Predicting Divertible Medicaid Emergency Department Costs

Justine S. Hastings*

Research Improving People's Lives, Providence RI
Department of Economics, University of Washington, Seattle, WA
National Bureau of Economic Research, Cambridge, MA

Mark Howison

Research Improving People's Lives, Providence, RI

Lowering health care costs is a policy priority for public health insurance programs. Policies that divert Emergency Department (ED) care to less costly, more effective health services are a promising avenue for cost savings. Using comprehensive, anonymized administrative data from the State of Rhode Island, which includes Medicaid and other social insurance programs, we demonstrate how government can identify and conduct efficient outreach to Medicaid recipients at risk of becoming high-cost ED users for potentially divertible care. The top predictors from our models include age, Medicaid enrollment and eligibility factors, and prior medical procedures. Our predictive models capture more future divertible spending than existing methods for identifying ED "super-utilizers" based on multiple prior ED visits. By using comprehensive administrative data that includes all of the state's social insurance programs, we can also establish connections between predicted high-cost ED users and existing case management in non-Medicaid programs. Policymakers could use these models to improve their identification of divertible spending and reduce the need for *de novo* outreach and case management programs.

CCS CONCEPTS • Applied computing → Life and medical sciences → **Health informatics** • Social and professional topics → Professional topics → Computing and business → **Economic impact** • Computing methodologies → Machine learning → Machine learning approaches → **Learning linear models**

Additional Keywords and Phrases: Fact-based policy, predictive modeling, health care

* Contact: connect@ripl.org

1 INTRODUCTION

There is a consensus view that many visits to the Emergency Department (ED) could be avoided by treatment in primary care [1]. Policymakers are therefore interested in developing policies to divert ED visits to less costly, more effective health services [2]. This is particularly true for the Medicaid program. Medicaid is the largest means-tested program in the United States in terms of spending, and Medicaid recipients have disproportionately large ED use when compared to privately-insured individuals [3], making lowering program costs a policy priority [4].

State and federal policymakers have begun trialing programs to identify and divert frequent ED users among Medicaid recipients [3,5]. These programs often rely on identifying individuals based on frequent ED visits, and then assigning caseworkers to help reduce future unnecessary or preventable ED use [5].

Randomized trial evidence on these programs has not shown consistent cost reduction [6]. More generally, approaches which rely on identifying individuals in a high-cost year to intervene in future years may be conceptually problematic; in the presence of mean reversion, the targeted years are likely to be lower cost than the year used for identification. These approaches may miss the highest expense year, and substantial costs accrue in order to identify targets for intervention.

A descriptive study of ED super-utilizers from a safety-net health system in Colorado found evidence of mean reversion through reductions in ED spending of 44.1 percent after one year and 58.6 percent after two years from a baseline year of frequent ED visits [7]. Additionally, the study found that while super-utilizers exhibited stable population-level characteristics over time, there was greater variability at the individual level, and only 6 percent of individuals continuously met the definition of super-utilization throughout the 24-month study period. Both the mean reversion in spending and the variability in super-utilizer status over time suggest that current approaches to identifying super-utilization based on frequency of visits are not optimal for helping policymakers reach the population who might benefit most from diversion to more appropriate care. A predictive modeling approach might better identify this population by anticipating divertible costs at the individual level.

In this study, we make use of a new database of administrative records from the state of Rhode Island to develop a data-driven method for predicting divertible ED use among Medicaid recipients. Improving the availability and quality of Medicaid care while controlling costs is a key initiative of the Office of the Governor in Rhode Island [8]. We use a machine-learning approach to predict which individuals who have no divertible ED spending in a given year will emerge as high-cost in the subsequent year. In principle, this allows us to identify individuals in the first year they become expensive, and to target individuals who may be easier to divert to lower-cost, effective care before they exhibit high ED spending. We identify characteristics which predict high costs and possible avenues for intervention using existing connections with case management in non-Medicaid programs.

2 DATA AND METHODS

We use anonymized Rhode Island state administrative records housed in a secure data enclave [9]. These records include administrative claims for 556,767 anonymous Medicaid enrollees in the state of Rhode Island between 2004 and 2014. The database also contains detailed information on these individuals' use of social benefit and insurance programs, incarceration history, quarterly wages, and demographics. Personally identifiable information has been removed from the data and replaced with anonymous identifiers that make it

possible for researchers with approved access to join and analyze records associated with the same individual while preserving anonymity [10].

