Received: 7 August 2020 | Revised: 1 December 2020 | Accepted: 11 December 2020

DOI: 10.1002/emp2.12346

JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE
Open Access

WILEY

ORIGINAL RESEARCH

Health Policy

Identifying subgroups and risk among frequent emergency department users in British Columbia

Jessica Moe MD, MSc[1] | Fiona O'Sullivan PhD[2] | Margaret J. McGregor MD, MHSc[3] | Michael J. Schull MD, MSc[4] | Kathryn Dong MD, MSc[5] | Brian R. Holroyd MD, MBA[6] | Eric Grafstein MD[7] | Corinne M. Hohl MD, MHSc[8] | Johanna Trimble[9] | Kimberlyn M. McGrail PhD[10]

[1] Department of Emergency Medicine, University of British Columbia, Department of Emergency Medicine, Vancouver General Hospital, Vancouver, British Columbia, Canada

[2] Department of Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada

[3] Department of Family Practice, University of British Columbia, Vancouver, British Columbia, Canada

[4] Institute for Clinical Evaluative Sciences, Department of Medicine, University of Toronto, Toronto, Ontario, Canada

[5] Department of Emergency Medicine, University of Alberta, Edmonton, Alberta, Canada

[6] Department of Emergency Medicine, Emergency Strategic Clinical Networ, Alberta Health Services, University of Alberta, Edmonton, Alberta, Canada

[7] Department of Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada

[8] Department of Emergency Medicine, University of British Columbia, Department of Emergency Medicine, Vancouver General Hospital, Vancouver, British Columbia, Canada

[9] Patients for Patient Safety Canada, Roberts Creek, Vancouver, British Columbia, Canada

[10] Population Data BC, School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada

Correspondence
Dr. Jessica Moe, MD, MSc, Department of Emergency Medicine, Vancouver General Hospital, 920 West 10th Avenue, Vancouver, British Columbia V5Z 1M9, Canada.
Email: jessica.moe@ubc.ca

Funding and support: This study received funding from the Vancouver Coastal Health Research Institute (Innovation and Translational Research Award) and Mentored Clinician Scientist Salary Support Award).

Abstract

Objective: Frequent emergency department (ED) users are heterogeneous. We aimed to identify subgroups and assess their mortality.

Methods: We identified patients ≥18 years with ≥1 ED visit in British Columbia from April 1, 2012 to March 31, 2015, and linked to hospitalization, physician billing, prescription, and mortality data. Frequent users were the top 10% of patients by ED visits. We employed cluster analysis to identify frequent user subgroups. We assessed 365-day mortality using Kaplan-Meier curves and conducted Cox regressions to assess mortality risk factors within subgroups.

Results: We identified 4 subgroups. Subgroup 1 ("Elderly") had median age 77 years (interquartile range [IQR]: 66–85), 5 visits/year (IQR: 4–6), median 8 prescription medications (IQR: 5–11), and 24.7% mortality. Subgroup 2 ("Mental Health and Alcohol Use") had median age 48 years (IQR: 34–61), 13 visits/year (IQR: 10–16), and 12.3% mortality. They made a median 31 general practitioner visits (IQR: 19–51); however, only 23.7% received a majority of services from 1 primary care physician. Subgroup

Supervising Editor: Chadd K. Kraus, DO, DrPH

3 ("Young Mental Health") had median age 39 years (IQR: 28–51), 5 visits/year (IQR: 4–6), and 2.2% mortality. Subgroup 4 ("Short-term") had median age 50 years (IQR: 34–65), 4 visits/year (IQR: 4–5) regularly spaced over a short term, and 1.4% mortality. Male sex (all subgroups), long-term care ("Mental Health and Alcohol Use;" "Young Mental Health"), and rural residence ("Elderly" in long-term care; "Young Mental Health") were associated with increased mortality.

Conclusions: Our results identify frequent user subgroups with varying mortality. Future research should explore subgroups' unmet needs and tailor interventions toward them.

KEYWORDS
cluster analysis, emergency medicine, epidemiology, frequent users, health policy, high service users

1 | INTRODUCTION

1.1 | Background

High health care users account for disproportionate costs: experience in the United States and internationally indicates that the top 10% of patients account for 68% of health care spending.[1,2] "Super-users" of health care, and particularly of emergency departments, are therefore a cost-containment priority.[2–4] Frequent ED users comprise 4.5%-8% of ED patients, yet account for 21%–30% of visits.[5,6] They incur disproportionately high costs because of high ED and other health service use.[7,8] They also make higher acuity visits and are admitted and die more often than non-frequent users.[9–13]

Despite the recognized impact of frequent ED use, there is a knowledge gap in how best to address this issue. Interventions described to date (eg, case management, care plans) are rarely tailored to patient-specific needs.[14–17] Greater primary care linkages have been proposed; however, over 93% of North American frequent users are already attached.[18–23] Interventions liaising frequent users with primary care physicians at discharge have had mixed effects on ED use.[24,25]

Existing literature indicates that frequent users are heterogeneous. They have a bimodal age distribution with peaks at 25–44 and >65 years.[5] They make varied ED presentations including for pain,[26] mental health,[27–29] chronic medical illness,[18,29–31] and substance use.[30–34] Costs may increase with combinations of characteristics: for instance, homelessness, multiple physical conditions, mental illness, and substance use.[7] Characteristics and risk also vary by degree of use. Previous analyses suggest that frequent users with ≤17 visits/year may be older, have more chronic medical illness, and be admitted and die more often than extremely frequent users (≥18 visits/year), who may be younger with more substance use and mental illness-related visits.[13,24,35] Clinical heterogeneity suggests different care needs; however, these have been incompletely explored to date.

Cluster analysis is a method that has been broadly applied to identify distinct meaningful groups, including patients at risk for disease, health care practitioners with common characteristics, and patterns at a cellular level.[36–39] Cluster analysis has not been applied to the question of identifying frequent user subgroups with distinct care needs.

1.2 | Importance

Frequent ED users are high-cost, high-risk, and heterogeneous patients. Plans to manage these patients rarely address root causes of frequent use. Interventions that improve health and reduce health care use could yield considerable cost savings and importantly improve patient care and outcomes. There is a need to develop a nuanced and multidimensional understanding of heterogeneous frequent user subgroups as a first step in planning targeted and effective interventions.

1.3 | Goals of this investigation

The study objective was, first, to identify and characterize frequent ED user subgroups based on demographic, clinical, and health care usage patterns within a comprehensive linked administrative database in British Columbia (BC), Canada. Our second objective was to assess risk factors for 365-day mortality among these frequent user subgroups.

2 | METHODS

2.1 | Study design, setting and participants

This was a retrospective administrative database study capturing patients who visited an ED in BC between April 1, 2012 and March 31, 2015. We split the data into 4 fiscal year groupings. We created a cohort by identifying all patients aged ≥18 years who made ≥1 ED visit based on ED visit records within the National Ambulatory Care Reporting System (NACRS) database.[40]

2.2 | Data sources

Our study database linked patient-level data for each patient in our cohort (NACRS) to hospitalization data within the Discharge Abstract Database (DAD),[41] physician billing data in the Medical Services Plan (MSP),[42] provincial prescription records (PharmaNet),[43] and mortality data (Vital Statistics).[44] Population Data BC (PopData) housed and linked databases using personal health number, age, sex, and postal code. PopData, which originated as the provincial Ministry of Health-funded BC Linked Health Data Set, is a multiuniversity resource supporting linkage and access to individual-level, de-identified data for research. PopData undertakes validation and quality assurance/control, and employs rigorous, standardized linkage procedures, using a combination of deterministic and probabilistic approaches.[45,46] Integrated data checks within PopData's linkage procedures minimize false positives and negatives, and the use of a linkage coordinating file eliminates the risk of propagation error.[46] All patients received an anonymized study identification number that was consistent across the data.[45] The University of British Columbia Clinical Research Ethics Board approved this study.

