BMJ Open

Statistical tools used for analyses of frequent users of emergency department: a scoping review

Yohann Chiu,[1] François Racine-Hemmings,[1] Isabelle Dufour,[1] Alain Vanasse,[1] Maud-Christine Chouinard,[2] Mathieu Bisson,[1] Catherine Hudon[1]

To cite: Chiu Y, Racine-Hemmings F, Dufour I, *et al*. Statistical tools used for analyses of frequent users of emergency department: a scoping review. *BMJ Open* 2019;**9**:e027750. doi:10.1136/bmjopen-2018-027750

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-027750).

Received 6 November 2018
Revised 22 March 2019
Accepted 18 April 2019

Check for updates

© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

[1]Department of Family Medicine and Emergency Medicine, Université de Sherbrooke, Sherbrooke, Quebec, Canada
[2]Department of Health Sciences, Université du Québec à Chicoutimi, Chicoutimi, Quebec, Canada

Correspondence to
Yohann Chiu;
yohann.chiu@usherbrooke.ca

ABSTRACT

Objective Frequent users represent a small proportion of emergency department users, but they account for a disproportionately large number of visits. Their use of emergency departments is often considered suboptimal. It would be more efficient to identify and treat those patients earlier in their health problem trajectory. It is therefore essential to describe their characteristics and to predict their emergency department use. In order to do so, adequate statistical tools are needed. The objective of this study was to determine the statistical tools used in identifying variables associated with frequent use or predicting the risk of becoming a frequent user.

Methods We performed a scoping review following an established 5-stage methodological framework. We searched PubMed, Scopus and CINAHL databases in February 2019 using search strategies defined with the help of an information specialist. Out of 4534 potential abstracts, we selected 114 articles based on defined criteria and presented in a content analysis.

Results We identified four classes of statistical tools. Regression models were found to be the most common practice, followed by hypothesis testing. The logistic regression was found to be the most used statistical tool, followed by χ^2 test and t-test of associations between variables. Other tools were marginally used.

Conclusions This scoping review lists common statistical tools used for analysing frequent users in emergency departments. It highlights the fact that some are well established while others are much less so. More research is needed to apply appropriate techniques to health data or to diversify statistical point of views.

Strengths and limitations of this study

- ► First overview of statistical tools used in frequent users analysis.
- ► Follows a well-defined methodological framework in an extensive body of literature.
- ► Quality assessment is not performed in a scoping review.
- ► Studies in other languages than English or French might have been missed.

BACKGROUND

Emergency department (ED) 'frequent users' are a sub-group of ED users that make repeated, multiple visits during a given amount of time. Though there is no consensus about definition for frequent users, thresholds in the literature range from two to more than 10 ED visits per year,[1 2] while the most common one is more than four ED visits per year.[1 2] Frequent users represent a small proportion of ED users but account for a large number of visits.[3–5] They often display complex characteristics such as low socioeconomic status combined with physical and mental health issues.[6] As such, their ED use is considered suboptimal,[7] as the best strategy would be to identify those patients at an earlier stage in their health problem trajectory, in order to treat them more efficiently.[8] Furthermore, frequent users' visits may lead to overcrowding in EDs and decreased quality of care.[2] Identifying factors that best describe those users and predict their ED use is therefore an essential task to improve ED care as well as frequent users' health problems. Adequate statistical tools are needed to that end. Although they are numerous, no literature review has been published yet about statistical tools used for analysing ED frequent users. Therefore, the aim of our study was to draw up a list of statistical tools used in identifying variables associated with frequent use or predicting the risk of becoming a frequent user.

METHODS

In order to list the statistical tools used in describing variables associated with and prediction of frequent ED use, we conducted a scoping review. We followed the 5-stage methodology of Arksey and O'Malley[9] adapted by Levac *et al*.[10] The methodological framework of a scoping review allows 'mapping rapidly the key concepts underpinning a research area and the main sources and types of evidence available',[11] thus allowing us to identify gaps in the literature and future research opportunities.

Stage 1: Identifying the research question

We defined our research question as follows: What statistical tools are used in the identification of variables associated with frequent ED users and in their prediction?

Stage 2: Identifying relevant studies

We searched PubMed, CINAHL and Scopus databases in February 2019, using search strategies developed with the help of an information specialist (see the online supplementary appendix for the complete search strategy). Keywords included variants of 'frequent users', 'emergency departments' and 'statistical tools'.

There were no restriction regarding the population age or sex, health conditions, study period or country.

Stage 3: Study selection

Articles written in French or in English were included using the following criteria:

- The study must focus on frequent users of EDs (studies focusing on re-visits or on frequent visits other than in EDs were excluded).
- The study must have an explicit definition of frequent users, such as four visits in 1 year (reviews were excluded).
- The study must use at least one statistical tool that is classified as inferential (not descriptive, as defined by The Cambridge Dictionary of Statistics[12]), such as hypothesis tests, regression models, decision trees or others.
- The study's objectives must include identifying variables associated with frequent use or predicting the risk of becoming a frequent user.

We collected 4534 potential abstracts (figure 1). Of those, 32 were duplicates and 4344 were excluded by an investigator (YC) after reading the title and the abstract. At this stage, studies were discarded if it was explicit from the title and the abstract that they were unfit for the scoping review (for instance studies about frequent use of inpatient services, systematic reviews, etc). In case of uncertainty, studies were kept for complete reading. Then, YC and FRH or ID independently evaluated the remaining 158 full text articles, of which 109 matched the above criteria. A third evaluator was consulted in case of discrepancy. Reasons for exclusion were: not in French or English (1), duplicate (3), systematic review (4), no inferential statistics (5), no explicit definition of frequent users (5), focus not on ED (14), no description or prediction of frequent users (17). A reference search among the references of the 109 included articles yielded five relevant articles. Thus, 114 articles were included in this study, of which YC, ID and MB examined the full texts.

Stage 4: Charting the data

YC, MB and ID independently extracted the corresponding data. Reported characteristics were the first (two) author(s), the publication year, the study location, the population, the frequent users' definition, the objectives, the sample size and the statistical tools used concerning the research question.

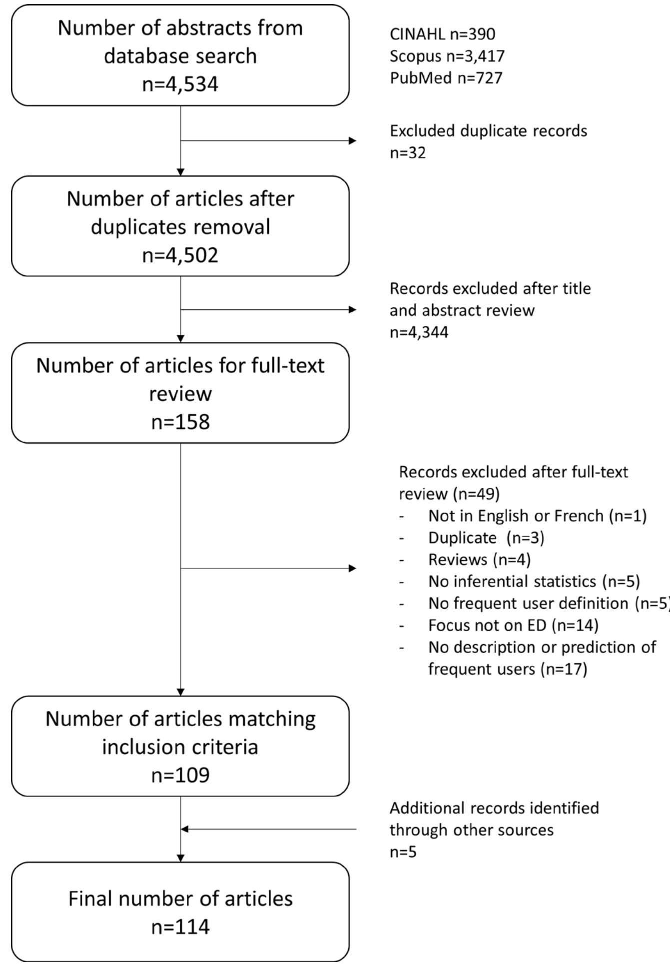

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. ED, emergency department.