Our objective was to define a panel of Rhode Island Medicaid enrollees in each quarter from 2004 and 2013; identify claims for potentially divertible Emergency Department (ED) care; define “cost emergence” as an enrollee’s change in potentially divertible ED spending from zero in the 12 preceding months to non-zero in the subsequent 12 months; and predict cost emergence using information known in administrative records for the enrollee in the preceding 12 months. Although anonymized administrative records as recent as 2015 were available for study, Medicaid expansion under the Affordable Care Act took place in Rhode Island during 2014. Therefore, we restricted the study to Medicaid enrollees in the pre-expansion period of 2004 to 2013 to prevent any confounding effects of changes in Medicaid eligibility under the Affordable Care Act.

Data are available through individual data sharing agreements with each of the following Rhode Island agencies: the Department of Human Services (DHS), the Department of Labor and Training (DLT), the Department of Corrections (DOC), and the Executive Office of Health and Human Services (EOHHS). Source code for all analyses is available from GitHub at <https://github.com/ripl-org/predict-medicaid-ed>.

2.1 Divertible ED spending

We identify ED visits based on revenue, procedure, and place of service codes in institutional and professional claims data, following the “Yale definition” [11]. We define divertible ED visits as those involving care that was potentially treatable in Primary Care (PC) using the “Minnesota algorithm” to assess the severity of procedures during the visit and whether they require facilities beyond those commonly available in an outpatient PC setting [12]. In contrast, visits are defined as requiring the Emergency Department if they result in admission, are assigned a high complexity or critical care management code or are assigned a moderate complexity management code but also include one of 120 procedure codes that were identified by ED physicians as indicating the appropriateness of the ED for treatment.

2.2 Predictive modeling

For each calendar year, we estimate the risk that each Medicaid enrollee with no divertible spending in that year will exhibit non-zero divertible ED spending in the subsequent year. Because spending in the subsequent year clusters around zero and on average only 6.8% of enrollees will have non-zero spending, we use an approach similar to the *tobit* to separately model the extensive and intensive margins [13]. We decompose the risk into the product of the predicted probability of non-zero spending (the extensive margin) and the predicted magnitude of spending conditional on non-zero spending (the intensive margin) with:

$$\log \left(\frac{\Pr(Y > 0)}{1 - \Pr(Y > 0)} \right) = \beta X + \varepsilon \quad (1)$$

$$Y|_{Y>0} = \gamma X|_{Y>0} + \zeta \quad (2)$$

where Y is the amount of cost emergence in the subsequent 12 months for an individual, X is a design matrix of variables known in the 12 preceding months from administrative records for an individual, and ε and ζ are error terms. Equation (1) models the extensive margin as the probability that the individual had non-zero cost

emergence. Equation (2) models the intensive margin of the amount of cost emergence conditional on non-zero cost emergence.

Following a decomposition like the one created by McDonald and Moffitt for the *tobit* [13], we can define the predicted values $E(Y)$ of cost emergence as the product of predictions $E(\text{Pr}) = e^{\beta X} / (1 - e^{\beta X})$ from equation (1) and predictions $E(Y^*) = \gamma X$ from equation (2), applied to all individuals in X . The coefficient for the i th variable in this combined model then follows from the product rule for partial derivatives:

$$\frac{\partial E(Y)}{\partial X_i} = E(Y^*) \frac{\partial E(\text{Pr})}{\partial X_i} + E(\text{Pr}) \frac{\partial E(Y^*)}{\partial X_i} = (\gamma \bar{X}_i) \beta + \left(\frac{e^{\beta \bar{X}_i}}{1 - e^{\beta \bar{X}_i}} \right) \gamma \quad (3)$$

where $\bar{X}_i$ is the mean value for the i th variable in the design matrix X .

To estimate these models, we fit both a basic linear regression over 11 expert-selected variables chosen to match those in a similar study by Tamang *et al.* [14] as well as regularized linear regressions over a much larger set of 1,963 variables using the LASSO [15]. We fit the models with training data from 2004-2012, tune them on validation data from 2012-2013, and report the results of the model predictions on testing data from 2013-2014. The testing data were withheld from analysis prior to preparation of this manuscript.

The design matrix X for the basic linear regression includes 11 variables: six age bins, sex, a missing indicator for sex, the count of distinct diseases and chronic conditions using the Clinical Classifications Software [16] and the Chronic Condition Indicators [17] classifications, and the count of distinct procedures. The additional 1,952 variables for the LASSO model come from demographics, Medicaid claims, social insurance programs, incarceration records, and employment history. The demographic variables include race/ethnicity, sex, age, marital status, body mass index, and neighborhood characteristics for the home Census block group (including median income and fraction of households below the federal poverty level). The variables from Medicaid claims including indicators for enrollment eligibility, payer codes, and managed care; payments for all institutional, professional, and pharmacy claims; the number of distinct drugs prescribed and daily supply; and indicators for diagnosis, procedure, and revenue codes. The variables for social insurance programs measure payments for the Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), Temporary Disability Insurance (TDI), and Unemployment Insurance (UI). The variables for incarceration history include indicators for charges and sentences (including for particular offenses of assault, sex offenses, drug offenses, homicide, public offenses, property offenses, or robbery) and for release or committal to a corrections facility. The employment history variables measure wages, hours, and number of employers; sector of employment; and unemployment rates both for the state of Rhode Island and nationally in the primary sector of employment. The full list of 1,963 feature names with descriptions is available from <https://github.com/ripl-org/predict-medicaid-ed/blob/main/features.csv>.