2.3 | Study definitions and variables

2.3.1 | Frequent users

We defined frequent users as adult patients ≥18 years who were within the top 10% of ED use when all patients who visited an ED were ordered by number of visits made within each fiscal year, consistent with a definition established by the Canadian Institute for Health Information (Supplementary Figure 1).[47]

To determine visit counts, we first attempted to identify and remove scheduled revisits. Clinicians on our team determined that the majority of scheduled ED revisits in BC are made for intravenous antibiotics for cellulitis. We therefore developed an algorithm to identify ED visit strings occurring within 48 hours of one another with an initial diagnosis of cellulitis. To verify the accuracy of our approach, we examined diagnoses associated with 48-hour repeat ED visits: cellulitis was the most common specified diagnosis (12.1%), followed in frequency by abdominal pain (4.0%), follow-up examination (3.0%), and other medical care (2.3%).

We selected the top 10% of patients based on their ED visit count for each fiscal year, after presumed scheduled revisits had been removed. We present a flow chart outlining study inclusion in Supplementary Figure S2. In a sensitivity analysis, we also identified and characterized the top 5% and 1% of patients within our frequent user cohort.

2.3.2 | Neighborhood income quintile and rural/urban residence

PopData determined neighborhood income quintile using a postal code-based algorithm and rural residence based on "0" in the second position of the postal code. Both are standardized methods employed by Statistics Canada.[48]

The Bottom Line

Over a 3-year period, frequent emergency department users in British Columbia represented 10% of ED visits, and could be classified into 4 frequent user subgroups with unique demographic, clinical, and visit pattern profiles. Frequent users had higher mortality (11.3%) compared to non-frequent users (3.8%). Tailored interventions should be aimed at addressing the unique needs of these subgroups.

2.3.3 | Index visit

We chose April 1, 2013 to March 31, 2014 as our study year, as this was the most recent year of available data that would provide us with a complete baseline period and follow-up period for all included patients. We defined frequent users' final ED visit between April 1, 2013 and March 31, 2014 as each individual's index visit. We chose patients' final visits as their index visits to mirror the clinical situation in which a treating emergency physician might use a frequent user patient's visit history in the preceding year to assess their present risk, including reasons for past presentations and ED visit patterns. We examined baseline demographic and clinical characteristics for each patient in a preindex period that was unique for each individual patient (the 365-day period before the index visit). We determined all independent variables for cluster and survival analyses during this preindex period. We examined mortality within the 365-day period following each individual's index visit. We structured our analysis such that it would mirror information that might be available to an emergency physician when caring for a patient in the ED and reflect outcomes of interest to emergency physicians (eg, within a time frame that could be influenced by interventions initiated during that visit). Although the exact 365-day preindex period used to determine baseline variables and the 365-day postindex period to determine mortality were unique for each patient, data available for our study ranged from April 1, 2012 to March 31, 2015.

2.3.4 | Majority source of care

We used the general practitioner specialty code to identify primary care physicians and visits using MSP data. We then created a Majority Source of Care variable that described continuity of primary care. Majority Source of Care is a binary variable that identified whether ≥50% of patients' services were provided by a single general practitioner, among patients who received ≥3 family doctor services in 1 year. This variable is a recognized standard for measuring primary care continuity.[49–51]

2.3.5 | Long-term care residence

We used a previously developed algorithm based on MSP billing codes to identify patients' residence in a long-term care (LTC) facility.[52] We identified LTC residence if the patient received the following fee item codes attributable to a general practitioner: "Visit nursing home 1 or multiple patients," "Nursing home visit," "LTC institution visit," and "LTC facility visit."

2.3.6 | Regularity index

The regularity index is a previously developed measure quantifying the spacing between visits.[53] This index describes whether patients made visits that were regularly spaced relative to one another or visited sporadically. We calculated the regularity index using the following equation: $^{1}/_{1+\text{variance of visits}}$, with visit variance calculated using the time in days between each visit. We measured this over a 365-day time horizon preceding each patient's index visit. This generated a score from 0 to 1, where indices closer to 1 reflected more regular spacing. For example, 2 people who made 12 ED visits will have different "regularity" if one makes those visits randomly (irregular) and the other makes 1 visit every month throughout the year (regular).

2.3.7 | Measure of complexity: number of prescription medications

We summed the number of distinct generic medication names listed in provincial PharmaNet records. We used number of prescription medications as a measure of patient complexity.[54]

2.3.8 | Measure of complexity: number of diagnostic categories

We assessed the number of diagnostic categories (*International Classification of Diseases, 10th edition* [ICD-10] chapters)[55] attributable to ED diagnoses for which frequent users' presented as a measure of complexity. This method is a previously validated measure of the complexity of patients' health needs.[56,57]

2.4 | Primary data analysis

2.4.1 | Cluster analysis and clustering variables

We employed cluster analysis to explore whether subgroups existed within our frequent user cohort. Cluster analysis is a methodology to identify de novo patterns and to organize data into subgroups that maximize within-subgroup similarities and between-subgroup differences.[58] We employed an iterative process to determine the number and nature of included variables. We integrated information from our team's clinical experience to identify meaningful and distinct subgroups, as is commonly done in cluster analysis.[58] Furthermore, we applied information from a comprehensive characterization of our cohort using the same linked provincial data set to inform variable selection (in submission). For instance, our parallel analysis identified that frequent users had a high prevalence of mental illness, had greater numbers of prescription medications, and made more primary care visits than non-frequent users.[59] We included variables in our clustering algorithm pertaining to patients' ED visit patterns and clinical characteristics to reflect information available to emergency physicians at the point of- care to guide clinical decisionmaking. We excluded patients with missing information from this analysis.

We included the following 10 variables, assessed during the 365-day preindex period, in our clustering algorithm:

1. Total number of ED visits.
2. Number of months in the year that the patient visited an ED, a measure of visit spread.
3. Regularity index.[53]
4. Number of ICD-10 ED discharge diagnosis pertaining to mental and behavioral disorders (Chapter V). Substance use disorders were included in this category.
5. Number of ICD-10 ED discharge diagnosis pertaining to circulatory system diseases (Chapter IX).
6. Number of ICD-10 ED discharge diagnosis pertaining respiratory system diseases (Chapter X).
7. Number of ICD-10 ED discharge diagnosis pertaining digestive system diseases (Chapter XI).
8. Number of distinct ICD-10 ED discharge diagnostic chapters assigned to each patient.
9. Number of different medications prescribed.
10. Age at time of index visit.

We scaled all variables such that they all had a mean of zero and a standard deviation of 1 to prevent any single variable from overinfluencing the clustering algorithm. We employed the *Kmeans* method to identify variables that best differentiated our subgroups, as this approach was most suited to the size of our data set and number of variables we included in our algorithm.[60,61] We evaluated the optimal subgroup number using the elbow method and pseudo-F test (Supplementary Figures 3–4 and Supplementary Table S1).[62–64] Using these methods, we determined that 4 subgroups were the optimal number of groupings to describe our data.

2.5 | Outcome measures and survival analysis

Our primary outcome was 365-day mortality among frequent user subgroups. We present Kaplan-Meier curves as graphical representations of 365-day mortality stratified by subgroups. To determine patient characteristics associated with a higher mortality hazard, we calculated adjusted hazard ratios for mortality using multivariable Cox proportional hazard regressions stratified by subgroups. We

used a backwards stepwise method to select which covariates to include. We considered a combination of clinical importance, optimizing the Bayesian Information Criterion number and minimizing collinearity to create our final models (Supplementary Tables S2–3). We calculated Schoenfeld residuals to test the proportional hazards assumption for variables included in our Cox models (Supplementary Table S4). Subgroup 1 demonstrated non-proportional hazards; we therefore stratified our Cox model in this subgroup by LTC residence. We included the following covariates in our final models: sex, LTC residence, number of general practitioner visits, urban/rural residence, and neighborhood income quintile. We also included the number of individual general practitioners visited in our Cox model for subgroup 4, based on superior model performance using our variable selection algorithm. We examined outliers using DfBeta residuals.[60] We excluded patients with missing information from the survival analysis.