Stage 5: Collating, summarising and reporting the results

The results are reported via a content analysis.[13]

Patient and public involvement

Patients or public were not involved in this study.

RESULTS

The studies' main characteristics are presented in table 1. Out of 114 studies, 65 were conducted in the USA, 17 in Canada and 8 in Australia (figure 2). The various statistical tools were classified into four main categories: regression, hypothesis testing, machine learning and other tools.

Regression

Regression tools consist of a set of processes aimed at quantifying the relationships between a dependent variable and other explanatory variables.[14] They are useful for description and prediction. Some regression models may be regularised, which in this case means avoiding

Table 1 Main characteristics of the 86 included studies

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Aagaard *et al* 2014[15] Denmark	Psychiatric	≥5 visits per year	To identify predictors of frequent use of a psychiatric ER.	8034	Logistic regression
Adams *et al* 2000[16] Australia	Adults with asthma	≥2 visits per year	To identify whether factors other than severity and low socioeconomic status were associated with this disproportionate use.	293	Logistic regression
Ahn *et al* 2018[91] Australia	General population aged≤70 years	≥4 visits per year	To examine the characteristics of frequent visitors to EDs and develop a predictive model to identify those with high risk of a future representations to ED among younger and general population.	170 134	Maximum likelihood monotone coarse classifier algorithm Logistic regression Mixed-effects model
Alghanim and Alomar 2015[17] Saudi Arabia	All	≥3 visits per year	To determine the prevalence of frequent use of EDs in public hospitals, to determine factors associated with such use, and to identify patients' reasons for frequent use.	666	X^2 test Logistic regression
Alpern *et al* 2014[18] USA	All	≥4 visits per year	To describe the epidemiology of and risk factors for recurrent and high frequency use of the ED by children.	695 188	Negative binomial regression Logistic regression Generalised estimating equations
Andren and Rosenqvist 1987[113] Sweden	All	≥4 visits per year	To follow a cohort of heavy ED users with regard to changes in medical and psycho-social profiles and ED use and to identify predictors for a maintained high use of ED services and the relationship between changes in access to social networks and utilisation of medical care services.	232	Decision trees Linear regression
Andrews *et al* 2018[92] USA	Medicaid enrollees with addiction	≥2 visits during a 2 year-period	To examine whether the number of outpatient addiction programmes accepting Medicaid in South Carolina counties is linked to repeat use of the ED for addiction-related conditions.	2401	Logistic regression
Arfken *et al* 2004[19] USA	Psychiatric	≥6 visits per year	To identify risk factors for people who use psychiatric emergency services repeatedly and to estimate their financial charges.	74	Logistic regression
Batra *et al* 2017[83] USA	Women	≥3 visits per 3 months	To use population data to identify patient characteristics associated with a postpartum maternal ED visit within 90 days of discharge after birth.	1 071 232	Logistic regression Receiver operating characteristic curve
Beck *et al* 2016[105] UK	Mental health	≥3 visits in 3 months	To statistically identify characteristics associated with a shorter time to re-attendance and a higher number of overall ED admissions with a Mental Health Liaison Service referral.	24 010	Cox regression Negative binomial regression
Bieler *et al* 2012[20] Switzerland	All	≥4 visits per year	To identify the social and medical factors associated with frequent ED use and to determine if frequent users were more likely to have a combination of these factors in a universal health insurance system.	719	Wilcoxon rank-sum test Logistic regression
Billings and Raven 2013[21] USA	All	≥3 visits per year ≥5 visits per year ≥8 visits per year ≥10 visits per year	To examine whether it is possible to predict who will become a frequent ED user with predictive modelling and to compare ED expenditures to total Medicaid services expenditures.	212 259	Logistic regression
Birmingham *et al* 2017[124] USA	All	≥4 visits per year	To characterise frequent ED users, including their reason for presenting to the ED and to identify perceived barriers to care from the users' perspective.	1523	t-test X^2 test Wilcoxon rank-sum test
Blair *et al* 2018[112] UK	Children	≥4 visits per year	To describe the sociodemographic and clinical characteristics of preschoolers who attend ED a large District General Hospital.	10 169	X^2 test Poisson regression Mann-Whitney U test
Blonigen *et al* 2017[82] USA	Veteran psychiatric	≥5 visits per year	To identify patient-level factors associated with ED use among veteran psychiatric patients and to examine factors associated with different subgroups of ED users including 'high utilisers'.	226 122	X^2 test Zero-truncated negative binomial regression Logit regression

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Boyer *et al* 2011[22] France	Psychiatric	≥6 visits per year	To examine characteristics of frequent visitors to a psychiatric emergency service in a French public teaching hospital over 6 years.	1285	Logistic regression
Brennan *et al* 2014[23] USA	Psychiatric	≥4 visits per year	To assess the incidence of psychiatric visits among frequent ED users and utilisation among frequent psychiatric users.	788 005	Kruskal-Wallis test Mann-Whitney U test Logistic regression
Buhumaid *et al* 2015[24] USA	Psychiatric	≥4 visits per year	To evaluate demographic factors associated with increased ED use among people with psychiatric conditions.	569	Logistic regression
Burner *et al* 2018[84] USA	People with diabetes	≥3 visits per 6 months	To describe characteristics of patients with poorly controlled diabetes who have high ED utilisation, and compare them with patients with lower ED utilisation.	108	Logistic regression
Cabey *et al* 2014[25] USA	All	90th percentile	To define the threshold and population factors associated with paediatric ED use above the norm during the first 36 months of life.	16 664	Non-parametric distribution fit Logistic regression Bootstrap Clopper-Pearson method
Castner *et al* 2015[26] USA	People with psychiatric and substance abuse diagnoses	≥3 visits per year	To stratify individuals by overall health complexity and examine the relationship of behavioural health diagnoses (psychiatric and substance abuse) as well as frequent treat-and-release ED utilisation in a cohort of Medicaid recipients.	56 491	Logistic regression
Chambers *et al* 2013[27] Canada	Homeless	90th percentile	To identify predictors of ED use among a population-based prospective cohort of homeless adults in Toronto, Ontario.	1165	Logistic regression
Chang *et al* 2014[28] USA	Psychiatric	≥4 visits per year or ≥3 visits during two consecutive months	To identify the patient characteristics associated with frequent ED use and develop a tool to predict risk for returning in the next month.	863	X^2 test Logistic regression
Christensen *et al* 2017[107] USA	All	≥4 visits per year	To determine the patient characteristics and healthcare utilisation patterns that predict frequent ED use (≥4 visits per year) over time to assist healthcare organisations in targeting patients for care management.	13 265	Zero-inflated Poisson regression Receiver operating characteristic curve
Chukmaitov *et al* 2012[29] USA	People with ambulatory care-sensitive conditions	≥4 visits per year	To study characteristics of all, occasional and frequent ED visits due to ambulatory care-sensitive conditions.	4 914 933 (number of visits)	Logistic regression
Colligan *et al* 2016[30] USA	Medicare beneficiaries	≥4 visits per year	To examine factors associated with persistent frequent ED use during a 2 year period among Medicare beneficiaries.	5 400 237	Logistic regression Wald test
Colligan *et al* 2017[96] USA	Medicare beneficiaries	≥4 visits per year	To examine factors related to frequent ED use in a large, nationally representative sample of Medicare beneficiaries.	5 778 038	X^2 test Analysis of variance Logistic regression Wald test
Cunningham *et al* 2017[97] USA	All	95th percentile ≥10 visits per year	To compare frequent and infrequent ED visitors' primary care utilisation and perceptions of primary care access, continuity and connectedness and to examine primary care utilisation and perceptions as predictors of ED use.	1113	t-test X^2 test Logistic regression
Das *et al* 2017[31] USA	Children with asthma	≥2 visits per year	To explore the predictability of frequent ED use among children with asthma using data from an EHR from one medical centre.	2691	Wilcoxon rank-sum test X^2 test LASSO logistic regression Regularised logistic regression Decision trees Random forests Support vector machines

