visits taking place within the UWM in the subsequent 12 months. d_2 is the number of their hospital visits taking place anywhere in the subsequent 12 months.

Since an average hospital visit costs much more than an average visit of another type, this percentage signifies the proportion of those patients' care obtained from the UWM.

When deciding the optimal value of the distance threshold d to use, we struck a balance between 2 goals:

(1) **Goal 1**: For patients fulfilling the constraint, the proportion of their hospital visits taking place within the UWM should be as large as possible. This will maximize the completeness of UWM medical data and minimize bias in the results of analyses done on those data. As these patients each have a UWM PCP, we expect most of their outpatient visits to occur within the UWM in the subsequent 12 months.
(2) **Goal 2**: The percentage of patients fulfilling the constraint should be as large as possible. This will help maximize the impact of the application using UWM medical data.

To show how our method performs for every UWM hospital, for all patients in the subgroup fulfilling the constraint, we employed PreManage data to calculate:

(1) the percentage of their hospital visits taking place at the UWM hospital in the subsequent 6 months = $n_3/d_1\times100\%$. Here, n_3 is the number of their hospital visits taking place at the UWM hospital in the subsequent 6 months. Recall d_1 is the number of their hospital visits taking place anywhere in the subsequent 6 months.
(2) the percentage of their hospital visits taking place at the UWM hospital in the subsequent 12 months = $n_4/d_2\times100\%$. Here, n_4 is the number of their hospital visits taking place at the UWM hospital in the subsequent 12 months. Recall d_2 is the number of their hospital visits taking place anywhere in the subsequent 12 months.

Ethics approval

The UWM's institutional review board approved this retrospective cohort study.

Results

Table 1 shows our patient cohort's demographic and clinical characteristics.

Table 1. Demographic and clinical characteristics of the adult patients who visited UWM facilities during 2018 with information kept in the UWM's enterprise data warehouse (N=343,681).

Characteristic	*n* (%)
Age	
18 to <40	120,422 (35.04%)
40 to 65	149,418 (43.48%)

>65	73,841 (21.49%)
Gender	
Male	159,964 (46.54%)
Female	183,701 (53.45%)
Unknown or not reported	16 (0.00%)
Race	
Black or African American	25,513 (7.42%)
American Indian or Alaska native	4,795 (1.40%)
Asian	34,474 (10.03%)
Native Hawaiian or other Pacific islander	2,843 (0.83%)
Multiple races	1 (0.00%)
Unknown or not reported	45,094 (13.12%)
White	230,961 (67.20%)
Ethnicity	
Non-Hispanic	271,582 (79.02%)
Hispanic	21,718 (6.32%)
Unknown or not reported	50,381 (14.66%)
Insurance	
Private	163,908 (47.69%)
Public (Medicare and Medicaid)	160,026 (46.56%)
Self-paid or charity	19,747 (5.75%)
Disease	
Asthma	14,644 (4.26%)
Chronic kidney disease	25,363 (7.38%)
Chronic obstructive pulmonary disease	7,579 (2.21%)
Type 2 diabetes	56,551 (16.45%)
Type 1 diabetes	2,879 (0.84%)

Figures 3 and 4 present the percentage of patients in each of the 5 patient subgroups fulfilling the parameterized PCP constraint. The percentage rises with increase in d, at first swiftly when d is small and then at a slower pace when d grows larger. Recall that d is the largest permitted distance in miles between the patient's home and the nearest UWM hospital.