We performed all analyses in R (R Development Core Team, 2011) and used the R-packages "cluster" and "survival."[65,66] We used $\alpha < 0.05$ as the threshold for statistical significance. We provide further detailed methods in the Supplementary Materials.

3 | RESULTS

Overall, 1,196,353 patients made at least 1 ED visit, and 205,136 were classified as frequent users over our 4 years of data. During our study year April 1, 2013 to March 31, 2014, we identified 58,491 frequent users with a median age of 53.03 years (interquartile range [IQR]: 34.93, 72.19), 25,784 (46.6%) of whom were male.

3.1 | Frequent user subgroups

Our cluster analysis identified 4 frequent user subgroups. We present their characteristics pertaining to the clustering variables, demographics, and health care use in Table 1.

Subgroup 1 ("Elderly") had a median age of 77 years (IQR: 66–85), was more frequently female, and made a median of 5 visits/year (IQR: 4–6). They had high indicators of complexity: median of 8 prescriptions (IQR: 5–11) and made visits related to a median of 3 diagnostic categories (IQR: 2–4). Endocrine and metabolic disorders accounted for 8.3% of their visits, and circulatory complaints for 5.5% of their visits.

Subgroup 2 ("Mental Health and Alcohol Use") had a median age of 48 years (IQR: 34–61), was more frequently male, and made a median of 13 visits/year (IQR: 10–16). Mental health accounted for 13.3% of visits. Alcohol-related visits accounted for 6.1% of visits, and 3.9% related specifically to alcohol intoxication.[67,68] Patients in Subgroup 2 were relatively medically complex, with a median 8 prescription medications (IQR: 5–11), and ED visits related to a median of 4 different diagnostic categories (IQR: 3–6). They made a median 31 general practitioner visits (IQR: 19–51); however, only 23.7% received a majority of services from 1 primary care physician.

Subgroup 3 ("Young Mental Health") had a median age of 39 years ([IQR]: 28–51), was more frequently female, and made a median of 5 visits/year (IQR: 4–6). They had moderate complexity relative to other subgroups, as indicated by a median 4 prescription medications (IQR: 2–6) and a median of 2 different diagnostic categories (IQR: 1–3). Mental illness accounted for 5.6% of their ED visits; specifically, alcohol-related presentations accounted for 1.7% of their visits.[67,68]

Subgroup 4 ("Short-term") had a median age of 50 years (IQR: 34–65), was more frequently male, and made a median of 4 visits/year (IQR: 4–5). They had low indicators of complexity (median 3 prescriptions [IQR: 2–5], and made visits related to a median of 2 diagnostic categories [IQR: 1–2]). They made very regularly spaced visits over a short time (ED visits within a median of 1 month [IQR: 0.8–1] over the baseline year). Common diagnoses in this group included urinary tract infection (5.4%), follow-up examination (4.8%), pyelonephritis (4.5%), and abscess (3.6%).

Our sensitivity analysis among the top 5% of frequent users identified 4 subgroups with similar characteristics (Supplementary Table S5). Among the top 1% of frequent users, 3 of 4 identified subgroups were similar ("Elderly;" "Mental Health and Alcohol Use;" "Young Mental Health"); however, we no longer identified a "Short-term" subgroup. Instead, we identified a "Respiratory Illness" subgroup comprising older patients (median age 60 [IQR: 48–70]) who had made a median of 7 (IQR: 5–9) respiratory-related visits/year (Supplementary Table S6).

3.2 | Frequent users' mortality

We observed high mortality overall among frequent users (11.3% [6635/58,491]), compared to non-frequent users in our study year (3.81% [18,993/497,936]). Furthermore, our analysis revealed different 365-day mortality among frequent user subgroups following the index visit, demonstrated by Kaplan-Meier curves (Figure 1). We observed the highest mortality in Subgroup 1 ("Elderly") at 24.7%, followed by 12.3% in Subgroup 2 ("Mental Health and Alcohol Use"), 2.2% in Subgroup 3 ("Young Mental Health"), and 1.4% in Subgroup 4 ("Short-term"). Causes of death also varied among subgroups (Table 2). The top causes of mortality were lung cancer for Subgroup 1 ("Elderly"), chronic obstructive pulmonary disease for Subgroup 2 ("Mental Health and Alcohol Use"), lung cancer for Subgroup 3 ("Young Mental Health"), and atherosclerotic heart disease for Subgroup 4 ("Short-term"). Figure 2 illustrates percentages of patients who died within each subgroup, with subgroups separated visually by median age and size. We present Kaplan-Meier curves for all variables, stratified by subgroup in Supplementary Figures S5–9.

In our adjusted analysis, male sex was associated with an increased hazard ratio for mortality among all subgroups. LTC residence was associated with an increased hazard ratio for mortality in the "Mental Health and Alcohol Use" and "Young Mental Health" subgroups. Residing in the lowest and second-highest neighborhood income quintiles were protective of mortality in the "Mental Health and Alcohol Use" subgroup, and residing in a rural location increased the hazard ratio for mortality in the subgroups of "Elderly" not residing in LTC and "Young Mental Health." Increased numbers of general practitioner physician visits were associated with a slight increased hazard ratio for mortality

TABLE 1 Demographic and visit characteristics for each subgroup

	Elderly	Mental health and alcohol use	Young mental health	Short-term
Subgroup characteristics				
Age, median (IQR)	77 (66–85)	48 (34–61)	39 (28–51)	50 (34–65)
Number of visits to the ED, median (IQR)	5 (4–6)	13 (10–16)	5 (4–6)	4 (4–5)
Number of discharge diagnosis mental illness chapters, median (IQR)	0 (0–0)	0 (0–3)	0 (0–0)	0 (0–0)
Number of discharge diagnosis circulatory chapters, median (IQR)	0 (0–1)	0 (0–0)	0 (0–0)	0 (0–0)
Number of discharge diagnosis respiratory chapters, median (IQR)	0 (0–0)	0 (0–1)	0 (0–0)	0 (0–0)
Number of discharge diagnosis digestive chapters, median (IQR)	0 (0–0)	0 (0–1)	0 (0–0)	0 (0–0)
Number of prescription medications, median (IQR)	8 (5–11)	8 (5–11)	4 (2–6)	3 (2–5)
Number of different discharge diagnosis chapters, median (IQR)	3 (2–4)	4 (3–6)	2 (1–3)	2 (1–2)
Number of months in the year visited ED, median (IQR)	4 (3–4)	7 (6–8)	3 (3–4)	1 (1–1)
Regularity index, median (IQR)	3.0×10^{-4} (1.0×10^{-4}-9×10^{-4}	1.3×10^{-3} (7×10^{-4}-2.4×10^{-3})	3×10^{-4} (1×10^{-4}-8×10^{-3})	1 (0.8–1)
Patient demographic characteristics				
Total patient number, n (% of all top 10% users)	21761 (39.3)	4278 (7.7)	28164 (50.8)	1166 (2.1)
Sex				
Female, n (%)	11525 (53.0)	2093 (48.9)	15388 (54.6)	569 (48.8)
Male, n (%)	10228 (47.0)	2183 (51.0)	12776 (45.4)	597 (51.2)
Unknown, n (%)	8 (0)	<5 (0)	<5 (0)	<5 (0)
Neighborhood income quintile				
1st quintile, n (%)	6112 (28.1)	1754 (41.1)	8734 (31.1)	242 (20.8)
2nd quintile, n (%)	4626 (21.3)	850 (19.9)	5934 (21.1)	233 (20.0)
3rd quintile, n (%)	3935 (18.1)	699 (16.3)	4942 (17.5)	233 (20.0)
4th quintile, n (%)	3462 (15.9)	454 (10.6)	4372 (15.5)	208 (17.8)
5th quintile, n (%)	3361 (15.4)	397 (9.3)	3685 (13.1)	239 (20.4)
Unknown, n (%)	255 (1.2)	118 (2.8)	404 (1.4)	10 (0.9)
NA, n (%)	10 (0)	6 (0)	93 (0.3)	<5 (0.1)
Urban/rural				
Urban, n (%)	20820 (95.7)	4116 (96.2)	26489 (94.1)	1090 (93.5)
Rural, n (%)	938 (4.3)	162 (3.8)	1661 (5.9)	76 (6.5)
Unknown, n (%)	<5 (0)	0 (0)	11 (0)	0 (0)
ED visit characteristics (all visits)				
Arrive by ambulance				
Ground ambulance, n (%)	43382 (36.6)	22866 (35.6)	22660 (15.8)	178 (3.1)
No ambulance, n (%)	75273 (63.4)	41422 (64.4)	120864 (84.2)	5554 (96.9)
Triage level (CTAS)				
1 (Resuscitation), n (%)	889 (0.7)	332 (0.5)	532 (0.4)	2 (0.0)
2 (Emergent), n (%)	26681 (22.5)	10736 (16.7)	19676 (13.7)	248 (4.3)
3 (Urgent), n (%)	62803 (52.9)	34293 (53.3)	69415 (48.3)	1293 (22.6)