We estimate the basic regression model using a logistic regression for equation (1) and linear regression for equation (2) with the R package *glm* and with the subset of the design matrix X that contains the 11 variables described above. The coefficients for the extensive and intensive models are combined according to equation (3) with bootstrapped 95% confidence intervals and are available from <https://github.com/ripl-org/predict-medicaid-ed/blob/main/coef.basic.csv>.

We estimate the LASSO model with regularized logistic regressions for equation (1) and regularized regressions for equation (2) using the R package *glm*, which tunes the regularization with a path descent over 100 lambda values [18]. We select the lambda value that yields the best area under the curve of the cumulative

sum of actual cost emergence when ordering individuals from highest to lowest predicted cost emergence, as measured on the validation set. This objective function optimizes the model for the policy application of prioritizing individuals for outreach by predicted cost emergence.

Furthermore, we use a bootstrap procedure called the bootstrapped LASSO, or BOLASSO for short [19]. The advantage of this bootstrap procedure is that it consistently identifies a model in the presence of correlated variables. We run 100 bootstrap replicates of the LASSO implementations described above, resampling from X in each replicate, and select the variables occurring with non-zero coefficient in more than 90 of the 100 replicates. The selected variables are then pooled across both the extensive and intensive models to estimate a final pair of post-LASSO models for equations (1) and (2) with the R package *glm* on the BOLASSO-selected variables. The coefficients from the post-LASSO extensive and intensive models are combined according to equation (3) with bootstrapped 95% confidence intervals and are available from <https://github.com/ripl-org/predict-medicaid-ed/blob/main/coef.postlasso.csv>.

2.3 Limitations

Our findings are contingent on the accuracy of the Yale definition for identifying ED visits in claims data, and of the Minnesota algorithm for classifying which of those visits were potentially treatable in primary care. We chose the Minnesota algorithm based on its reported improvements over the NYU algorithm [20,21], which uses only the diagnoses codes from the visit instead of the final outcomes, such as the procedures performed and duration of the visit.

Another limitation of our study is that it is restricted to the population of Medicaid enrollees. Expanding the study to a data source such as an all-payers claims database would improve the representativeness of the population by including Medicare enrollees and the privately insured. However, a strength of restricting our study to the Medicaid population is our ability to securely join claims data to additional administrative records and establish connections with other government programs, which would be challenging with an all-payers claim database.

3 RESULTS

3.1 Annual spending

Figure 1 shows the change in divertible ED spending by year during our study period, which varied from \$6.0 to \$12.2 million. Our estimate of \$10.2 million in 2014 is similar to the estimate of \$18.0 million from a study by the Rhode Island Department of Health using a proprietary algorithm from 3M Inc. to identify divertible ED visits based on diagnoses codes [22]. We also examined mean reversion in our study population. The mean divertible ED spending in an enrollee's first year with such spending was \$338.05, compared to only \$50.41 in the second year and \$41.28 in the third year.

3.2 Characteristics of divertible spenders

Table 1 shows the baseline characteristics for enrollees in 2013 divided into different groups of policy interest. First, we examine the groups who had two or more or four or more ED visits in the prior year, which is representative of existing policies for identifying divertible ED use. For example, policies in Washington, Oregon, and Iowa target individuals with five or more [23], six or more [5], or twelve or more ED visits in the prior year

Figure 1. Divertible Emergency Department spending among Medicaid enrollees in Rhode Island from 2004 to 2014.

[24], respectively. Appendix Table 1 provides a comprehensive list of existing intervention programs for ED super-utilizer intervention across the US.

Next, we examine the groups of individuals who have non-zero or above median divertible ED spending in the subsequent year (2014), which are the targets for our predictive models. Individuals with multiple prior ED visits are more likely to be older, white, and married, and to have elevated prior year costs across institutional, professional, and pharmacy claims. Individuals with emerging non-zero divertible ED spending are more likely to be enrolled in a managed care plan and also have elevated prior year costs relative to the overall population, but not to the same degree as those with multiple ED visits. By construction, they also have higher potentially divertible ED spending in the subsequent year relative to both the overall population and the groups with multiple prior ED visits, which is our motivation for using predictive modeling to identify them ahead of time.