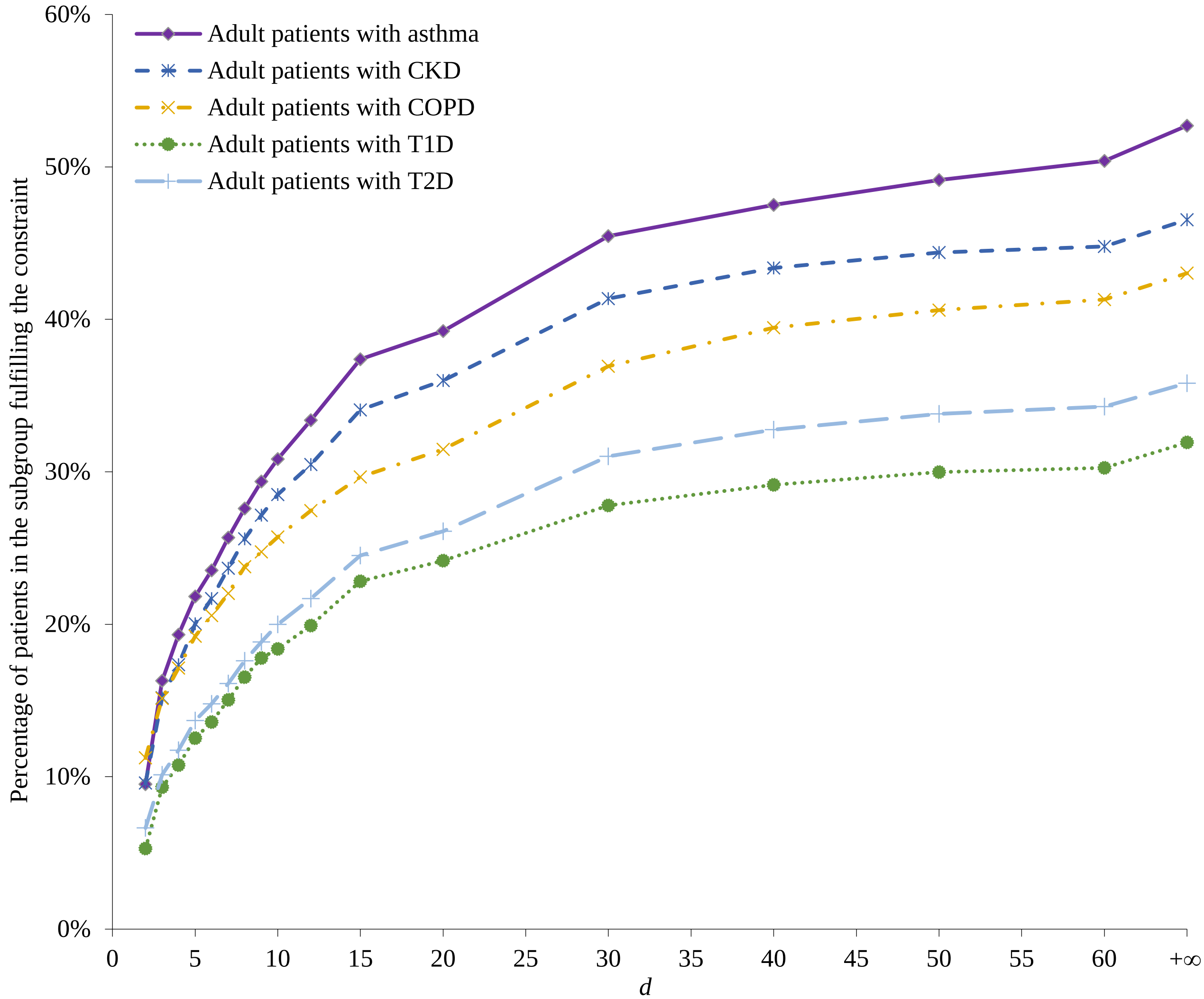

Figure 3. The percentage of patients in each of the 5 patient subgroups fulfilling the parameterized primary care physician (PCP) constraint. CKD: chronic kidney disease. COPD: chronic obstructive pulmonary disease. T1D: type 1 diabetes. T2D: type 2 diabetes.

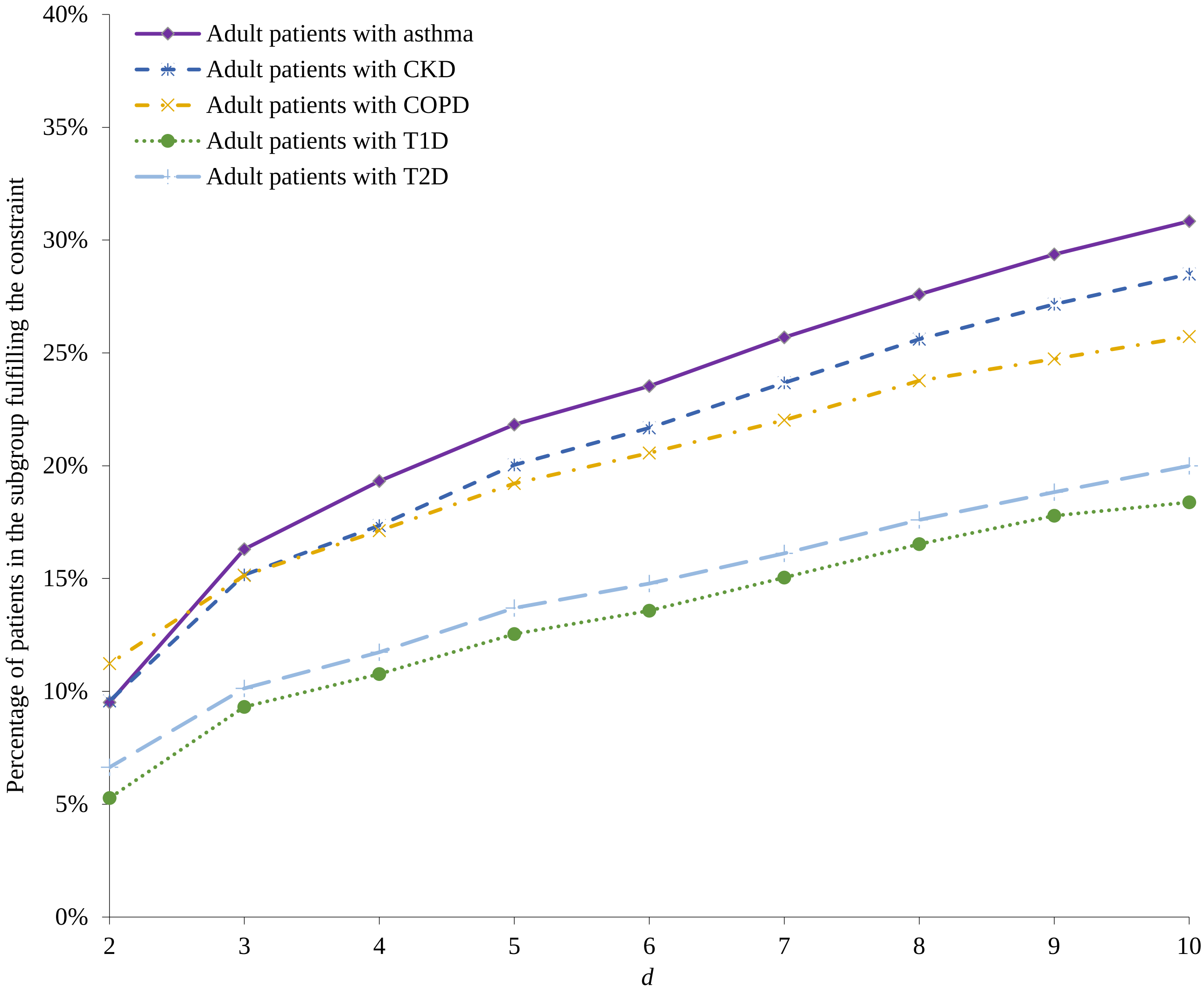

Figure 4. The percentage of patients in each of the 5 patient subgroups fulfilling the parameterized primary care physician (PCP) constraint, when *d* is ≤10. CKD: chronic kidney disease. COPD: chronic obstructive pulmonary disease. T1D: type 1 diabetes. T2D: type 2 diabetes.

For all patients in each of the 5 patient subgroups fulfilling the parameterized PCP constraint, Figures 5 and 6 present the percentages of their hospital visits taking place within the UWM in the subsequent 6 months and in the subsequent 12 months. With a small number of exceptions at small values of *d*, the percentage drops as *d* grows, swiftly when *d* is small and then at a slower pace when *d* grows larger. This reflects that all else being equal, patients residing closer to UWM hospitals are more inclined to visit them. No matter how small *d* is, the percentage never becomes 100%, partially because the patients can also visit multiple non-UWM hospitals within 1 mile of some of the UWM hospitals. For each positive *d*, the percentage is relatively similar across the 5 patient subgroups and the 2 follow-up periods.