(Continues)

TABLE 1 (Continued)

	Elderly	Mental health and alcohol use	Young mental health	Short-term
4 (Less-urgent), n (%)	24433 (20.6)	16061 (25.0)	45971 (32.0)	3128 (54.6)
5 (Non-urgent), n (%)	3289 (2.8)	2609 (4.1)	7205 (5.0)	844 (14.7)
Unknown, n (%)	593 (0.5)	281 (0.4)	773 (0.6)	217 (3.8)
Top 5 ICD-10 ED diagnosis, (non-missing), n (%)				
	Chest pain, unspecified (R074), 4357 (3.7)	Other and unspecified abdominal pain (R104), 3804 (5.9)	Other and unspecified abdominal pain (R104), 7566 (5.3)	Other specified medical care (Z5188), 363 (6.3)
	Other and unspecified abdominal pain (R104), 3685 (3.1)	Mental and behavioral disorders due to use of alcohol, acute intoxication (F100), 2485 (3.9)	Chest pain, unspecified (R074), 3546 (2.5)	Urinary tract infection (N390), 311 (5.4)
	Urinary tract infection (N390), 2896 (2.4)	Chest pain, unspecified (R074), 2040 (3.2)	Cellulitis (L039), 2758 (1.9)	Follow-up examination after unspecified treatment for other conditions (Z099), 273 (4.8)
	Cellulitis (L039), 2485 (2.1)	Anxiety disorder, unspecified (F419), 1326 (2.1)	Lower back pain (M545), 2568 (1.8)	Acute pyelonephritis (N10), 259 (4.5)
	Pneumonia, unspecified (J189), 2285 (1.9)	Lower back pain (M545), 1220 (1.9)	Urinary tract infection (N390), 2213 (1.5)	Cutaneous abscess, furuncle and carbuncle, unspecified (L029), 209 (3.6)
Top 5 ICD-10 ED diagnostic categories, (non-missing), n (%)				
	Symptoms, signs and abnormal clinical and laboratory findings (XVIII), 22660 (19.1)	Symptoms, signs and abnormal clinical and laboratory findings (XVIII), 12387 (19.3)	Symptoms, signs and abnormal clinical and laboratory findings (XVIII), 22787 (15.9)	Diseases of the musculoskeletal system and connective tissue (XIV), 769 (13.4)
	Endocrine, nutritional and metabolic diseases (IX), 9874 (8.3)	Mental and behavioral disorders (V), 8545 (13.3)	Injury, poisoning, and certain other consequences of external causes (XIX), 16291 (11.3)	Factors influencing health status and contract with health services (XXI), 730 (12.7)
	Injury, poisoning, and certain other consequences of external causes (XIX), 9731 (8.2)	Injury, poisoning, and certain other consequences of external causes (XIX), 5827 (9.1)	Mental and behavioral disorders (V), 8001 (5.6)	Injury, poisoning, and certain other consequences of external causes (XIX), 571 (10.0)
	Diseases of the circulatory system (X), 6490 (5.5)	Diseases of the respiratory system (XI), 3590 (5.6)	Diseases of the musculoskeletal system and connective tissue (XIV), 7060 (4.9)	Symptoms, signs, and abnormal clinical and laboratory findings (XVIII), 411 (7.2)
	Diseases of the skin and subcutaneous tissue (XIII), 5265 (4.4)	Factors influencing health status and contract with health services (XXI), 3455 (5.4)	Diseases of the skin and subcutaneous tissue (XIII), 6205 (4.3)	Diseases of the digestive system (XII), 377 (6.6)

(Continues)

TABLE 1 (Continued)

	Elderly	Mental health and alcohol use	Young mental health	Short-term
Discharge disposition				
Discharged home or place of residence, n (%)	79153 (66.7)	50847 (79.1)	124562 (78)	5556 (25)
Admitted or transferred, n (%)	39094 (32.9)	12585 (19.6)	18087 (12.6)	169 (2.9)
Left before completion of treatment, n (%)	360 (0.4	870 (1.4)	908 (0.6)	7 (0.1)
Died, n (%)	81 (0.1)	10 (0)	15 (0)	<5 (0)
Other health care utilization				
Number of admissions per person in last 365 days, median (IQR)	2 (1–3)	3 (2–5)	1 (1–2)	1 (1–1)
Median time admitted (days, [IQR])	5 (2–12)	3 (1–8)	3 (1–7)	2 (1–4)
Long-term care residence, n (%)	1340 (6.2)	59 (1.4)	72 (0.3)	<5
Number of general practitioner visits, median (IQR)	25 (16–39)	31 (19–51)	13 (8–21)	9 (6–14)
Number of individual general practitioner visited, median (IQR)	8 (5–11)	14 (9–19)	7 (5–9)	5 (3–7)
Majority source of care, n (%)	12288 (56.5)	1012 (23.7)	10040 (35.6)	410 (35.2)

CTAS, Canadian Triage and Acuity Scale; ED, emergency department; ICD, International Classification of Diseases; IQR, interquartile range.

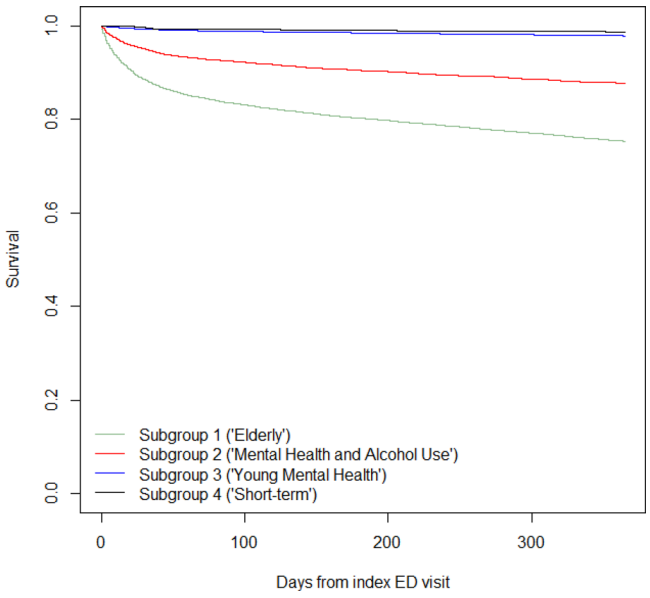

FIGURE 1 Kaplan-Meier curves stratified by subgroup

among patients in the "Elderly" subgroup not residing in LTC, the "Mental Health and Alcohol Use," and "Young Mental Health" subgroups (Table 3).