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Doran *et al* 2013[33] USA	All	2–4 visits per year 5–10 visits per year 11–25 visits per year ≥25 visits	To identify sociodemographic and clinical factors most strongly associated with frequent ED use within the Veterans Health Administration nationally.	930712	Logistic regression
Doran *et al* 2014[32] USA	All	≥3 visits per year	To examine patients' reasons for using the ED for low-acuity health complaints, and determine whether reasons differed for frequent ED users versus non-frequent ED users.	940	Logistic regression
Doupe *et al* 2012[4] Canada	All	≥7 visits per year	To identify factors that define frequent and highly frequent ED users.	105687	Logistic regression Receiver operating characteristic curve
Fernandes *et al* 2003[34] Brazil	All	≥3 visits per year	To identify characteristics related to poor disease control and frequent visits to the ED to apply appropriate clinical management.	86	X^2 test Logistic regression
Flood *et al* 2017[85] USA	Children	≥4 visits per year	To identify factors associated with high ED utilisation among children in vulnerable families.	2631	X^2 test t-test Logistic regression
Freitag *et al* 2005[77] USA	People with chronic daily headache	≥3 visits per year	To examine the characteristics of chronic daily headache sufferers who use EDs and identify factors predictive of ED visits.	785	Wilcoxon rank-sum test t-test X^2 test Poisson regression Negative binomial regression Logistic regression
Friedman *et al* 2009[78] USA	People with severe headache	≥4 visits per year	To determine frequency of ED use and risk factors for use among patients suffering severe headache.	13451	Markov chain Monte Carlo imputation Logistic regression
Frost *et al* 2017[35] Canada	All	≥3 visits per year	To determine whether machine learning techniques using text from a family practice electronic medical record can be used to predict future high ED use and total costs by patients who are not yet high ED users or high cost to the healthcare system.	43111	Logistic regression
Girts *et al* 2002[114] USA	People with a diagnosis of psychosis	≥2 visits per 6 months	To develop a predictive model of ED utilisation for patients where a diagnosis of psychosis could be identified from a claim associated with a medical service provider visit.	764	t-test Linear regression
Grinspan *et al* 2015[36] USA	People with epilepsy	≥4 visits per year	To describe (1) the predictability of frequent ED use (a marker of inadequate disease control and/or poor access to care), and (2) the demographics, comorbidities and use of health services of frequent ED users, among people with epilepsy.	8041	X^2 test Logistic regression Regularised logistic regression Elastic net logistic regression Decision trees Random forests AdaBoost Support vector machines Receiver operating characteristic curve
Gruneir *et al* 2018[93] Canada	Nursing home residents	≥3 visits per year	To describe repeat ED visits over 1 year, identify risk factors for repeat use and characterise 'frequent' ED visitors.	25653	Logistic regression Andersen-Gill model
Hardie *et al* 2015[108] USA	All	≥4 visits per year	To describe frequent users of ED services in a rural community setting and the association between counts of patient's visits and discrete diagnoses.	1652	Poisson regression
Hasegawa *et al* 2014[37] USA	People with acute asthma	≥2 visits per year	To examine the proportion and patient characteristics of adult patients with multiple ED visits for acute asthma and the associated hospital charges.	86224	X^2 test Kruskal-Wallis test Logistic regression

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Hasegawa *et al* 2014[76] USA	People with acute heart failure syndrome	≥2 visits per year	To examine the proportion and characteristics of patients with frequent ED visits for acute heart failure syndrome and associated healthcare utilisation.	113 033	X^2 test Kruskal-Wallis test Negative binomial regression Linear regression
Hasegawa *et al* 2014[102] USA	People with chronic obstructive pulmonary disease	≥2 visits per year	To quantify the proportion and characteristics of patients with frequent ED visits for acute exacerbation of chronic obstructive pulmonary disease and associated healthcare utilisation.	98 280	X^2 test Kruskal-Wallis test Logistic regression Negative binomial regression Linear regression
Huang *et al* 2003[38] Taiwan	All	≥4 visits per year	To characterise frequent ED users and to identify the factors associated with frequent ED use in a hospital in Taiwan.	800	X^2 test Logistic regression
Hudon *et al* 2016[39] Canada	All	≥3 visits per year	To identify prospectively personal characteristics and experience of organisational and relational dimensions of primary care that predict frequent use of ED.	1769	Mixed-effects logistic regression
Hudon *et al* 2017[5] Canada	People with diabetes	≥3 visits for three consecutive years	To explore the factors associated with chronic frequent ED utilisation in a population with diabetes.	62 316	Logistic regression Decision trees
Hunt *et al* 2006[3] USA	All	≥4 visits per year	To identify frequent users of the ED and determine the characteristics of these patients.	49 603	Logistic regression
Huynh *et al* 2016[103] Canada	People with substance use disorders	≥4 visits per year	To assess the characteristics of individuals with substance use disorders according to their frequency of ED utilisation, and to examine which variables were associated with an increase in ED visits using Andersen's model.	4526	X^2 test Analysis of variance Negative binomial regression Generalised estimating equations
Kanzaria *et al* 2017[86] USA	Adults aged 18–55 years	≥4 visits per year	To examine the persistence of frequent ED use over an 11-year period, describe characteristics of persistent versus non-persistent frequent ED users, and identify predictors of persistent frequent ED use.	173 273	Logistic regression
Kerr *et al* 2005[40] Canada	Injection drug users	≥3 visits during the two past years	To examine rates of primary care and ER use among injection drug users and to identify correlates of frequent ED use.	883	X^2 test Wilcoxon signed-rank test t-test Logistic regression
Kidane *et al* 2018[98] Canada	Patients who received oesophagectomy	≥3 visits per year	To evaluate healthcare resource utilisation, specifically ED visits within 1 year of oesophagectomy, and to identify risk factors for ED visits and frequent ED use.	3344	t-test Wilcoxon rank-sum test Fisher exact tests Logistic regression
Kim *et al* 2018[125] Canada	All	99th percentile	To describe patient and visit characteristics for Canadian ED highly frequent users and patient subgroups with mental illness, substance misuse or ≥30 yearly ED visits.	261	t-test Wilcoxon rank-sum test
Kirby *et al* 2010[41] Australia	People with chronic disease	≥3 visits per year	To explore the link between frequent readmissions in chronic disease and patient-related factors.	15 806	X^2 test Logistic regression
Kirby *et al* 2011[42] Australia	All	≥4 visits per year	To identify the factors associated with frequent re-attendances in a regional hospital thereby highlighting possible solutions to the problem.	15 806	Kruskal-Wallis test X^2 test Logistic regression
Klein *et al* 2018[126] USA	Adults who present to the ED repeatedly for acute alcohol intoxication	≥20 visits per year	To describe frequent ED users who present to the ED repeatedly for acute alcohol intoxication and their ED encounters.	325	Difference in proportions test
Ko *et al* 2015[43] Taiwan	All	≥4 visits per year	To describe the distribution of the frequency of ED visits among ED users in 2010 and to evaluate the association of frequent ED use with various patient characteristics.	170 457	Logistic regression