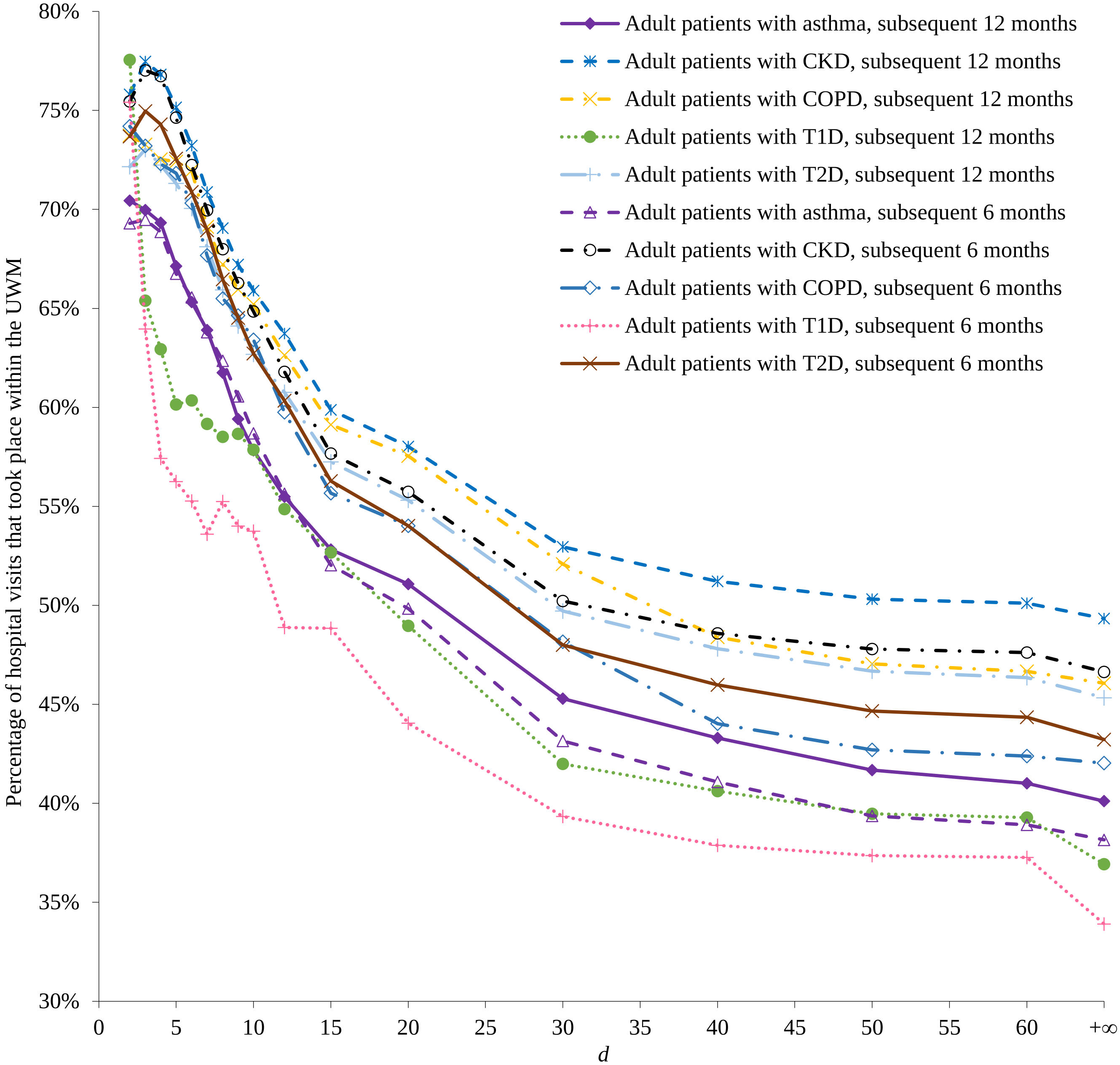

Figure 5. For all patients in each of the 5 patient subgroups fulfilling the parameterized primary care physician (PCP) constraint, the percentages of their hospital visits taking place within the University of Washington Medicine (UWM) in the subsequent 6 months and in the subsequent 12 months. CKD: chronic kidney disease. COPD: chronic obstructive pulmonary disease. T1D: type 1 diabetes. T2D: type 2 diabetes.