4 | LIMITATIONS

Our analysis is constrained by inherent limitations in large administrative database analyses. First, our use of NACRS to identify patients who made ED visits in BC will miss records from EDs not reporting to NACRS. Despite this, NACRS remains the most comprehensive provincial ED visit repository, and by 2015/16, our most recent year of data, 29 BC EDs contributed 1.57 million records to NACRS, accounting for 74% provincial coverage.[69] Second, our reliance on ED diagnoses to characterize frequent users will be affected by accuracy and completeness of discharge diagnostic information. We observed 31.8% and 30.1% missing data for discharge diagnoses for frequent and non-frequent users, respectively (we do not suspect that missing data are systematically different). Discharge diagnosis was a conditional mandatory field for Level 2 reporting facilities in BC during the study period, meaning that completion of either but not both of presenting complaint of discharge diagnosis was required.[70,71] Institutions have variable quality and standardization of NACRS coding in BC; training and medical expertise of coders can vary widely, introducing possible transcription and interpretation errors. Furthermore, additional diagnoses beyond the primary visit diagnosis may not be reliably captured. To mitigate coding inaccuracies in individual discharge diagnoses, we included broader diagnostic categories (ICD-10 diagnostic chapters) in our frequent user characterization and cluster analysis, an approach that has demonstrated improved reliability.[72] Third, although NACRS has a field for scheduled revisits, this is unreliably coded. We therefore developed an algorithm to identify presumed scheduled revisits for cellulitis treatment, based on our clinical experience. Although our method has not been validated, we examined all visits occurring within 48 hours of one another and did not observe that diagnoses other than cellulitis commonly appeared. Fourth, we could not explore important patient factors not captured within included databases (eg, housing insecurity/homelessness, ethnicity, employment, and individual/household income). Fifth, our analysis precedes the opioid overdose-related public health emergency in BC declared in

TABLE 2 Characteristics of patients who died within 365 days of index visit, by frequent ED user subgroups

	Elderly (Med age: 77)	Mental health and alcohol use (Med age: 48)	Young mental health (Med age: 39)	Short-term (Med age: 50)
Number of deaths (% of all frequent users within each subgroup)	5385 (24.7%)	527 (12.3%)	618 (2.2%)	16 (1.4%)
Sex				
Female, n (%)	2516 (46.7)	204 (38.7)	253 (40.9)	3 (18.8)
Male, n (%)	2865 (53.2)	321 (60.9)	365 (59.1)	13 (81.3)
Unknown, n (%)	<5 (0.1)	<5 (0.4)	<5 (0)	<5 (0)
Age at death, median (IQR)	80 (70–88)	65 (54–78)	53 (43–61)	71 (57–83)
Neighborhood income quintile				
1st quintile, n (%)	1444 (26.8)	199 (37.8)	212 (34.3)	<5 (25.0)
2nd quintile, n (%)	1202 (22.3)	113 (21.4)	137 (22.2)	5 (31.3)
3rd quintile, n (%)	995 (18.5)	103 (19.5)	93 (15.0)	<5 (6.3)
4th quintile, n (%)	879 (16.3)	45 (8.5)	84 (13.6)	<5 (0)
5th quintile, n (%)	809 (15.0)	50 (9.5)	78 (12.6)	6 (37.5)
Unknown, n (%)	56 (1.0)	16 (3.0)	13 (2.1)	<5 (0)
NA, n (%)	<5 (0)	<5 (0.2)	<5 (0.2)	<5 (0)
Urban/rural				
Urban, n (%)	5118 (95.0)	513 (97.3)	571 (92.4)	13 (81.3)
Rural, n (%)	267 (5.0)	14 (2.7)	47 (7.6)	<5 (18.9)
Top 5 causes of death, n (%)				
	Lung cancer, unspecified, (C349), 463 (8.6)	Unspecified cause of mortality, (R99), 43 (8.2)	Unspecified cause of mortality, (R99), 63 (10.2)	Atherosclerotic heart disease, (I251), 2 (12.5)
	Atherosclerotic heart disease, (I251), 245 (4.5)	COPD, unspecified, (J449), 31 (5.9)	Lung cancer, unspecified, (C349), 39 (6.3)	HIV, unspecified, (B24), 1 (6.3)
	Unspecified dementia, (F03), 190 (3.5)	Lung cancer, unspecified, (C349), 19 (3.6)	Malignant neoplasm of brain, (C719), 25 (4.0)	Liver cell carcinoma, (C220), 1 (6.3)
	COPD, unspecified, (J449), 186 (3.5)	Alcoholic cirrhosis of liver, (K703), 17 (3.2)	Malignant neoplasm of breast, (C509), 23 (3.7)	Lung cancer, unspecified, (C349), 1 (6.3)
	Acute myocardial infarction, unspecified, (I219), 182 (3.4)	Malignant neoplasm of prostate, (C61), 16 (3.0)	Alcoholic cirrhosis of liver, (K703), 22 (3.6)	Malignant melanoma of lower limb, (C437), 1 (6.3)
Long-term care residence, n (%)	644 (12.0)	26 (4.9)	9 (1.5)	<5 (0)
Number of general practitioner visits, median (IQR)	32 (20-51)	43 (25-71)	18 (11-29)	20 (12-27)
Number of individual general practitioner visited, median (IQR)	9 (6-13)	15 (10-21)	8 (5-10)	6 (5-9)
Majority source of care, n (%)	2729 (50.7)	153 (29.0)	254 (41.1)	11 (68.8)

COPD, chronic obstructive pulmonary disease; ED, emergency department; IQR, interquartile range.

April 2016.[73] The opioid epidemic may have changed the characteristics and risk profiles of frequent user subgroups. Nonetheless, our current analysis sheds important light on comorbidities (eg, mental illness and alcohol) and risk factors (eg, LTC residence) that likely continue to affect experiences and outcomes of patients with high health care use affected by opioid overdose. Finally, the generalizability of our BC analysis to other settings is unknown. Nonetheless, despite anticipated interjurisdictional nuances, commonalities between our

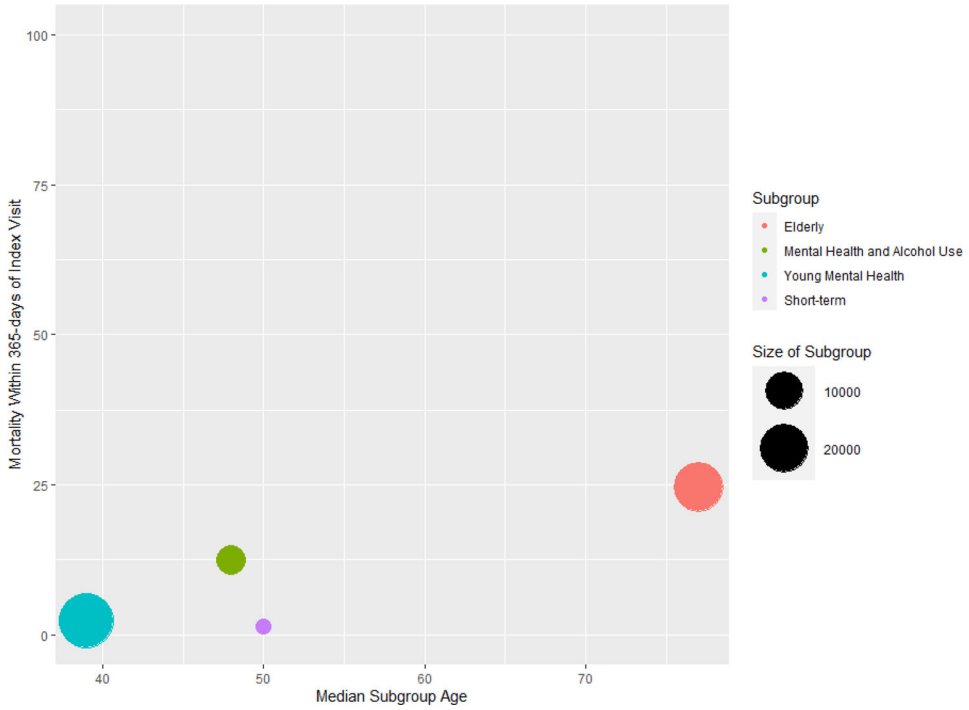

FIGURE 2 Bubble plot illustrating subgroup size, median age, and 365-day mortality

results and existing frequent user characterizations in other locations suggest that our findings capture overarching characteristics with wide relevance.[5,35] Furthermore, our novel application of cluster analysis to identify population-level patterns offers an important proof of concept that could be replicated in other population data sets to better understand nuances of frequent user subgroups in other settings.