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Ledoux and Minner 2006[44] Belgium	Psychiatric	≥4 visits per year	(1) To provide a naturalistic evaluation of patients repeating admissions in a psychiatric emergency ward (distinguishing between occasional repeaters and frequent repeaters), (2) to identify patients' characteristics that predict repeated use of a psychiatric ER and (3) to propose adapted treatment models.	2470	Mantel-Haenszel test Analysis of variance Logistic regression
Lee *et al* 2018[94] USA	Persons with systemic lupus erythematosus	≥3 visits per year	To identify lupus erythematosus patients who persistently frequented the ED over 4 years.	129	t-test X^2 test Fisher exact test Logistic regression
Legramante *et al* 2016[45] Italy	All	≥4 visits per year	To evaluate and characterise hospital visits of older patients (age 65 or greater) to the ED of a university teaching hospital in Rome, in order to identify clinical and social characteristics potentially associated with 'elderly frequent users'.	38016	t-test Logistic regression
Leporatti *et al* 2016[46] Italy	All	90th percentile ≥3 visits per year	To describe the characteristics of patients who frequently accessed accident and EDs located in the metropolitan area of Genoa.	147864	Zero-truncated negative binomial regression Logistic regression
Lim *et al* 2014[47] Singapore	People with asthma	≥4 visits per year	To describe the characteristics of frequent attenders who present themselves multiple times to the ED for asthma exacerbations.	155	t-test X^2 test Mann-Whitney U test Logistic regression
Limsrivilai *et al* 2017[48] USA	People with inflammatory bowel diseases	75th percentile of the annual medical charges	To identify predictive factors readily available in a standard electronic medical record to develop a multivariate model to predict the probability of inflammatory bowel diseases-related hospitalisation, ED visit and high total charges in the subsequent year.	1430	Receiver operating characteristic curve Logistic regression
Lin *et al* 2015[104] USA	Homeless people	≥3 visits per year	To examined factors associated with frequent hospitalisations and ED visits among Medicaid members who were homeless.	6494	X^2 test Analysis of variance Negative binomial regression
Liu *et al* 2013[49] USA	People with mental health, alcohol or drug-related diagnoses	≥4 visits per year	To determine whether frequent ED users are more likely to make at least one and a majority of visits for mental health, alcohol or drug-related complaints compared with non-frequent users.	65201	t-test X^2 test Logistic regression
Mandelberg *et al* 2000[50] USA	All	≥5 visits per year	To determine how the demographic, clinical and utilisation characteristics of frequent ED users differ from those of other ED patients.	43383	Logistic regression Survival analysis
Mann *et al* 2016[51] Canada	People with chronic pain	90th percentile	To investigate the role of chronic pain in healthcare visits and to document the frequency of healthcare visits and to identify characteristics associated with frequent visits.	1274	Logistic regression
Mann *et al* 2017[95] Canada	People with chronic pain	90th percentile	To describe factors associated with high clinic and ER use among individuals with chronic pain.	702	t-test Logistic regression
McMahon *et al* 2018[52] Ireland	All	≥4 visits per year	To examine the characteristics of the frequent ED attenders by age (under 65 and over 65 years).	19310	X^2 test Logistic regression
Meyer *et al* 2013[109] USA	Prisoners with Human Immunodeficiency Virus	≥2 visits per year	To characterise the medical, social and psychiatric correlates of frequent ED use among released prisoners with HIV.	151	t-test X^2 test Poisson regression
Milani *et al* 2016[53] USA	People with multimorbid chronic diseases	≥4 visits per year	To examine the association between multimorbid chronic disease and frequency ED visits in the past 6 months, by sex, in a community sample of adults from northern Florida.	7143	Breslow-Day test Logistic regression
Milbrett and Halm 2009[110] USA	All	≥6 visits per year	To describe the characteristics of patients who frequently use ED services and to determine factors most predictive of frequent ED use.	201	X^2 test Mann-Whitney U test Poisson regression

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Moe *et al* 2013[121] Canada	All	95th percentile	To develop uniform definitions, quantify ED burden and characterise adult frequent users of a suburban community ED.	14223	X^2 test Mann-Whitney U test
Mueller *et al* 2016[54] USA	Children with cancer	90th percentile ≥4visits per year	To (a) evaluate patient and ED encounter characteristics of frequent ED utilisers among children with cancer and (b) quantify healthcare services for frequent ED utilisers.	17943	X^2 test Logistic regression
Nambiar *et al* 2017[55] Australia	Adults who inject drugs	≥3visits per year	To describe demographic factors, patterns of substance use and previous health service use associated with frequent use of EDs in people who inject drugs.	612	Negative binomial regression Logistic regression
Nambiar *et al* 2018[106] Australia	Adults who inject drugs	≥3visits per year	To describe characteristics of state-wide ED presentations in a cohort of people who inject drugs, compare presentation rates to the general population and to examine characteristics associated with frequent ED use.	678	Negative-binomial regression Generalised estimating equations
Naseer *et al* 2018[87] Sweden	Older adults	≥4visits during a 4-year period	To assess the association of health related quality of life with time to first ED visit and/or frequent ED use in older adults during 4year period and if this association differs in 66–80 and 80+ age groups.	673	Cox proportional hazard model Logistic regression
Neufeld *et al* 2016[56] Canada	All	≥4visits per year	To describe factors predicting frequent ED use among rural older adults receiving home care services in Ontario, Canada.	12118	X^2 test Logistic regression
Neuman *et al* 2014[117] USA	All	≥4visits per year	To compare the characteristics and ED health services of children by their ED visit frequency.	1 896 547	Mantel-Haenszel test Receiver operating characteristic curve Generalised linear mixed-effects models
Ngamini-Ngui *et al* 2014[118] Canada	Patients with schizophrenia and a co-occurring substance use disorder	≥5visits per year	To assess factors associated over time with high use of EDs by Quebec patients who had schizophrenia and a co-occurring substance use disorder.	2921	Generalised estimating equations
Norman *et al* 2016[57] USA	All	≥4visits per year	To clearly define and describe characteristics of frequent EMS users in order to provide suggestions for efficient and cost-effective interventions that address the healthcare needs of these users.	539	Logistic regression
O'Toole *et al* 2007[79] USA	Substance users	≥3visits per year	To identify factors associated with 12month high frequency utilisation of ambulatory care, ED and inpatient medical care in a substance-using population.	326	t-test X^2 test Logistic regression
Palmer *et al* 2014[58] Canada	All	≥4visits per year	To determine if having a primary care provider is an important factor in frequency of ED use.	59803	X^2 test Wilcoxon rank-sum test Logistic regression
Panopalis *et al* 2010[59] USA	People with systemic lupus erythematosus	≥3visits per year	To describe characteristics of systemic lupus erythematosus patients who are frequent users of the ED and to identify predictors of frequent ED use.	807	One-way analysis of variance Logistic regression
Pasic *et al* 2005[80] USA	Psychiatric	2 SD above the mean number of visits ≥6visits per year ≥4visits in a quarter	To examine the sociodemographic and clinical characteristics of high utilisers of psychiatric emergency services.	17481	X^2 test Logistic regression
Paul *et al* 2010[60] Singapore	All	≥5visits per year	To determine factors associated with frequent ED attendance at an acute general hospital in Singapore.	82172	X^2 test Logistic regression
Peltz *et al* 2017[101] USA	Medicaid-insured children	≥4visits per year	To describe the characteristics of children who sustain high-frequency ED use over the following 2years.	470449	X^2 test Wilcoxon signed-rank test Logistic regression
Pereira *et al* 2016[61] USA	All	≥5visits per year	To develop machine learning models that can predict future ED utilisation of individual patients, using only information from the present and the past.	4 604 252	Decision trees AdaBoost Logistic regression