3.3 Potential cost savings

In Figure 2, we compare the “cost-capture” for each model by sorting individuals from highest to lowest predicted risk, and then calculating the total realized divertible ED spending for the top 1, 2, and 5 percent of individuals, who would be potential targets for outreach and case management. We also compare these savings to the two or more and four or more ED visit policies. The cost-capture is highest using our LASSO models, which capture 4.7, 8.3, and 16.8 percent of divertible ED spending with outreach to 1, 2, and 5 percent of Medicaid enrollees, respectively. With an estimated \$10.2 million in divertible ED spending in Rhode Island, this amounts to capturing \$459 thousand, \$836 thousand, and \$1.7 million, respectively, in divertible ED spending. The

Table 1. Baseline characteristics of Rhode Island Medicaid enrollees in 2013.

		All enrollees	2+ ED visits in prior year	4+ ED visits in prior year	Non-zero divertible ED spending in subsequent year	Above median divertible ED spending in subsequent year
N		234,421	37,611	14,533	16,381	8,159
Age	<18	41.3%	25.5%	12.4%	50.9%	54.7%
	18-60	42.6%	56.4%	62.5%	44.5%	42.5%
	>60	15.1%	17.9%	25.0%	4.6%	2.8%
Female		56.3%	63.4%	64.9%	58.8%	60.6%
Race/ Ethnicity	White	47.4%	53.0%	57.8%	42.8%	39.6%
	African American	11.3%	12.7%	12.5%	14.8%	14.8%
	Hispanic	17.5%	14.8%	10.8%	23.8%	26.0%
Subsequent year means:						
Annual divertible ED spending		\$36.01	\$81.61	\$99.44	\$515.28	\$842.74
Prior year means:						
Annual divertible ED spending		\$37.53	\$167.37	\$225.05	\$120.98	\$152.06
Received SNAP		67.8%	80.1%	79.1%	83.1%	83.6%
Enrolled in managed care		58.6%	52.2%	34.8%	72.8%	82.9%
Enrolled in premium payments		25.6%	21.5%	17.2%	30.8%	29.9%
Monthly institutional spending		\$853.40	\$3,075.83	\$5,026.55	\$1,362.56	\$1,483.09
Monthly professional spending		\$519.08	\$1,288.36	\$1,961.00	\$698.17	\$728.03
Monthly pharmacy spending		\$56.95	\$121.73	\$168.54	\$73.85	\$74.83
Had an outpatient office exam		64.6%	89.9%	92.8%	79.1%	82.5%
Had a preventative, dental, or vision exam		60.4%	71.4%	69.6%	72.1%	75.4%
Distinct diseases		7.0	17.1	22.9	10.6	11.0
Distinct chronic conditions		3.1	8.3	12.8	3.9	3.7
Distinct procedures		15.2	36.4	49.1	20.7	21.2
Had a geocoded home address		92.6%	94.1%	94.3%	93.8%	93.4%
Home blockgroup's median income		\$50,133	\$47,197	\$47,032	\$45,245	\$44,330
Home blockgroup's fraction below the poverty line		17.1%	18.6%	18.4%	20.2%	20.8%

predictions from the LASSO models have no identification costs in the prior year, unlike the identification methods based on prior ED visits.

Predictors of cost emergence

Figure 3 summarizes the full regression output from the LASSO models. The variables selected by the LASSO come from the Medicaid claims, social benefit and insurance program, employment, and demographic data. The top predictors that increase predicted divertible spending include enrollment in a managed care plan, age between 18 to 23, and receipt of prior diagnostic services. The top predictor that decreases predicted divertible spending is enrollment in a premium payment plan. Variables for social insurance programs and employment history also contribute to the prediction, but with magnitudes closer to zero.

Figure 2. Divertible Emergency Department spending potentially captured in 2014 (subsequent year) by outreach methods evaluated on data in 2013 (prior year).

3.4 Connections with other programs

Individuals with emerging non-zero divertible ED spending participate in other non-Medicaid programs, which indicates connections to existing case management: 83.1% of these individuals were enrolled in the Supplemental Nutrition Assistance Program (SNAP) during the year prior to cost-emergence, and 22.5% in the Temporary Assistance for Needy Families (TANF) program (see Table 1). Enrollment in both of these programs requires regular contact with an existing caseworker or case management system. In Rhode Island, continued enrollment in SNAP requires recertification every six months, which includes a telephone or in-person interview with a caseworker [25]. The application process for TANF requires an in-person screening interview and

Figure 3. Summary of coefficients from the LASSO regression models for emerging divertible ED costs.

development of an employment plan, which is then monitored on a biweekly basis for the duration of benefits, and continued benefits require filing an interim report by mail every six months [26].