5 | DISCUSSION

Our study confirms that distinct subgroups exist within the top 10% of the highest using ED patients, each with unique demographic, clinical, and visit pattern profiles. Our cluster analysis identified 4 frequent user subgroups: complex elderly, middle-aged with comorbid mental health and alcohol-related presentations, young with comorbid mental health, and middle-aged with increased visits over a short term. The final subgroup likely comprises patients who make multiple visits related to a discrete time-limited event, such as a complication, acute infection, or injury. Some of these patients likely represent those making scheduled revisits not captured by our algorithm focused on return visits for cellulitis. Our analyses also indicate that the 365-day mortality among identified subgroups vary widely, with elderly frequent users and middle-aged frequent users with mental health and alcohol use comorbidities having the highest risk. One quarter of patients in our "Elderly" subgroup, and 12.3% of patients in our middle-aged, "Mental Health and Alcohol Use" subgroup died within 365 days of the index visit. In adjusted analysis, characteristics associated with increased hazard ratios for mortality included male sex (all subgroups), LTC residence ("Mental Health and Alcohol Use" and "Young Mental Health" subgroups), and rural residence ("Elderly" subgroup not residing in LTC and "Young Mental Health" subgroup). ED clinicians should consider supportive interventions and closer follow-up plans for frequent users with these profiles.

Our results indicate that frequent user subgroups are important to consider distinctly: their unique characteristics and risk profiles likely indicate important differences in unmet needs and gaps in care that underlie frequent visits. Previous literature has identified heterogeneity with regard to age, reasons for presentation, and comorbidities.[5,74] Our study adds another level of understanding by using a clustering algorithm to identify distinct subgroups based on patterns within a comprehensive, linked provincial ED visit data set. Notably, our data-driven approach confirms much of what is apparent clinically: mental health and alcohol use are important comorbidities among frequent users; and the elderly frequent users with complex medical histories have poor outcomes. In addition to confirming our clinical experience, our study highlights important and novel insights that challenge clinical intuition: nearly 1 in 8 patients within our middle-aged, "Mental Health and Alcohol Use" subgroup died within 1 year. Our finding contradicts previous studies indicating that similar frequent users have lower rates of death.[13,35] The profile of frequent users who are perceived as "regular" patients at local EDs are often treated and discharged, with limited options for follow-up, referrals, or ongoing care, and with minimal attention paid to whether follow-up options actually meet individuals' needs (eg, cultural sensitivity, accessibility). That these patients die often, and are relatively young, indicate that interventions could potentially have a considerable impact on life-years saved. Our analysis corroborates previous findings that frequent ED users are also high users of primary care.[18–23] However, our study adds a unique

TABLE 3 Multivariable models of 365-day mortality stratified by frequent ED user subgroups

	Beta	95% CI	*P* value
Subgroup 1 (elderly) residing in long-term care			
Sex			
Female	Ref	Ref	Ref
Male	1.51	(1.29–1.77)	$3 \times 10^{-07***}$
Number of general practitioner physician visits	1.00	(0.99–1.00)	0.08
Urban/rural			
Urban	Ref	Ref	Ref
Rural	0.69	(0.44–1.09)	0.11
Neighborhood income quintile			
1st quintile	0.89	(0.71–1.11)	0.29
2nd quintile	1.15	(0.91–1.45)	0.24
3rd quintile	Ref	Ref	Ref
4th quintile	0.83	(0.64–1.09)	0.18
5th quintile	0.93	(0.70–1.22)	0.58
Subgroup 1 (elderly) not residing in long-term care			
Sex			
Female	Ref	Ref	Ref
Male	1.38	(1.31–1.47)	$<2 \times 10^{16***}$
Number of general practitioner physician visits	1.01	(1.01–1.01)	$<2 \times 10^{16***}$
Urban/rural			
Urban	Ref	Ref	Ref
Rural	1.19	(1.05–1.36)	0.01
Neighborhood income quintile			
1st quintile	0.95	(0.87–1.04)	0.27
2nd quintile	1.04	(0.95–1.14)	0.43
3rd quintile	Ref	Ref	Ref
4th quintile	1.06	(0.96–1.17)	0.26
5th quintile	1.00	(0.90–1.10)	0.94
Subgroup 2 (mental health and alcohol use)			
Sex			
Female	Ref	Ref	Ref
Male	1.62	(1.36–1.94)	$1 \times 10^{-07***}$
Long-term care			
False	Ref	Ref	Ref
True	2.63	(1.71–4.05)	$1.1 \times 10^{-05***}$
Number of general practitioner physician visits	1.01	(1.01–1.01)	$<2 \times 10^{16***}$
Urban/rural			
Urban	Ref	Ref	Ref
Rural	0.62	(0.35–1.07)	0.09
Neighborhood income quintile			
1st quintile	0.75	(0.59–0.95)	0.02*
2nd quintile	0.89	(0.69–1.17)	0.42
3rd quintile	Ref	Ref	Ref
4th quintile	0.64	(0.45–0.90)	0.01*
5th quintile	0.87	(0.62–1.22)	0.41
Subgroup 3 (young mental health)			
Sex			
Female	Ref	Ref	Ref
Male	1.79	(1.52–2.11)	$2.8 \times 10^{-12***}$
Long-term care			
False	Ref	Ref	Ref
True	4.34	(2.24–8.42)	$1.4 \times 10^{-05***}$
Number of general practitioner physician visits	1.01	(1.01–1.02)	$<2 \times 10^{16***}$
Urban/rural			
Urban	Ref	Ref	Ref
Rural	1.36	(1.01–1.83)	0.05*
Neighborhood income quintile			
1st quintile	1.25	(0.98–1.59)	0.08
2nd quintile	1.19	(0.91–1.55)	0.20
3rd quintile	Ref	Ref	Ref
4th quintile	1.01	(0.75–1.36)	0.93
5th quintile	1.14	(0.84–1.54)	0.40
Subgroup 4 (short-term)			
Sex			
Female	Ref	Ref	Ref
Male	5.80	(1.59–21.12)	0.008**
Long-term care			
False	Ref	Ref	Ref
True	0.18	(0.00–Inf)	1.00
Number of general practitioner physician visits	1.02	(0.99–1.05)	0.27
Number of individual general practitioner physicians visited	1.24	(1.01–1.52)	0.04**
Urban/rural			
Urban	Ref	Ref	Ref
Rural	2.38	(0.66–8.52)	0.18
Neighborhood income quintile			
1st quintile	3.03	(0.33–27.62)	0.29
2nd quintile	4.38	(0.50–38.12)	0.17
3rd quintile	Ref	Ref	Ref
4th quintile	0.00	(0.00–Inf)	1.00
5th quintile	5.44	(0.65–45.32)	0.10

ED, emergency department.

dimension—the *continuity*, not just quantity, of primary care likely matters.[75,76] Only 23.7% of patients in our "Mental Health and Alcohol Use" subgroup had 1 physician serving as a majority source of care for their primary care visits. Investing in building trusting therapeutic relationships between frequent users with mental health and alcohol use comorbidities and primary care physicians should be explored as an intervention that could improve outcomes and save health care system costs.