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Pines and Buford 2006[62] USA	People with asthma	90th percentile ≥3 visits per year	To determine socioeconomic and demographic factors that predict frequent ED use among asthmatics in southeastern Pennsylvania.	1799	t-test X^2 test Logistic regression
Quilty *et al* 2016[63] Australia	People without chronic health conditions	≥6 visits per year	To determine the clinical and environmental variables associated with frequent presentations by adult patients to a remote Australian hospital ED for reasons other than chronic health conditions.	273	t-test X^2 test Fisher exact tests Logistic regression
Rask *et al* 1998[81] USA	All	≥10 visits per 2 years	To describe primary care clinic use and emergency ED use for a cohort of public hospital patients seen in the ED, identify predictors of frequent ED use, and ascertain the clinical diagnoses of those with high rates of ED use.	351	X^2 test t-test Logistic regression
Rauch *et al* 2018[115] Germany	All	≥3 visits per year	To examine (1) what ambulatory care sensitive conditions are linked to frequent use, (2) how frequent users can be clustered into subgroups with respect to their diagnoses, acuity and admittance, and (3) whether frequent use is related to higher acuity or admission rate.	23 364	X^2 test t-test Linear regression Non-negative matrix factorisation
Sacamo *et al* 2018[111] USA	Persons with substance use	≥2 visits per 6 months	To examine associations of individuals and their social networks with high frequency ED use among persons reporting substance use.	653	Poisson regression
Samuels-Kalow *et al* 2017[64] USA	All	≥4 visits per year	To derive and test a predictive model for high frequency (four or more visits per year), low-acuity (emergency severity index 4 or 5) utilisation of the paediatric ED.	60 799 (number of visits)	Likelihood ratio test X^2 test Receiver operating characteristic curve Logistic regression
Samuels-Kalow *et al* 2018[88] USA	Patients with asthma exacerbation	≥4 visits per year	To create a predictive model to prospectively identify patients at risk of high-frequency ED utilisation for asthma and to examine how that model differed using state wide versus single-centre data.	254 132	X^2 test Fisher exact tests Wilcoxon rank-sum test Hosmer-Lemeshow test Receiver operating characteristic curve Logistic regression
Samuels-Kalow *et al* 2018[119] USA	Children	≥3 visits per year	To develop a population-based model for predicting Medicaid-insured children at risk for high frequency of low-resource-intensity ED visits.	743 016	X^2 test Receiver operating characteristic curve Logistic regression
Schlichting *et al* 2017[99] USA	Children	≥2 visits per year	To examine the utilisation of the ED by children with different forms of insurance and describe factors associated with repeat ED use and high reliance on the ED in a nationally representative sample of children in the USA.	47 926	Logistic regression
Schmoll *et al* 2015[65] France	Psychiatric	≥9 visits during the six past years	To describe demographic and clinical characteristics of frequent visitors to a psychiatric emergency ward in a French Academic hospital over 6 years in comparison to non-frequent visitors.	8800	t-test X^2 test Logistic regression
Soler *et al* 2004[66] Spain	People with chronic obstructive pulmonary disease	≥3 visits per year	To identify factors associated with frequent use of hospital services (emergency care and admissions) in patients with chronic obstructive pulmonary disease.	64	t-test X^2 test Kolmogorov-Smirnov test Mann-Whitney U test Logistic regression
Street *et al* 2018[123] Australia	Adults aged≥65 years	≥4 visits per year	To characterise older people who frequently use ED and compare patient outcomes with older non-frequent ED attenders.	21 073	X^2 test Wilcoxon rank-sum test Ordinal regression
Sun *et al* 2003[67] USA	All	≥4 visits per year	To identify predictors and outcomes associated with frequent ED users.	2333	Likelihood ratio test X^2 test Hosmer-Lemeshow test Logistic regression Bootstrap

Continued

Table 1 Continued

Authors, year and country	Population	Frequent user definition	Study main objectives	Study cohort size	Statistical tools used
Supat *et al* 2018[100] USA	Children	≥6 visits per year	To assess paediatric ED utilisation in California and to describe those identified as frequent ED users.	690 130	Logistic regression
Tangherlini *et al* 2010[68] USA	All	≥4 visits per year	To identify the factors that lead to increased use of EMS by patients≥65 years of age in an urban EMS system.	10 918	Kruskal-Wallis test X^2 test Logistic regression
Thakarar *et al* 2015[69] USA	Homeless	≥2 visits per year	To identify risk factors for frequent ER visits and to examine the effects of housing status and HIV serostatus on ER utilisation.	412	X^2 test Logistic regression
Vandyk *et al* 2014[70] Canada	Mental health	≥5 visits per year	To explore the population profile and associated socio demographic, clinical and service use factors of individuals who make frequent visits (5+ annually) to hospital EDs for mental health complaints.	536	Hosmer-Lemeshow test Logistic regression
Vinton *et al* 2014[71] USA	Chronic diseases and mental health	≥4 visits per year	To compare the characteristics of US adults by frequency of ED utilisation, specifically the prevalence of chronic diseases and outpatient primary care and mental health utilisation.	157 818	Logistic regression
Vu *et al* 2015[72] Switzerland	Mental health and substance users	≥4 visits per year	To determine the proportions of psychiatric and substance use disorders suffered by EDs' frequent users compared with the mainstream ED population, to evaluate how effectively these disorders were diagnosed in both groups of patients by ED physicians, and to determine if these disorders were predictive of a frequent use of ED services.	389	Fisher exact tests X^2 test Logistic regression
Wajnberg *et al* 2012[122] USA	All	≥4 visits over 6 months	To determine factors associated with frequent ED utilisation by older adults.	5718	X^2 test t-test
Watase *et al* 2015[73] Japan	Adults with asthma	≥2 visits per year	To characterise the adult patients who frequently presented to the ED for asthma exacerbation in Japan.	1002	One-way analysis of variance X^2 test Kruskal-Wallis test Logistic regression Negative binomial regression
Weidner *et al* 2018[89] USA	Patients with colorectal cancer	≥3 visits per year	To assess ED utilisation in patients with colorectal cancer to identify factors associated with ED visits and subsequent admission, as well as identify a high-risk subset of patients that could be targeted to reduce ED visits.	13 446	X^2 test t-test Logistic regression Negative binomial regression
Wong *et al* 2018[116] Singapore	Patients with cancer	≥4 visits per year	To identify factors associated with patients becoming ED frequent attenders after a cancer-related hospitalisation.	47 235	Cox regression Survival analysis
Woo *et al* 2016[74] Korea	All	≥4 visits per year	To understand whether the findings about frequent ED users in prior studies in the US healthcare system would be replicated in the Korean population, and whether these findings are independent of insurance status or ethnicity.	156 246	t-test X^2 test Linear regression Logistic regression
Wu *et al* 2016[75] USA	All	≥16 visits during the two past years	To assess the feasibility of using routinely gathered registration data to predict patients who will visit EDs with high frequency.	1 272 367	Logistic regression Receiver operating characteristic curve
Zook *et al* 2018[90] USA	Native American children	≥4 visits per year	To determine differences in ED use by Native American children in rural and urban settings and identify factors associated with frequent ED visits.	39 220	Logistic regression Hierarchical model Multiple imputations