4 DISCUSSION

Approaches that target individuals after a fixed number of ED visits rely on observing a period of high spending to identify those individuals. Waiting to accrue costs means that programs are paying to identify individuals for intervention. With an accurate enough predictive model, those costs could be saved without relying on accrued costs for targeting.

Our predictive modeling approach provides a proactive framework for identifying individuals for intervention and uncovers existing connections to social services and case management. These existing connections could be efficient avenues for providing education on and outreach for ED alternatives, such as urgent care retail clinics and community primary care centers, which have been shown to decrease ED usage and be more cost effective for preventable and minor acute conditions [2,27]. Additionally, expanded access to medical care can increase the incidence of ED visits, as shown in the Oregon Health Insurance Experiment [28]. This further reinforces the need for smart policy that optimally allocates limited ED resources.

If our predictive model extends to other states with similar efficacy, the potential savings could be much larger than what we can estimate in Rhode Island. For example, in FY 2014, total Medicaid spending was \$61.6 billion in California and \$31.2 billion in Texas [29]. Under the assumption that the ratio of divertible ED spending to total Medicaid spending is the same across states, divertible ED spending would have totaled approximately \$246.4 million in California and \$124.8 million in Texas. A similarly effective cost-capture as the Rhode Island model, capturing 4.7, 8.3, and 16.8 percent of divertible ED spending with outreach to 1, 2, and 5 percent of Medicaid enrollees, respectively, would amount to capturing \$11.1, \$20.2, and \$41.1 million in divertible ED costs in California and \$5.6, \$10.2, and \$20.8 million in divertible ED costs in Texas.

5 CONCLUSION

Our data-driven approach demonstrates how divertible Medicaid ED use can be predicted instead of waiting for costs to accrue. By utilizing their own data, policymakers may be able to save unnecessary costs and better serve the health needs of program enrollees. Moreover, interconnected administrative data points to existing outreach programs that could be used instead of starting new outreach. We add to a literature showing the value of administrative data and machine learning in improving medical care [30].

ACKNOWLEDGMENTS

We thank the Carla and David Crane Foundation of the Jewish Community Federation and Endowment Fund, the Laura and John Arnold Foundation, and the Smith Richardson Foundation for financial support. We thank the Office of the Governor of Rhode Island and the Rhode Island Executive Office of Health and Human Services for supporting fact-based policymaking.

REFERENCES

- [1] John Billings and Maria C. Raven. 2013. Dispelling An Urban Legend: Frequent Emergency Department Users Have Substantial Burden Of Disease. *Health Affairs* 32, 12 (Dec. 2013), 2099–2108. DOI:<https://doi.org/10.1377/hlthaff.2012.1276>
- [2] Diane Alexander, Janet Currie, and Molly Schnell. 2019. Check up before you check out: Retail clinics and emergency room use. *Journal of Public Economics* 178 (Oct. 2019), 104050. DOI:<https://doi.org/10.1016/j.jpubeco.2019.104050>
- [3] Cindy Mann. 2014. Reducing Nonurgent Use of Emergency Departments and Improving Appropriate Care in Appropriate Settings. Retrieved from https://www.hhs.gov/guidance/sites/default/files/hhs-guidance-documents/CIB-01-16-14_11.pdf.
- [4] Congressional Budget Office. 2017. Federal Spending for Means-Tested Programs, 2007 to 2027. Retrieved April 17, 2021 from <https://www.cbo.gov/publication/52405>.
- [5] Cindy Mann. 2013. Targeting Medicaid Super-Utilizers to Decrease Costs and Improve Quality. Retrieved April 17, 2021 from <https://www.medicaid.gov/federal-policy-guidance/downloads/cib-07-24-2013.pdf>.
- [6] Fabrice Althaus, Sophie Paroz, Olivier Hugli, William A. Ghali, Jean-Bernard Daeppen, Isabelle Peytremann-Bridevaux, and Patrick Bodenmann. 2011. Effectiveness of Interventions Targeting Frequent Users of Emergency Departments: A Systematic Review. *Annals of Emergency Medicine* 58, 1 (Jul. 2011), 41–52. DOI:<https://doi.org/10.1016/j.annemergmed.2011.03.007>
- [7] Tracy L. Johnson, Deborah J. Rinehart, Josh Durfee, Daniel Brewer, Holly Batal, Joshua Blum, Carlos I. Oronce, Paul Melinkovich, and Patricia Gabow. 2015. For Many Patients Who Use Large Amounts Of Health Care Services, The Need Is Intense Yet Temporary. *Health*