In conclusion, our study provides profiles of frequent ED users who have a high overall mortality compared to non-frequent users, but where risk is concentrated among patients who are complex and elderly and among middle-aged frequent users with mental health and alcohol use comorbidities. Because of frequent users' heterogeneity, one-size-fits-all interventions are unlikely to succeed. Future studies should explore unmet needs and drivers of ED use among the subgroups identified, and should collaboratively develop, pilot and test interventions targeted to specific frequent user subgroups.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

JM conceived the study, designed the analysis, obtained research funding, analyzed the data, interpreted results, and provided overall study oversight. FOS designed the analysis, analyzed the data, and interpreted results. MJM, MJS, KD, BRH, EG, CMH, and JH provided feedback on study design, data analysis, and results interpretation. KMM served as a methodological expert, designed the analysis, analyzed the data, and provided feedback on results interpretation. JM drafted the manuscript and all authors contributed substantially to its revision. JM takes responsibility for the paper as a whole.

DISCLAIMER

All inferences, opinions, and conclusions drawn in this article are those of the authors and do not reflect the opinions or policies of the Data Steward(s).

REFERENCES

1. Wammes JJG, Wees PJvander, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ Open*. 2018;8(9):e023113.
2. Johnson TL, Rinehart DJ, Durfee J, et al. For many patients who use large amounts of health care services, the need is intense yet temporary. *Health Aff (Millwood)*. 2015;34(8):1312-1319.
3. Kronick RG, Bella M, Gilmer TP, Somers SA. *The faces of medicaid II: Recognizing the care needs of people with multiple chronic conditions*. Center for Health Care Strategies, Inc; 2007.
4. Jiang JH, Weiss AJ, Barrett ML. Characteristics of Emergency Department visits for super-utilizers by payer. *HCUP statistical brief #221. Agency for Healthcare Research and Quality*. 2014.
5. LaCalle E, Rabin E. Frequent users of emergency departments: the myths, the data, and the policy implications. *Ann Emerg Med*. 2010;56(1):42-48.
6. Canadian Institute for Health Information. Emergency Department Visits in 2014–2015. 2015; Available from: https://secure.cihi.ca/free_products/NACRS_ED_QuickStats_Infosheet_2014-15_ENweb.pdf.
7. Mitchell MS, Leon CLK, Byrne TH, Lin W, Bharel M. Cost of health care utilization among homeless frequent Emergency Department users. *Psychological Services*. 2017;14(2):193-202.
8. Korczak V, Shanthosh J, Jan S, Dinh M, Lung T. Costs and effects of interventions targeting frequent presenters to the emergency department: a systematic and narrative review. *BMC Emergency Medicine*. 2019;19(1):1-83.
9. Fuda KK, Immekus R. Frequent users of Massachusetts emergency departments: a statewide analysis. *Ann Emerg Med*. 2006;48(1). 16.e1-16.e8.
10. Hansagi H, Olsson M, Sjöberg S, Tomson Y, Göransson S. Frequent use of the hospital emergency department is indicative of high use of other health care services. *Ann Emerg Med*. 2001;37(6):561-567.
11. Moe J, Kirkland S, Ospina MB, et al. Mortality, admission rates and outpatient use among frequent users of emergency departments: a systematic review. *Emergency Medicine Journal: EMJ*. 2016;33(3):230-236.
12. Pines JM, Asplin BR, Kaji AH, et al. Frequent users of emergency department services: gaps in knowledge and a proposed research agenda. *Acad Emerg Med*. 2011;18(6). e64-e69.
13. Ruger JP, Richter CJ, Spitznagel EL, Lewis LM. Analysis of costs, length of stay, and utilization of emergency department services by frequent users: implications for health policy. *Acad Emerg Med*. 2004;11(12):1311-1317.
14. Moe J, Kirkland SW, Rawe E, et al. Effectiveness of interventions to decrease emergency department visits by adult frequent users: a systematic review. *Acad Emerg Med*. 2017;24(1):40-52.
15. Althaus F, Paroz S, Hugli O, et al. Effectiveness of interventions targeting frequent users of emergency departments: a systematic review. *Ann Emerg Med*. 2011;58(1):41-52. e42.
16. Soril LJJ, Leggett LE, Lorenzetti DL, Noseworthy TW, Clement FM. Reducing frequent visits to the Emergency Department: a systematic review of interventions. *PLoS One*. 2015;10(4):e0123660.
17. Kumar GS, Klein R. Effectiveness of case management strategies in reducing emergency department visits in frequent user patient populations: a systematic review. *J Emerg Med*. 2013;44(3):717-729.
18. Lucas RH, Sanford SM. An analysis of frequent users of emergency care at an urban university hospital. *Ann Emerg Med*. 1998;32(5):563-568.
19. Sandoval E, Smith S, Walter J, et al. A comparison of frequent and infrequent visitors to an urban Emergency Department. *J Emerg Med*. 2010;38(2):115-121.
20. Ovens HJ, Chan BT. Heavy users of emergency services: a population-based review. *CMAJ*. 2001;165(8):1049-1050.
21. Zuckerman S, Shen Y. Characteristics of occasional and frequent emergency department users: do insurance coverage and access to care matter?. *Med Care*. 2004;42(2):176-182.
22. Blank FSJ, Li H, Henneman PL, et al. A descriptive study of heavy emergency department users at an academic emergency department reveals heavy ED users have better access to care than average users. *J Emerg Nurs*. 2005;31(2):139-144.
23. Palmer E, Leblanc-Duchin D, Murray J, Atkinson P. Emergency department use: is frequent use associated with a lack of primary care provider?. *Canadian family physician Médecin de famille canadien*. 2014;60(4):e223.
24. McCusker J, Verdon J. Do geriatric interventions reduce emergency department visits? A systematic review. The journals of gerontology. *Series A, Biological sciences and medical sciences*. 2006;61(1):53-62.
25. Hastings SN, Heflin MT. A systematic review of interventions to improve outcomes for elders discharged from the Emergency Department. *Acad Emerg Med*. 2005;12(10):978-986.
26. Milbrett P, Halm M. Characteristics and predictors of frequent utilization of emergency services. *J Emerg Nurs*. 2009;35(3):191-198.
27. Hunt KA, Weber EJ, Showstack JA, Colby DC, Callaham ML. Characteristics of frequent users of emergency departments. *Ann Emerg Med*. 2006;48(1):1-8.