ED, emergency department; EMS, emergency medical services; ER, emergency room.

Chiu Y, *et al*. *BMJ Open* 2019;**9**:e027750. doi:10.1136/bmjopen-2018-027750

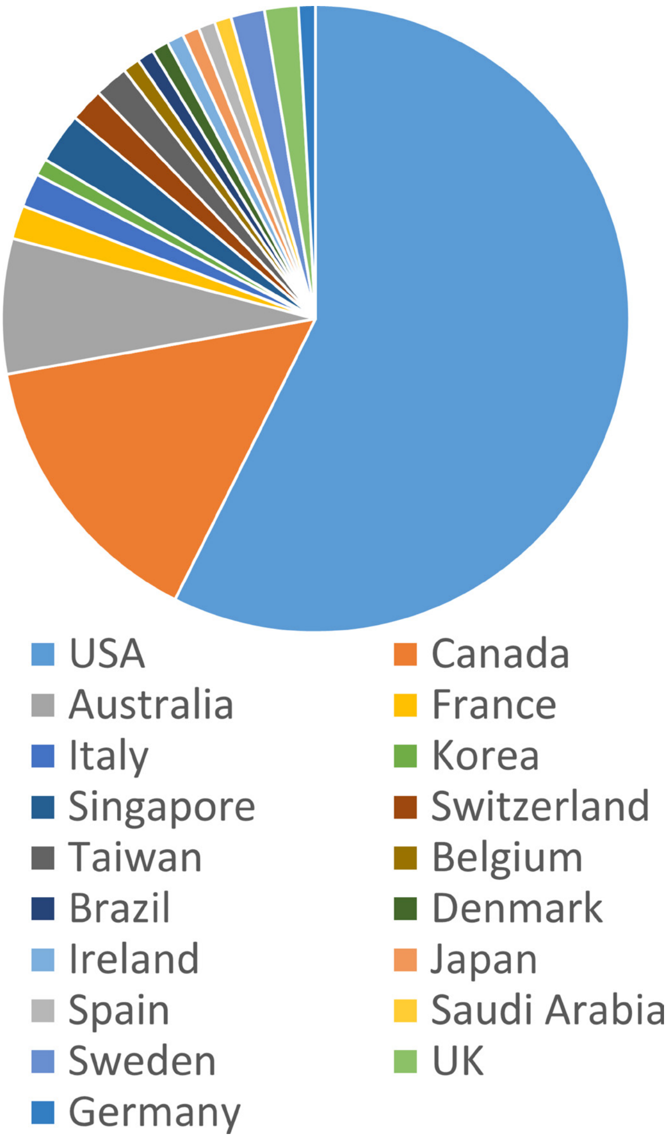

Figure 2 Number of studies by country.

overfitting with too many explanatory variables, or *zero-truncated*, which means that the model is not allowed to take null values.

Out of the four categories (regression, hypothesis testing, machine learning and other tools), the most reported tool was the logistic regression (90 studies,[3–5 15–101] two of which are regularised by LASSO or elastic net techniques), followed by the binomial regression (13 studies,[18 46 55 73 76 77 82 89 102–106] 2 of which are zero-truncated). To a lesser extent, the Poisson regression (seven studies,[77 107–112] one of which is zero-truncated), the linear regression (six studies[74 76 102 113–115]), the analysis of variance (six studies[44 59 73 96 103 104]), the Cox regression (four studies[87 93 105 116]) and hierarchical models (one study[90]) were also used. In those studies, the results are often associated with ORs. The mixed-effects models were mentioned three times.[39 91 117] Regression parameters were estimated by generalised estimating equations in four studies[18 103 106 118] while parameter confidence intervals were estimated by the bootstrap procedure (two studies[25 67]) and the Clopper-Pearson method (one study[25]). The receiver operating characteristic curve, or equivalently the sensitivity, specificity or area under the curve ('c-statistic'), was computed in 10 studies.[4 36 48 64 75 83 88 107 117 119] Finally, two studies performed imputation to account for missing data (Markov chain Monte Carlo and multiple imputations[78 90]).

Hypothesis testing

Statistical tests aim at testing a specific hypothesis about data and rely on probability distributions.[120] In the selected studies, the tests aimed mainly at comparing two samples (frequent users and non-frequent users).

The most common statistical tests were the χ^2 test (53 studies[17 28 31 34 36–38 40–42 47 49 52 54 56 58 60 62–69 72–74 76 77 79–82 85 88 89 94 96 97 101–104 109 110 112 115 119 121–124]) and the t-test (24 studies[40 45 47 49 62 63 65 66 74 77 79 81 85 89 94 95 97 98 109 114 115 122 124 125]) which measured association between variables or goodness-of-fit. As an alternative to the χ^2 test for association, five studies used the Fisher exact test.[63 72 94 98 119] Sample mean differences were assessed by 23 studies with the Mann-Whitney U test (also called the Wilcoxon rank-sum test[20 23 31 47 58 66 77 98 110 119 121 123–125]), its variant for dependent samples the Wilcoxon signed rank test,[40 101] or the Kruskal-Wallis test.[23 37 42 68 73 76 102] The difference in proportions test,[126] Mantel-Haenszel test (test for differences in contingency tables, two studies[44 117]), the likelihood ratio test (significance test for nested models, two studies[64 67]), the Hosmer-Lemeshow test (goodness-of-fit for logistic regression, two studies[67 70]), the Wald test (significance test for regression coefficients, two studies[30 96]) and the Breslow-Day test (test for homogeneity in contingency tables OR[53]) were also used to a lesser degree. Finally, one study checked the assumption of normality with the Kolmogorov-Smirnov test.[66]

Machine learning

Machine learning tools are a set of algorithms that can learn and adapt to data in order to classify or predict, for instance.[127] In the selected studies, the machine learning tools aimed mainly at classifying users (frequent vs non-frequent).