- Affairs* 34, 8 (Aug. 2015), 1312–1319. DOI:<https://doi.org/10.1377/hlthaff.2014.1186>
- [8] Anya Rader Wallack. 2016. Reinventing Medicaid – Expanding Access, Improving Care, Controlling Costs. Retrieved April 17, 2021 from <https://governor.ri.gov/documents/press/022016MedicaidM.pdf>.
 - [9] Justine S. Hastings. 2019. *Fact-Based Policy: How Do State and Local Governments Accomplish It?* Policy Proposal 2019-01. The Hamilton Project, Washington, DC. Retrieved from https://www.hamiltonproject.org/assets/files/Hastings_PP_web_20190128.pdf.
 - [10] Justine S. Hastings, Mark Howison, Ted Lawless, John Ucles, and Preston White. 2019. Unlocking Data to Improve Public Policy. *Communications of the ACM* 62, 10 (Oct. 2019), 48–53. DOI:<https://doi.org/10.1145/3335150>
 - [11] Arjun K. Venkatesh, Hao Mei, Keith E. Kocher, Michael Granovsky, Ziad Obermeyer, Erica S. Spatz, Craig Rothenberg, Harlan M. Krumholz, and Zhenqi Lin. 2017. Identification of Emergency Department Visits in Medicare Administrative Claims: Approaches and Implications. *Academic Emergency Medicine* 24, 4 (Apr. 2017), 422–431. DOI:<https://doi.org/10.1111/acem.13140>
 - [12] Molly Moore Jeffery, M. Fernanda Bellolio, Julian Wolfson, Jean M. Abraham, Bryan E. Dowd, and Robert L. Kane. 2016. Validation of an Algorithm to Determine the Primary Care Treatability of Emergency Department Visits. *BMJ Open* 6, 8 (Jul. 2016), e011739. DOI:<https://doi.org/10.1136/bmjopen-2016-011739>
 - [13] John F. McDonald and Robert A. Moffitt. 1980. The Uses of Tobit Analysis. *The Review of Economics and Statistics* 62, 2 (May 1980), 318–321. DOI:<https://doi.org/10.2307/1924766>
 - [14] Suzanne Tamang, Arnold Milstein, Henrik Toft Sørensen, Lars Pedersen, Lester Mackey, Jean-Raymond Betterton, Lucas Janson, and Nigam Shah. 2017. Predicting Patient ‘Cost Blooms’ in Denmark: A Longitudinal Population-Based Study. *BMJ Open* 7, 1 (Jan. 2017), e011580. DOI:<https://doi.org/10.1136/bmjopen-2016-011580>
 - [15] Robert Tibshirani. 1996. Regression Shrinkage and Selection via the Lasso. *J. R. Statist. Soc. B* 58, 1 (1996), 267–288.
 - [16] Healthcare Cost and Utilization Project. 2017. Clinical Classifications Software (CCS) for ICD-9-CM. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved April 17, 2021 from <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>.
 - [17] Healthcare Cost and Utilization Project. 2016. Chronic Condition Indicator (CCI) for ICD-9-CM. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved April 17, 2021 from <https://www.hcup-us.ahrq.gov/toolssoftware/chronic/chronic.jsp>.
 - [18] Matt Taddy. 2017. One-Step Estimator Paths for Concave Regularization. *Journal of Computational and Graphical Statistics* 26, 3 (Jul. 2017), 525–536. DOI:<https://doi.org/10.1080/10618600.2016.1211532>
 - [19] Francis R. Bach. 2008. Bolasso: model consistent Lasso estimation through the bootstrap. In *Proceedings of the 25th International Conference on Machine Learning (ICML '08)*, July 5 - 9, 2008, Helsinki, Finland. ACM Inc., New York, NY, 33–40. DOI:<https://doi.org/10.1145/1390156.1390161>
 - [20] Dustin W. Ballard, Mary Price, Vicki Fung, Richard Brand, Mary E. Reed, Bruce Fireman, Joseph P. Newhouse, Joseph V. Selby, and John Hsu. 2010. Validation of an Algorithm for Categorizing the Severity of Hospital Emergency Department Visits. *Medical Care* 48, 1 (Jan. 2010), 58–63. DOI:<https://doi.org/10.1097/MLR.0b013e3181bd49ad>
 - [21] Maria C. Raven, Robert A. Lowe, Judith Maselli, and Renee Y. Hsia. 2013. Comparison of Presenting Complaint vs Discharge Diagnosis for Identifying ‘Nonemergency’ Emergency Department Visits. *JAMA* 309, 11 (Mar. 2013), 1145–1153. DOI:<https://doi.org/10.1001/jama.2013.1948>
 - [22] Rhode Island Department of Health. 2015. Potentially Preventable Emergency Room Visits. Retrieved April 17, 2021 from <https://health.ri.gov/data/potentiallypreventableemergencyroomvisits/>.
 - [23] Jesse Pines, Nathan Schlicher, Elise Presser, Meaghan George, Mark McClellan. 2015. Washington State Medicaid: Implementation and Impact of “ER is for Emergencies” Program. Center for Health Policy at Brookings, The Brookings Institution, Washington, DC. Retrieved from <https://www.brookings.edu/wp-content/uploads/2016/07/050415EmerMedCaseStudyWash.pdf>.
 - [24] A. Clinton MacKinney and Christopher Stamy. 2021. Serving High Need/High Cost Patients in the Emergency Department. Retrieved from <https://cph.uiowa.edu/ruralhealthvalue/files/Super%20Utilizers.pdf>.
 - [25] Rhode Island Department of Human Services. 2020. SNAP Recertification Form. Retrieved April 17, 2021 from <http://www.dhs.ri.gov/Programs/1SNAP-2RecertificationFormRev03-20.pdf>.
 - [26] Rhode Island Department of Human Services. 2020. Rhode Island Works Program Rules and Regulations. Retrieved April 17, 2021 from <http://www.dhs.ri.gov/Regulations/218-RICR-20-00-2RIWorksProgramFeb2020.pdf>.
 - [27] Donald Wesson, Heather Kitzman, Kenneth H. Halloran, and Kristen Tecson. 2018. Innovative Population Health Model Associated With Reduced Emergency Department Use And Inpatient Hospitalizations. *Health Affairs* 37, 4 (Apr. 2018), 543–550. DOI:<https://doi.org/10.1377/hlthaff.2017.1099>
 - [28] Sarah L. Taubman, Heidi L. Allen, Bill J. Wright, Katherine Baicker, and Amy N. Finkelstein. 2014. Medicaid Increases Emergency-Department Use: Evidence from Oregon’s Health Insurance Experiment. *Science* 343, 6168 (Jan. 2014), 263–268. DOI:<https://doi.org/10.1126/science.1246183>
 - [29] Kaiser Family Foundation. 2014. Medicaid Spending by Enrollment Group. Retrieved April 17, 2021 from <https://www.kff.org/medicaid/state-indicator/medicaid-spending-by-enrollment-group/>.
 - [30] Janet Currie and W. Bentley MacLeod. 2016. Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians. *Journal of Labor Economics* 35, 1 (Jun. 2016), 1–43. DOI:<https://doi.org/10.1086/687848>
 - [31] Frederick Isasi. 2015. *Using Data to Better Serve the Most Complex Patients: Highlights from NGA’s Intensive Work with Seven States*. National Governors Association Center for Best Practices, Washington, DC. Retrieved April 17, 2021 from <https://www.nga.org/wp-content/uploads/2019/08/1509UsingDataBetterServeComplexPatients.pdf>.