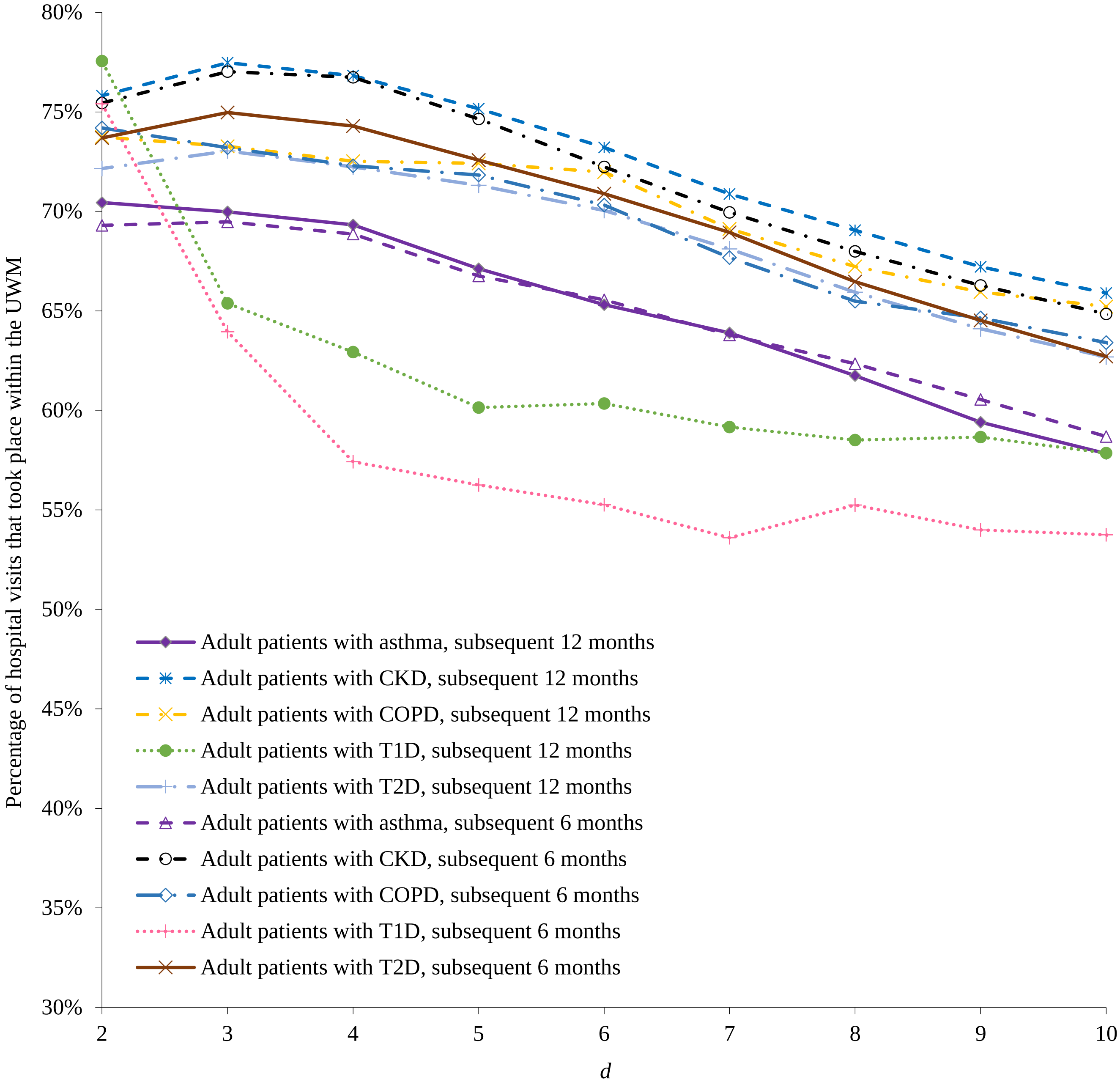

Figure 6. For all patients in each of the 5 patient subgroups fulfilling the parameterized primary care physician (PCP) constraint when d is ≤10, the percentages of their hospital visits taking place within the University of Washington Medicine (UWM) in the subsequent 6 months and in the subsequent 12 months. CKD: chronic kidney disease. COPD: chronic obstructive pulmonary disease. T1D: type 1 diabetes. T2D: type 2 diabetes.

By striking a balance between the 2 goals mentioned at the end of the Methods section, we chose d=5 as the optimal value to use for each patient subgroup and each follow-up period. This value is the same as that chosen in our prior paper [15]. For d=5 and for each of the 5 patient subgroups, Table 2 shows the percentage of patients fulfilling the parameterized PCP constraint, the percentages of patient hospital visits taking place within the UWM in the subsequent 6 months and in the subsequent 12 months, and the percentages of hospital visits by patients fulfilling the constraint that took place within the UWM in the subsequent 6 months and in the subsequent 12 months.

Table 2. For *d*=5 and for each of the 5 patient subgroups, the percentage of patients fulfilling the parameterized primary care physician (PCP) constraint, the percentages of patient hospital visits taking place within the University of Washington Medicine (UWM) in the subsequent 6 months and in the subsequent 12 months, and the percentages of hospital visits by patients fulfilling the constraint that took place within the UWM in the subsequent 6 months and in the subsequent 12 months.

Patient subgroup	Percentage of patients fulfilling the parameterized PCP constraint	Percentage of patient hospital visits taking place within the UWM in the subsequent 6 months	Percentage of hospital visits by patients fulfilling the constraint that took place within the UWM in the subsequent 6 months	Percentage of patient hospital visits taking place within the UWM in the subsequent 12 months	Percentage of hospital visits by patients fulfilling the constraint that took place within the UWM in the subsequent 12 months
Adult patients with asthma	21.81% (3,194/14,640)	37.11% (2,648/7,135)	66.75% (797/1,194)	37.66% (6,857/18,206)	67.13% (1,997/2,975)
Adult patients with chronic kidney disease	20.03% (5,081/25,363)	40.77% (7,503/18,404)	74.64% (2,178/2,918)	42.52% (19,558/45,994)	75.16% (5,634/7,496)
Adult patients with chronic obstructive pulmonary disease	19.21% (1,456/7,579)	38.85% (2,587/6,659)	71.82% (831/1,157)	41.47% (7,026/16,941)	72.42% (2,179/3,009)
Adult patients with type 1 diabetes	12.54% (361/2,879)	23.78% (317/1,333)	56.25% (63/112)	25.38% (845/3,330)	60.14% (169/281)
Adult patients with type 2 diabetes	13.69% (7,744/56,551)	35.58% (10,926/30,707)	72.57% (2,847/3,923)	36.69% (29,272/79,775)	71.31% (7,177/10,065)

For every UWM hospital and all patients in each of the 5 patient subgroups fulfilling the parameterized PCP constraint, Figures 7-11 in the Appendix show the percentages of their hospital visits taking place at the UWM hospital in the subsequent 6 months and in the subsequent 12 months. For every (patient subgroup, positive *d*) pair, each percentage differs across the 3 UWM hospitals. As *d* grows, each percentage drops at similar paces across the 3 UWM hospitals. For each (positive *d*, UWM hospital) pair, the percentage is relatively similar across the 5 patient subgroups and the 2 follow-up periods.