Two studies used random forests[31 36] along with support vector machines. Decision trees, which include classification and regression trees, were implemented by five studies.[5 31 36 61 113] Adaptive boosting, or AdaBoost, is a meta-algorithm that combines with other algorithms and helps for better performances. It was computed in two studies.[36 61]

Other tools

Two studies used survival analysis,[50 116] while another one fitted a non-parametric distribution to their data.[25] Finally, maximum likelihood monotone coarse classifier algorithm was used as a binning method[91] and non-negative matrix factorisation as a clustering technique.[115]

DISCUSSION

The most exploited statistical tools arguably came from regression analysis. This may be because regression is well established in medical statistics or also because it is the most natural tool when trying to find significant variables to explain a dependent variable (in this case, to be a frequent user). Moreover, it allows predicting easily the risk of a new user becoming a frequent user, depending on its covariates. Other tools from hypothesis testing or machine learning also proved to be popular, although to a much lesser extent. Combining these statistical techniques may help in discovering significant and complementary patterns, compared with using tools from one class only. In our scoping review, two studies mixed statistical tools from regression, hypothesis testing and machine learning.[31 36] In those studies, the author evaluated various performance criteria. While logistic regression performed well, other techniques such as random forests or LASSO regression were also competitive. Besides the fact that logistic regression can display modest performances,[128] random forests and LASSO regression can complete logistic regression. The first technique can be used to assess the importance of each independent variable in the model, while the second technique can be useful for automatic selection of features. Likewise, using a variety of statistical tools can help complete or confirm results obtained with established methodologies. Different tools from one class can also be mixed in order to achieve different stages of the analysis (for instance, different types of regression[82]).

The analysis of frequent ED users could benefit from using more machine learning techniques. Those were found to be not as common as regression or hypothesis testing, although they are especially appropriate when dealing with classification, prediction or big data. Tools such as support vector machines (which were used by two studies in this scoping review[31 36]), artificial neural networks or Bayesian networks are common classifiers and predictors in the artificial intelligence community.[129] They are popular for instance in cancer diagnostic and prognosis, which strongly rely on classification and prediction.[130–132] In particular, support vector machines, decision trees or self-organising maps can deal with binary outcomes, which is usually the case for frequent use outcomes. They usually require large datasets in order to overcome overfitting, but this is becoming less and less of an issue in health sciences.[133] Nevertheless, machine learning tools often use a black box approach as there are many intermediary steps leading to the final solution. While each step usually consists of simple arithmetic operations, their multiple interactions can be more difficult to interpret. In spite of this opacity, they still display good performances in classifying and predicting. In some cases, they may be more accurate than the widely used logistic regression.[134] Those methods would thus turn out to be less useful in data exploration.[135] Machine learning tools are getting popular in other fields in health sciences, such as critical care,[136] cardiology[137] or emergency medicine.[138] The authors state that their fields would benefit from this growing popularity, though results need to be analysed and interpreted in collaboration with clinicians.

Other tools exist that may also be suitable for describing the associated variables or the prediction of frequent ED users but were not reported in the literature. Among those, principal component analysis (PCA) is a dimensional reduction and visualisation technique, sometimes used with cluster or discriminant analysis.[139] Based on all the original explanatory variables, PCA constructs new ones by summing and weighing them differently. More weight is given to relevant variables so that those latter become dominant in the new constructions while still including all variables. For instance, Burgel *et al*[140] built chronic obstructive pulmonary disease clinical phenotypes by constructing new relevant variables with PCA and by grouping similar subjects in this new space with cluster analysis.[140] Moreover, PCA has already been used for the construction of questionnaires and diagnosis tools in a medical context,[141 142] both of which can prove useful in the identification of frequent users.

As mentioned, regression techniques were common in the selected studies. Yet, quantile regression (QR)[143] was not mentioned. QR is a generalisation of mean regression in the sense that its focus is not only the mean of the dependent variable distribution (such as in classical linear regression) but any quantile of it. QR thus represents an alternative to define frequent users by the high quantiles of ED visit distribution (eg, the 90th quantile). Eight studies[25 27 46 48 51 54 62 121] defined frequent users with quantiles, but they did not use QR. QR would allow for finer investigations in the different quantiles of ED users in relationship to the explanatory variables. For instance, the association between age and the number of ED visits may be significantly different across the 10th (low users) and 90th (frequent users) quantiles. Such a heterogeneous association would be uncovered by QR, while usually unseen with a classical mean regression. Ding *et al*[144] used QR to characterise waiting room and treatment times in EDs.[144] They explored the lowest, median and highest of those times and highlighted predictors that were significant only in particular quantiles. Usually, QR requires a continuous dependent variable as opposed to a logistic regression, though it is possible to combine these two regressions.[145] Furthermore, defining frequent users by quantiles would allow for better comparison between studies as there is no common definition for frequent users.

Strengths and limitations

To the best of our knowledge, this scoping review is the first to list statistical tools that are used in the identification of variables associated with frequent ED use and the prediction of frequent users. Besides, it was conducted following a well-defined methodological framework. The search strategies were designed with an information specialist in three different databases. Two independent evaluators selected the articles and extracted the data while a third independent evaluator settled disagreements, ensuring that all included studies were relevant. One limitation of our study is that quality assessment is not performed in a scoping review. However, this should not alter the results, since the aim was to list which statistical tools have been applied in the literature. Moreover, the majority of articles were in English which may introduce a selection bias (for instance, one excluded article was in Spanish). More than half of the reviewed studies were indeed conducted in the USA, making the results difficult to compare to other countries.

CONCLUSIONS

Frequent ED users represent a complex issue, and their analysis require adequate statistical tools. In this context, this scoping review shows that some tools are well established, such as logistic regression and χ^2 test, while others such as support vector machines are less so, though they would deserve to get more attention. It also outlines some research opportunities with other tools not yet explored.

Acknowledgements We would like to thank information specialist Josée Toulouse for her help in defining the search strategies and Tina Wey (PhD) for revising the text.