- [32] Karen W. Linkins, Jennifer J. Brya, and Daniel W. Chandler. 2008. Frequent Users of Health Services Initiative: Final Evaluation Report. The Lewin Group, Falls Church, VA. Retrieved April 17, 2021 from <https://www.chcf.org/wp-content/uploads/2017/12/PDF-FUHSIEvaluationReport.pdf>.
- [33] Robert L. Okin, Alicia Boccellari, Francisca Azocar, Martha Shumway, Kathy O'Brien, Alan Gelb, Michael Kohn, Phyllis Harding, and Christine Wachsmuth. 2000. The effects of clinical case management on hospital service use among ED frequent users. *American Journal of Emergency Medicine* 18, 5 (Sep. 2020), 603–608. DOI: <https://doi.org/10.1053/ajem.2000.9292>
- [34] Clemens S. Hong, Allison L. Siegel, and Timothy G. Ferris. 2014. *Caring for High-Need, High-Cost Patients: What Makes for a Successful Care Management Program?* The Commonwealth Fund Issue Briefs, Pub. 1764, Vol. 19. The Commonwealth Fund, New York, NY. Retrieved April 17, 2021 from <https://www.commonwealthfund.org/publications/issue-briefs/2014/aug/caring-high-need-high-cost-patients-what-makes-successful-care>.
- [35] Michael Hochman. 2016. Improvement happens – learning to better care for ‘super-utilizers’ at Denver Health: an interview with Holly Batel. *Journal of General Internal Medicine* 31, 10 (May 2016), 1246–1253. DOI: <https://doi.org/10.1007/s11606-016-3838-4>
- [36] Connecticut General Assembly. 2014. Hospital Emergency Department Use and Its Impact on the State Medicaid Budget. Retrieved from <https://www.cga.ct.gov/pri/docs/2013/HEDU%20FINAL%20PRI%20HOSP%20ED%20USE%20REPORT%20FOR%20PRINTING.pdf>.
- [37] Kentucky Health Information Exchange. 2014. KHIE clinical notifications. *Kentucky Health Information Exchange Newsletter* 1, 10 (Aug 2014), 3. Retrieved December 5, 2017 from https://khie.ky.gov/tech/Documents/KHIE%20Newsletter_issue10%20August%202014.pdf
- [38] New York State Department of Health. 2017. DSRIP – Medicaid Accelerated eXchange (MAX) Series Program. Retrieved April 17, 2021 from https://www.health.ny.gov/health_care/medicaid/redesign/dsrip/pps_workshops/docs/2017-01_imp_care.pdf.