Discussion

Principal results

For each of the 5 major chronic diseases most relevant to care management (asthma, COPD, CKD, T1D, and T2D), our constraint-based method can pinpoint a reasonably large subset of patients apt to obtain care mostly within the UWM. Using our method to pinpoint a subset of UWM adult patients with asthma, we roughly doubled the percentage of patient hospital visits taking place within the UWM in the subsequent 6 months from 37.11% (2,648/7,135) to 66.75% (797/1,194), and the corresponding percentage for the subsequent 12 months from 37.66% (6,857/18,206) to 67.13% (1,997/2,975). The results for adult patients with CKD, adult patients with COPD, adult patients with T1D, and adult patients with T2D are relatively similar. As the patients fulfilling the constraint all have a UWM PCP, we expect a majority of their outpatient visits to happen within the UWM in the subsequent 12 months, although we did not examine this in our study.

Comparison with our prior work

The performance numbers shown in this paper are relatively similar to those that our prior paper [15] showed for the group of all adult patients and the 6-month follow-up period from April to October 2017. There, using our constraint-based method with *d*=5 to pinpoint 16.01% (55,707/348,054) of UWM adult patients, we roughly doubled the percentage of patient hospital visits taking place within the UWM in the subsequent 6 months from 31.80% (39,171/123,162) to 69.38% (10,501/15,135).

Differences in the results for patients with T1D and patients with T2D

T1D tends to occur in younger people than T2D. Many young adults are students at the University of Washington and several other universities in the Seattle metropolitan area. During the summer and other university breaks, many of these students return to their hometowns outside the Seattle metropolitan area that the UWM primarily serves. The hospital visits they incur during these periods are likely to be outside of the UWM. Partly due to this, as shown in Table 2, the percentage of hospital visits by UWM adult patients with T1D taking place within the UWM in each follow-up period is about 30% less than the corresponding percentage for UWM adult patients with T2D. Looking only at patients fulfilling the parameterized PCP constraint with *d*=5, the percentage of hospital visits by patients with T1D taking place within the UWM in each follow-up period is about 15-30% less than the corresponding percentage for patients with T2D.

Possible use of our results

We showed that for each of 5 major chronic diseases most relevant to care management, the UWM offers most of the care and has decently complete medical data for patients fulfilling the parameterized PCP constraint with d=5. For these patients, we can build a predictive model to identify future high-cost patients among them and intervene preemptively via care management to avert poor outcomes [1-5]. As patients residing farther from the 3 UWM hospitals are inclined to obtain a smaller percentage of their care from the UWM, the UWM could consider adopting differing preventive interventions for patients residing at different distances from the UWM hospitals. This could help care management gain better results. For patients obtaining only a small percentage of their care from the UWM, it is hard for the UWM to adopt costly preventive interventions economically.

Like many other healthcare systems, the UWM has no complete claims data on its patients' healthcare use outside of the UWM. If a healthcare system has complete claims data on its patients' outside healthcare use, we could employ claims data instead of PreManage data to do a similar study.

A healthcare system with no access to PreManage data could also adopt our method. Without using PreManage data, one could assess our method's performance by asking some patients of the healthcare system about the care they obtained everywhere.

Limitations

This study has 2 limitations, which could be interesting topics for future work:

(1) This study assesses our method at the UWM, which primarily serves an urban region. To know how our method generalizes, we need to redo our analysis at other healthcare systems, some providing care to rural regions and others primarily serving urban regions. Patients concentrate more in urban regions than in rural regions. For a healthcare system primarily serving rural regions, we expect d's optimal value to be >5.

(2) For a healthcare system having incomplete medical data for its patients, we can employ our method to identify a subset of patients apt to obtain care mostly within the healthcare system and assess the degree of data incompleteness for this subset. Analyzing incomplete data could lead to biased results, which are better than no result if the degree of bias is small. At present, we know neither the exact relation between data incompleteness and bias nor the extent of data incompleteness that can be tolerated before the results of data analysis become invalid. To assess whether our method could safely enable the data analysis task in such a healthcare system, we could obtain a more complete data set from Kaiser Permanente or any other similar healthcare system, drop differing portions of the data set, and assess the impact on the analysis results.

Conclusions

For each of the 5 major chronic diseases most relevant to care management (asthma, COPD, CKD, T1D, and T2D), our constraint-based method can pinpoint a reasonably large subset of patients apt to obtain care mostly within the UWM. This opens the door to building models to predict an individual patient's cost on incomplete data, which was formerly deemed infeasible.

Acknowledgments

We thank Hye-Eun Yoon for useful discussions. GL was partially supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL142503. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. YT did the work at the University of Washington when she was a visiting PhD student there.

Authors' contributions

YT participated in doing data analysis, performing literature review, and writing the paper's first draft. GL conceptualized and designed the study, participated in doing data analysis and performing literature review, and wrote most of the paper. ZL and PT provided feedback on various medical issues, contributed to conceptualizing the presentation, and revised the paper.

Conflicts of Interest

None declared.

List of abbreviations

CKD: chronic kidney disease

COPD: chronic obstructive pulmonary disease
eGFR: estimated glomerular filtration rate
HbA1c: hemoglobin A1c
ICD-9: International Classification of Diseases, Ninth Revision
ICD-10: International Classification of Diseases, Tenth Revision
PCP: primary care physician
T1D: type 1 diabetes
T2D: type 2 diabetes
UWM: University of Washington Medicine