28. Byrne M, Murphy AW, Plunkett PK, McGee HM, Murray A, Bury G. Frequent attenders to an emergency department: a study of primary health care use, medical profile, and psychosocial characteristics. *Ann Emerg Med*. 2003;41(3):309-318.
29. Moe J, Bailey AL, Oland R, Levesque L, Murray H. Defining, quantifying, and characterizing adult frequent users of a suburban Canadian emergency department. *CJEM*. 2013;15(4):214-226.
30. Mandelberg JH, Kuhn RE, Kohn MA. Epidemiologic analysis of an urban, public emergency department's frequent users. *Acad Emerg Med*. 2000;7(6):637-646.
31. Huang J, Tsai W, Chen Y, Hu W, Yang D. Factors associated with frequent use of emergency services in a medical center. *J Formos Med Assoc*. 2003;102(4):222.
32. Brubacher JR, Mabie A, Ngo M, et al. Substance-related problems in patients visiting an urban canadian emergency department. *CJEM*. 2008;10(3):198-204.
33. Mehl-Madrona LE. Prevalence of psychiatric diagnoses among frequent users of rural emergency medical services. *Canadian journal of rural medicine: the official journal of the Society of Rural Physicians of Canada*. 2008;13(1):22.
34. Curran GM, Sullivan G, Williams K, Han X, Allee E, Kotrla KJ. The association of psychiatric comorbidity and use of the emergency department among persons with substance use disorders: an observational cohort study. *BMC emergency medicine*. 2008;8(1):17.
35. Doupe MB, Palatnick W, Day S, et al. Frequent users of emergency departments: developing standard definitions and defining prominent risk factors. *Ann Emerg Med*. 2012;60(1):24.
36. Guo Q, Lu X, Gao Y, et al. Cluster analysis: a new approach for identification of underlying risk factors for coronary artery disease in essential hypertensive patients. *Sci Rep*. 2017;7(1):43965.
37. McGrail K, Lavergne R, Lewis S, Peterson S, Barer M, Garrison S. Classifying physician practice style: a new approach using administrative data in British Columbia. *Med Care*. 2015;53(3):276-282.
38. Williamson DJ, Burn GL, Simoncelli S, et al. Machine learning for cluster analysis of localization microscopy data. *Nat Commun*. 2020;11(1). 1493-x.
39. Tothill RW, Tinker AV, George J, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin Cancer Res*. 2008;14(16):5198-5208.
40. Canadian Institute for Health Information [creator] (2011). National Ambulatory Care Reporting System (NACRS). Population Data BC. Data Extract. MOH. 2011; https://www.popdata.bc.ca/data/health/dad. Updated 2018. Accessed June18, 2020.
41. Canadian Institute for Health Information [creator] (2011). Discharge Abstract Database (Hospital Separations). Population Data BC [publisher]. Data Extract. MOH. 2011; https://www.popdata.bc.ca/data/health/dad. Updated 2018. Accessed April 29, 2020.
42. British Columbia Ministry of Health [creator] (2011). Medical Services Plan (MSP) Payment Information File. Population Data BC [publisher]. Data Extract. MOH. 2011; https://www.popdata.bc.ca/data/health/msp. Updated 2018. Accessed April 29, 2020.
43. British Columbia Ministry of Health [creator] (2011). PharmaCare. V2. Population Data BC [publisher]. Data Extract. MOH. 2011; https://www.popdata.bc.ca/data/health/PharmaNet. Updated 2019. Accessed April 29, 2020.
44. BC Vital Statistics Agency [creator] (2011). Vital Statistics Deaths. Population Data BC [publisher]. Data Extract BC Vital Statistics Agency. 2011; https://www.popdata.bc.ca/data/population. Updated 2018. Accessed April 29, 2020.
45. Ark TK, Kesselring S, Hills B, McGrail K. Population data british columbia: a data resource for research. *International Journal of Population Data Science*. 2020;4(2).
46. Chamberlayne R, Green B, Barer ML, Hertzman C, Lawrence WJ, Sheps SB. Creating a population-based linked health database: a new resource for health services research. *Can J Public Health*. 1998;89(4):270-273.
47. Canadian Institute for Health Information. *Dynamic cohort of complex, high system users - 2011–2015: A CIHR-CIHI collaboration*. CIHR-CIHI 2017.
48. Statistics Canada. Postal code conversion file (PCCF), reference guide, 2017. 2017 December 13.
49. General Practice Services Committee. General practice services committee… improving primary care in British Columbia: Annual report 2009/10. 2009 /10.
50. Lavergne R, McGrail K, Peterson S. *Defining and measuring full service family practice in BC*. 1991–2006: Centre for Health Services and Policy Research, University of British Columbia; 2013.
51. Lavergne MR, Peterson S, McKendry R, Sivananthan S, McGrail K. Full-service family practice in British Columbia: policy interventions and trends in practice, 1991–2010. *Healthcare Policy/Politiques de Santé*. 2014;9(4):32-47.
52. Wong ST, Katz A, Williamson T, et al. Can linked electronic medical record and administrative data help us identify those living with frailty?. *International Journal of Population Data Science*. 2020. Submitted.
53. Gibson DAJ, Moorin RE, Preen D, Emery J, Holman CDJ. Enhanced primary care improves GP service regularity in older patients without impacting on service frequency. *Australian journal of primary health*. 2012;18(4):295.
54. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *Am J Epidemiol*. 2001;154(9):854-864.
55. World Health Organization. ICD-10 Version:2019. Available from: https://icd.who.int/browse10/2019/en. [cited April 29, 2020].
56. Park S. *Patient complexity in primary care: Grouping systems with a special focus on validity and applications in primary care settings*. UBC Department of Family Medicine. 2016 December 31.
57. Quail JM, Lix LM, Osman BA, Teare GF. Comparing comorbidity measures for predicting mortality and hospitalization in three population-based cohorts. *BMC health services research*. 2011;11(1):146.
58. NCSS Statistical Software. Chapter 445: Hierarchical Clustering/Dendograms. Available from: https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Hierarchical_Clustering-Dendrograms.pdf. [cited April 29,2020].
59. Moe J, O'Sullivan F, McGregor M, et al. Characterizing frequent emergency department users in BC: A retrospective analysis of linked provincial databases. (in submission) 2020.
60. Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett*. 2010;31(8):651-666.
61. Sisodia D, Singh L, Sisodia S, Saxena K. Clustering techniques: a brief survey of different clustering algorithms. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*. 2012;1(3):82-87.
62. Caliński T, Harabasz J. A dendrite method for cluster analysis. *Communications in Statistics*. 1974;3(1):1-27.
63. Sarstedt M, Mooi E. *A concise guide to market research*. 2nd ed. 2014 ed. Berlin: Heidelberg: Springer Berlin Heidelberg; 2014.
64. StataCorp. *Stata: Release 13. Statistical software*. College Station, Texas: StataCorp LP; 2013.
65. Rousseeuw P, Struyf A, Hubert M, Package 'cluster'. 2019; Available from: https://cran.r-project.org/web/packages/cluster/cluster.pdf. [cited May 9,2020].
66. Therneau TM, Package 'survival'. 2020; Available from: https://cran.r-project.org/web/packages/survival/survival.pdf. [citedMay 9,2020].
67. Saunders JB, Room R. Enhancing the ICD system in recording alcohol's involvement in disease and injury. *Alcohol Alcohol*. 2012;47(3):216-218.
68. *Hospital stays for harm caused by substance use, Appendices to indicator library*. 2020.

69. Canadian Institute for Health Information. NACRS Emergency Department Visits and Length of Stay by Province/Territory, 2016–2017. 2017; Available from: https://www.cihi.ca/en/nacrs-emergency-department-visits-and-length-of-stay-2016-2017. [cited May 9,2020].
70. Canadian Institute for Health information. NACRS FAQ. Available from: https://silo.tips/download/nacrs-frequently-asked-questions.
71. Canadian Institute for Health information. NACRS Data Elements 2015–2016. Available from: https://www.cihi.ca/sites/default/files/document/nacrs_data_elements_2015_16_en.pdf.
72. Wockenfuss R, Frese T, Herrmann K, Claussnitzer M, Sandholzer H. Three- and four-digit ICD-10 is not a reliable classification system in primary care. *Scand J Prim Health Care*. 2009;27(3):131-136.
73. Province of British Columbia. Provincial health officer declares public health emergency. 2016; Available from: https://news.gov.bc.ca/releases/2016HLTH0026-000568. [cited June 19,2020].
74. van Tiel S, Rood P, Bertoli-Avella A, et al. Systematic review of frequent users of emergency departments in non-US hospitals: state of the art. *Eur J Emerg Med*. 2015;22(5):306-315.
75. Gill JM, 3, Mainous, AG, Nsereko M. The effect of continuity of care on emergency department use. *Arch Fam Med*. 2000;9(4):333-338.
76. Ionescu-Ittu R, McCusker J, Ciampi A, et al. Continuity of primary care and emergency department utilization among elderly people. *Canadian Medical Association Journal*. 2007;177(11):1362-1368.

AUTHOR BIOGRAPHY

Jessica Moe, MD, MSC, is an Assistant Professor in the Department of Emergency Medicine, Vancouver General Hospital, Vancouver, British Columbia, Canada.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Moe J, O'Sullivan F, McGregor MJ, et al. Identifying subgroups and risk among frequent emergency department users in British Columbia. *JACEP Open*. 2021;2:e12346. https://doi.org/10.1002/emp2.12346