APPENDICES

Appendix Table 1. Existing intervention programs for Emergency Department super-utilizers.

State	Program Name	Super-utilizer Definition	Ref.
Alaska	Coordinated Care Initiative	5+ ED visits in 18 months	[31]
California	Frequent Users of Health Services Initiative: Alameda	10+ ED visits/year or 4+ ED visits/year for 2 consecutive years	[32]
	Frequent Users of Health Services Initiative: LA	5+ ED visits/year and 2+ of: mental illness, substance use with or without co-occurring physical illness, homelessness, uninsured or underinsured, income <200% of FPL	[32]
	Frequent Users of Health Services Initiative: Sacramento	4+ ED visits/year, aged 18-64, and any of: history of mental illness or drug/alcohol abuse, no primary care provider, underinsured, uninsured	[32]
	Frequent Users of Health Services Initiative: Santa Clara	8+ ED visits/year, aged 18+, county resident, uninsured or Medicaid	[32]
	Frequent Users of Health Services Initiative: Santa Cruz	5+ ED visits/year, with at least 1 in the last 6 months with documented primary or secondary diagnoses of psychosocial disorders, mental illness or substance abuse disorders	[32]
	Frequent Users of Health Services Initiative: Tulare	8+ ED visits/year or 5+ visits in last 6 months; aged 18+	[32]
	San Francisco General Hospital Pilot Intervention	5+ ED visits/year	[33]
	Sutter Care Coordination Program	Any of: 2+ ED visits/year; unplanned readmission within 30 days; 2+ admissions/year; 7+ medications; diagnosis of CHF, COPD, or pneumonia; 3+ chronic conditions	[34]
	Humboldt County Super Utilizer Program	2+ ED visits/month for 6 months	[24]
Colorado	Denver Health Medical Plan	3+ ED visits/year	[35]
	NGA's Developing State-Level Capacity to Support Super-Utilizers Policy Academy	6+ ED visits/year and 30+ prescription drugs/year	[31]
Connecticut		10+ ED visits/year	[36]
Oregon	CareOregon	6+ ED visits./year	[5]
Iowa	ED Consistent Care Plan	12+ ED visits/year	[24]
Kentucky	KY ER SMART	10+ ED visits/year or 3+ acute hospital admissions/year	[37]
Maine	Maine Community Care Teams	3+ ED visits in 6 months and 5+ ED visits/year	[5]
Michigan	Spectrum Health	10+ ED visits/year or 3+ inpatient admissions/year with behavioral health conditions and social issues	[5,24]
Minnesota	Hennepin Health	5+ ED visits or medical detox stays in 6 months, 3+ inpatient admissions/year, 2+ psychiatric admissions	[5]
New Jersey	Camden Coalition	2+ chronic disease-related admissions in 6 months	[34]
New York	Medicaid Accelerated eXchange (MAX) Series Program	10+ ED visits/year for large/medium facilities or 8+ ED visits/year for smaller facilities	[38]
North Carolina	Community Care of North Carolina Priority Patients Program and Transitional Care Program	Estimated spending (3M Inc. model) using claims data to target those with the largest difference between actual and expected spending for potentially preventable inpatient	[5]

		admissions or ED visits. If claims history is insufficient, the program targets beneficiaries with 1+ inpatient admission in 6 months; 2+ ED visits in 6 months; and any of: congestive heart failure, diabetes, ischemic vascular disease, asthma, or chronic obstructive pulmonary disease.	
Vermont	Vermont Chronic Care Initiative	Top 5% of highest utilizers or demonstrate high utilization patterns including multiple ED visits and inpatient admissions	[5]
Washington	ER is for Emergencies	5+ ER visits or transports in the past year	[23]
Wisconsin	Milwaukee Health Care Partnership: Care Coordination	3+ visits in 6 months and annual Medicaid costs of \$100,000+	[31]