References

1. Mays GP, Claxton G, White J. Managed care rebound? Recent changes in health plans' cost containment strategies. Health Aff (Millwood) 2004;Suppl Web Exclusives:W4-427-36. PMID:15451964
2. Curry N, Billings J, Darin B, Dixon J, Williams M, Wennberg D. Predictive Risk Project Literature Review. London: King's Fund. 2005. http://www.kingsfund.org.uk/sites/files/kf/field/field_document/predictive-risk-literature-review-june2005.pdf.
3. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, Blumenthal D. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007;22 Suppl 3:391-5. PMID:18026807
4. Caloyeras JP, Liu H, Exum E, Broderick M, Mattke S. Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. Health Aff (Millwood) 2014;33(1):124-31. PMID:24395944
5. Nelson L. Congressional Budget Office. Lessons from Medicare's demonstration projects on disease management and care coordination. 2012. https://www.cbo.gov/sites/default/files/112th-congress-2011-2012/workingpaper/WP2012-01_Nelson_Medicare_DMCC_Demonstrations_1.pdf.
6. Fishbane S, Agoritsas S, Bellucci A, Halinski C, Shah HH, Sakhiya V, Balsam L. Augmented nurse care management in CKD stages 4 to 5: a randomized trial. Am J Kidney Dis 2017;70(4):498-505. PMID:28396108
7. Greineder DK, Loane KC, Parks P. A randomized controlled trial of a pediatric asthma outreach program. J Allergy Clin Immunol 1999;103(3 Pt 1):436-40. PMID:10069877
8. Kelly CS, Morrow AL, Shults J, Nakas N, Strope GL, Adelman RD. Outcomes evaluation of a comprehensive intervention program for asthmatic children enrolled in Medicaid. Pediatrics 2000;105(5):1029-35. PMID:10790458
9. Axelrod RC, Zimbro KS, Chetney RR, Sabol J, Ainsworth VJ. A disease management program utilizing life coaches for children with asthma. J Clin Outcomes Manag 2001;8(6):38-42.
10. Rubin RJ, Dietrich KA, Hawk AD. Clinical and economic impact of implementing a comprehensive diabetes management program in managed care. J Clin Endocrinol Metab 1998;83(8):2635-42. PMID:9709924
11. Dorr DA, Wilcox AB, Brunker CP, Burdon RE, Donnelly SM. The effect of technology-supported, multidisease care management on the mortality and hospitalization of seniors. J Am Geriatr Soc 2008;56(12):2195-202. PMID:19093919
12. Beaulieu N, Cutler DM, Ho K, Isham G, Lindquist T, Nelson A, O'Connor P. The business case for diabetes disease management for managed care organizations. Forum for Health Economics & Policy 2006;9(1):1-37.
13. Levine SH, Adams J, Attaway K, Dorr DA, Leung M, Popescu P, Rich J. Predicting the financial risks of seriously ill patients. California HealthCare Foundation, 2011. https://www.chcf.org/wp-content/uploads/2017/12/PDF-PredictiveModelingRiskStratification.pdf.
14. Duncan I. Healthcare Risk Adjustment and Predictive Modeling, 2nd ed. New Hartford, CT: ACTEX Learning; 2018. ISBN:1635884136
15. Luo G, Tarczy-Hornoch P, Wilcox AB, Lee ES. Identifying patients who are likely to receive most of their care from a specific health care system: demonstration via secondary analysis. JMIR Med Inform 2018;6(4):e12241. PMID:30401670
16. Bourgeois FC, Olson KL, Mandl KD. Patients treated at multiple acute health care facilities: quantifying information fragmentation. Arch Intern Med 2010;170(22):1989-95. PMID:21149756
17. Finnell JT, Overhage JM, Grannis S. All health care is not local: an evaluation of the distribution of emergency department care delivered in Indiana. AMIA Annu Symp Proc. 2011;2011:409-16. PMID:22195094
18. Ash A, McCall N. Risk assessment of military populations to predict health care cost and utilization. 2005. http://www.rti.org/pubs/tricare_riskassessment_final_report_combined.pdf.
19. Iezzoni LI. Risk adjustment for measuring health care outcomes, 4th ed. Chicago, IL: Health Administration Press; 2013. ISBN:1567934374
20. Collective Medical Technologies Inc. Collective Medical and Consonus Healthcare announce partnership to improve post-acute transitions of care. 2018. https://collectivemedical.com/resources/press-release/collective-medical-and-consonus-healthcare-announce-partnership-to-improve-post-acute-transitions-of-care.

21. Schatz M, Cook EF, Joshua A, Petitti D. Risk factors for asthma hospitalizations in a managed care organization: development of a clinical prediction rule. Am J Manag Care 2003;9(8):538-47. PMID:12921231
22. Desai JR, Wu P, Nichols GA, Lieu TA, O'Connor PJ. Diabetes and asthma case identification, validation, and representativeness when using electronic health data to construct registries for comparative effectiveness and epidemiologic research. Med Care 2012;50 Suppl:S30-5. PMID:22692256
23. Wakefield DB, Cloutier MM. Modifications to HEDIS and CSTE algorithms improve case recognition of pediatric asthma. Pediatr Pulmonol 2006;41(10):962-71. PMID:16871628
24. Lee Y, Park S, Lee S, Kim Y, Kang MW, Cho S, Park S, Han K, Kim YC, Han SS, Lee H, Lee JP, Joo KW, Lim CS, Kim YS, Kim DK. Lipid profiles and risk of major adverse cardiovascular events in CKD and diabetes: A nationwide population-based study. PLoS One 2020;15(4):e0231328. PMID:32271842
25. Levey AS, Becker C, Inker LA. Glomerular filtration rate and albuminuria for detection and staging of acute and chronic kidney disease in adults: a systematic review. JAMA 2015;313(8):837-46. PMID:25710660
26. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, Kusek JW, Van Lente F; Chronic Kidney Disease Epidemiology Collaboration. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med 2006;145(4):247-54. PMID:16908915
27. National Quality Forum. NQF #1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization. http://www.qualityforum.org/Projects/n-r/Pulmonary_Endorsement_Maintenance/1891_30_Day_RSRR_COPD.aspx.
28. Nguyen HQ, Chu L, Amy Liu IL, Lee JS, Suh D, Korotzer B, Yuen G, Desai S, Coleman KJ, Xiang AH, Gould MK. Associations between physical activity and 30-day readmission risk in chronic obstructive pulmonary disease. Ann Am Thorac Soc 2014;11(5):695-705. PMID:24713094
29. Lindenauer PK, Grosso LM, Wang C, Wang Y, Krishnan JA, Lee TA, Au DH, Mularski RA, Bernheim SM, Drye EE. Development, validation, and results of a risk-standardized measure of hospital 30-day mortality for patients with exacerbation of chronic obstructive pulmonary disease. J Hosp Med 2013;8(8):428-35. PMID:23913593
30. Cooke CR, Joo MJ, Anderson SM, Lee TA, Udris EM, Johnson E, Au DH. The validity of using ICD-9 codes and pharmacy records to identify patients with chronic obstructive pulmonary disease. BMC Health Serv Res 2011;11:37. PMID:21324188
31. Nichols GA, Schroeder EB, Karter AJ, Gregg EW, Desai J, Lawrence JM, O'Connor PJ, Xu S, Newton KM, Raebel MA, Pathak RD, Waitzfelder B, Segal J, Lafata JE, Butler MG, Kirchner HL, Thomas A, Steiner JF; SUPREME-DM Study Group. Trends in diabetes incidence among 7 million insured adults, 2006-2011: the SUPREME-DM project. Am J Epidemiol 2015;181(1):32-9. PMID:25515167
32. Klompas M, Eggleston E, McVetta J, Lazarus R, Li L, Platt R. Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data. Diabetes Care 2013;36(4):914-21. PMID:23193215
33. Schroeder EB, Donahoo WT, Goodrich GK, Raebel MA. Validation of an algorithm for identifying type 1 diabetes in adults based on electronic health record data. Pharmacoepidemiol Drug Saf 2018;27(10):1053-9. PMID:29292555
34. R's geosphere package homepage. 2020. https://cran.r-project.org/web/packages/geosphere/index.html.

Appendix

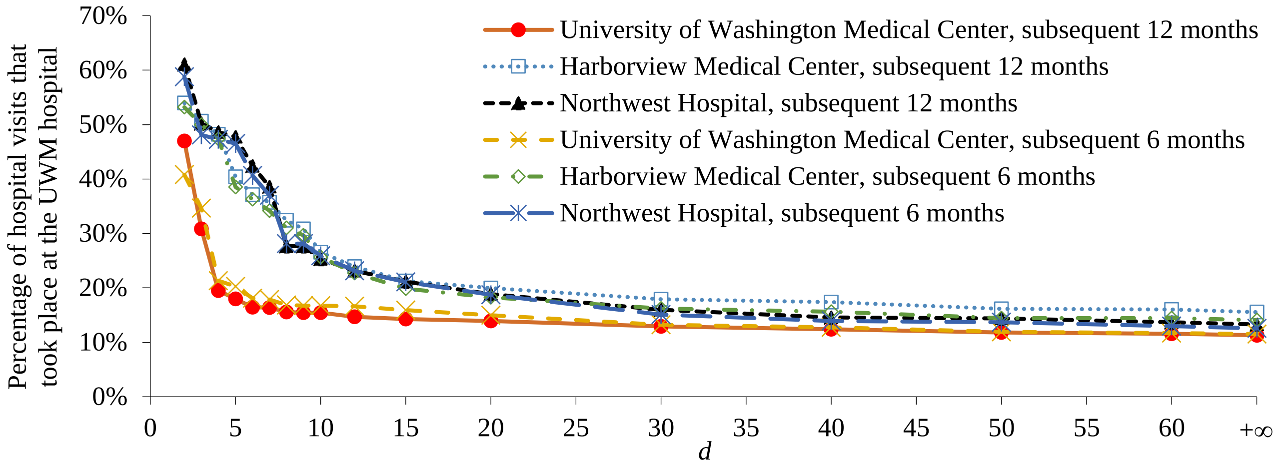

Figure 7. For every University of Washington Medicine (UWM) hospital and all adult patients with asthma fulfilling the parameterized primary care physician (PCP) constraint, the percentages of their hospital visits taking place at the UWM hospital in the subsequent 6 months and in the subsequent 12 months.

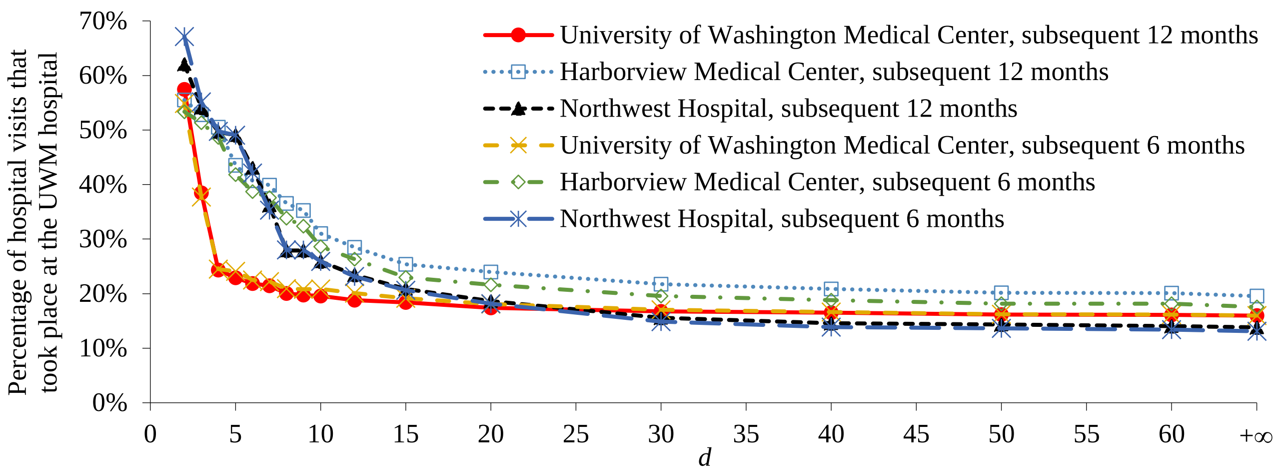

Figure 8. For every University of Washington Medicine (UWM) hospital and all adult patients with chronic kidney disease fulfilling the parameterized primary care physician (PCP) constraint, the percentages of their hospital visits taking place at the UWM hospital in the subsequent 6 months and in the subsequent 12 months.

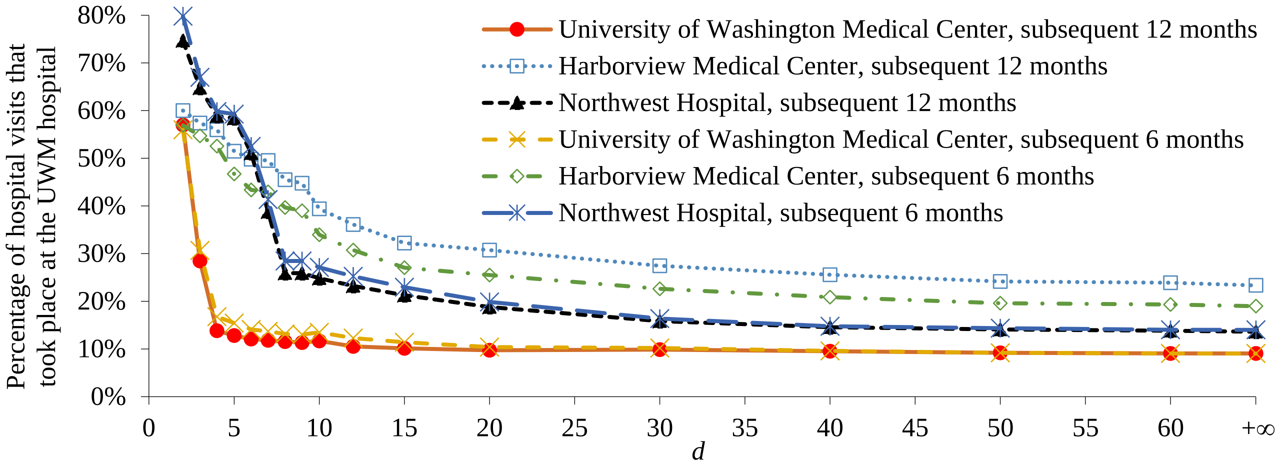

Figure 9. For every University of Washington Medicine (UWM) hospital and all adult patients with chronic obstructive pulmonary disease fulfilling the parameterized primary care physician (PCP) constraint, the percentages of their hospital visits taking place at the UWM hospital in the subsequent 6 months and in the subsequent 12 months.

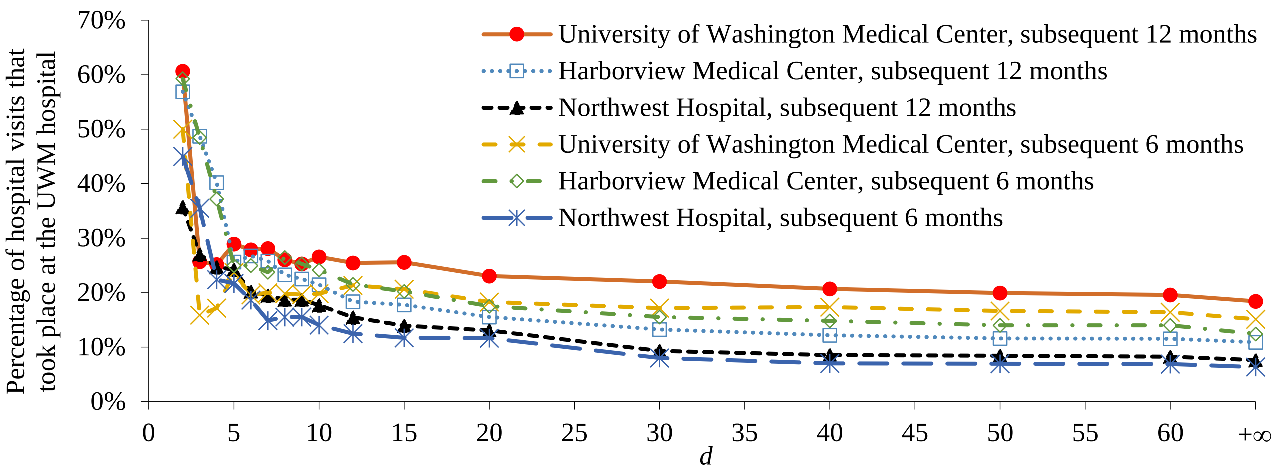

Figure 10. For every University of Washington Medicine (UWM) hospital and all adult patients with type 1 diabetes fulfilling the parameterized primary care physician (PCP) constraint, the percentages of their hospital visits taking place at the UWM hospital in the subsequent 6 months and in the subsequent 12 months.

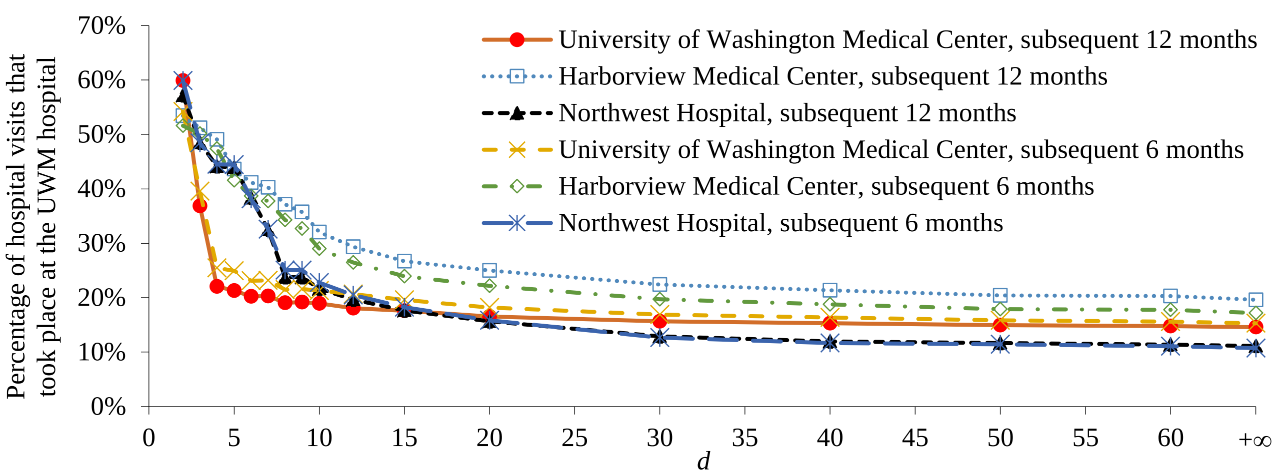

Figure 11. For every University of Washington Medicine (UWM) hospital and all adult patients with type 2 diabetes fulfilling the parameterized primary care physician (PCP) constraint, the percentages of their hospital visits taking place at the UWM hospital in the subsequent 6 months and in the subsequent 12 months.