Contributors YC and CH designed the study with FR-H, ID and AV. YC, ID, CH and MB collected and analysed the data. YC and CH wrote the first draft of the manuscript. FR-H, ID, AV, M-CC and MB contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Funding This work was financed by grants from the Fonds de recherche du Québec – Santé and the Centre de recherche du Centre hospitalier universitaire de Sherbrooke. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement There are no unpublished additional data from the study.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1. 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;44:717–29.
2. LaCalle E, Rabin E. Frequent users of emergency departments: the myths, the data, and the policy implications. *Ann Emerg Med* 2010;56:42–8.
3. Hunt KA, Weber EJ, Showstack JA, *et al*. Characteristics of frequent users of emergency departments. *Ann Emerg Med* 2006;48:1–8.
4. 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;60:24–32.
5. Hudon C, Courteau J, Krieg C, *et al*. Factors associated with chronic frequent emergency department utilization in a population with diabetes living in metropolitan areas: a population-based retrospective cohort study. *BMC Health Serv Res* 2017;17:525.
6. Krieg C, Hudon C, Chouinard MC, *et al*. Individual predictors of frequent emergency department use: a scoping review. *BMC Health Serv Res* 2016;16:594.
7. Ruger JP, Richter CJ, Spitznagel EL, *et al*. Analysis of costs, length of stay, and utilization of emergency department services by frequent users: implications for health policy. *Acad Emerg Med* 2004;11:1311–7.
8. Bodenheimer T, Berry-Millett R. Care management of patients with complex health care needs. *Policy* 2009;1:6.
9. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8:19–32.
10. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010;5:69.
11. Mays N, Roberts E, Popay J. Synthesising research evidence. Studying the organisation and delivery of health services. *Research methods* 2001:188–220.
12. Everitt BS, Skrondal A. *The Cambridge Dictionary of Statistics*. 4th edn. Cambridge: Cambridge University Press Cambridge, 2010.
13. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nurs Health Sci* 2013;15:398–405.
14. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2 edn. New York: Springer International Publishing, 2015.
15. Aagaard J, Aagaard A, Buus N. Predictors of frequent visits to a psychiatric emergency room: a large-scale register study combined with a small-scale interview study. *Int J Nurs Stud* 2014;51:1003–13.
16. Adams RJ, Smith BJ, Ruffin RE. Factors associated with hospital admissions and repeat emergency department visits for adults with asthma. *Thorax* 2000;55:566–73.
17. Alghanim SA, Alomar BA. Frequent use of emergency departments in Saudi public hospitals: Implications for primary health care services. *Asia-Pac J Public Health* 2015;27:NP2521–NP30.
18. Alpern ER, Clark AE, Alessandrini EA, *et al*. Pediatric Emergency Care Applied Research Network (PECARN). Recurrent and high-frequency use of the emergency department by pediatric patients. *Acad Emerg Med* 2014;21:365–73.
19. Arfken CL, Zeman LL, Yeager L, *et al*. Case-control study of frequent visitors to an urban psychiatric emergency service. *Psychiatr Serv* 2004;55:295–301.
20. Bieler G, Paroz S, Faouzi M, *et al*. Social and medical vulnerability factors of emergency department frequent users in a universal health insurance system. *Acad Emerg Med* 2012;19:63–8.
21. Billings J, Raven MC. Dispelling an urban legend: frequent emergency department users have substantial burden of disease. *Health Aff* 2013;32:2099–108.
22. Boyer L, Dassa D, Belzeaux R, *et al*. Frequent visits to a French psychiatric emergency service: diagnostic variability in psychotic disorders. *Psychiatr Serv* 2011;62:966–70.
23. Brennan JJ, Chan TC, Hsia RY, *et al*. Emergency department utilization among frequent users with psychiatric visits. *Acad Emerg Med* 2014;21:1015–22.
24. Buhumaid R, Riley J, Sattarian M, *et al*. Characteristics of frequent users of the emergency department with psychiatric conditions. *J Health Care Poor Underserved* 2015;26:941–50.
25. Cabey WV, MacNeill E, White LN, *et al*. Frequent pediatric emergency department use in infancy and early childhood. *Pediatr Emerg Care* 2014;30:710–7.
26. Castner J, Wu YW, Mehrok N, *et al*. Frequent emergency department utilization and behavioral health diagnoses. *Nurs Res* 2015;64:3–12.
27. Chambers C, Chiu S, Katic M, *et al*. High utilizers of emergency health services in a population-based cohort of homeless adults. *Am J Public Health* 2013;103(S2):S302–10.
28. Chang G, Weiss AP, Orav EJ, *et al*. Predictors of frequent emergency department use among patients with psychiatric illness. *Gen Hosp Psychiatry* 2014;36:716–20.
29. Chukmaitov AS, Tang A, Carretta HJ, *et al*. Characteristics of all, occasional, and frequent emergency department visits due to ambulatory care-sensitive conditions in Florida. *J Ambul Care Manage* 2012;35:149–58.
30. Colligan EM, Pines JM, Colantuoni E, *et al*. Risk Factors for Persistent Frequent Emergency Department Use in Medicare Beneficiaries. *Ann Emerg Med* 2016;67:721–9.
31. Das LT, Abramson EL, Stone AE, *et al*. Predicting frequent emergency department visits among children with asthma using EHR data. *Pediatr Pulmonol* 2017;52:880–90.
32. Doran KM, Colucci AC, Wall SP, *et al*. Reasons for emergency department use: do frequent users differ? *Am J Manag Care* 2014;20:e506–e14.
33. Doran KM, Raven MC, Rosenheck RA. What drives frequent emergency department use in an integrated health system?

National data from the Veterans Health Administration. *Ann Emerg Med* 2013;62:151–9.
34. Fernandes AK, Mallmann F, Steinhorst AM, *et al*. Characteristics of acute asthma patients attended frequently compared with those attended only occasionally in an emergency department. *J Asthma* 2003;40:683–90.
35. Frost DW, Vembu S, Wang J, *et al*. Using the Electronic Medical Record to Identify Patients at High Risk for Frequent Emergency Department Visits and High System Costs. *Am J Med* 2017;130:601.e17–601.e22.
36. Grinspan ZM, Shapiro JS, Abramson EL, *et al*. Predicting frequent ED use by people with epilepsy with health information exchange data. *Neurology* 2015;85:1031–8.
37. Hasegawa K, Tsugawa Y, Brown DF, *et al*. A population-based study of adults who frequently visit the emergency department for acute asthma. California and Florida, 2009-2010. *Ann Am Thorac Soc* 2014;11:158–66.
38. Huang JA, Tsai WC, Chen YC, *et al*. Factors associated with frequent use of emergency services in a medical center. *J Formos Med Assoc* 2003;102:222–8.
39. Hudon C, Sanche S, Haggerty JL. Personal Characteristics and Experience of Primary Care Predicting Frequent Use of Emergency Department: A Prospective Cohort Study. *PLoS One* 2016;11:e0157489.
40. Kerr T, Wood E, Grafstein E, *et al*. High rates of primary care and emergency department use among injection drug users in Vancouver. *J Public Health* 2005;27:62–6.
41. Kirby SE, Dennis SM, Jayasinghe UW, *et al*. Patient related factors in frequent readmissions: the influence of condition, access to services and patient choice. *BMC Health Serv Res* 2010;10:216.
42. Kirby SE, Dennis SM, Jayasinghe UW, *et al*. Frequent emergency attenders: is there a better way? *Aust Health Rev* 2011;35:462–7.
43. Ko M, Lee Y, Chen C, *et al*. Prevalence of and Predictors for Frequent Utilization of Emergency Department: A Population-Based Study. *Medicine* 2015;94:e1205.
44. Ledoux Y, Minner P. Occasional and frequent repeaters in a psychiatric emergency room. *Soc Psychiatry Psychiatr Epidemiol* 2006;41:115–21.
45. Legramante JM, Morciano L, Lucaroni F, *et al*. Frequent use of emergency departments by the elderly population when continuing care is not well established. *PLoS One* 2016;11:e0165939.
46. Leporatti L, Ameri M, Trinchero C, *et al*. Targeting frequent users of emergency departments: Prominent risk factors and policy implications. *Health Policy* 2016;120:462–70.
47. Lim SF, Wah W, Pasupathi Y, *et al*. Frequent attenders to the ED: patients who present with repeated asthma exacerbations. *Am J Emerg Med* 2014;32:895–9.
48. Limsrivilai J, Stidham RW, Govani SM, *et al*. Factors That Predict High Health Care Utilization and Costs for Patients With Inflammatory Bowel Diseases. *Clin Gastroenterol Hepatol* 2017;15:385–92.
49. Liu SW, Nagurney JT, Chang Y, *et al*. Frequent ED users: are most visits for mental health, alcohol, and drug-related complaints? *Am J Emerg Med* 2013;31:1512–5.
50. Mandelberg JH, Kuhn RE, Kohn MA. Epidemiologic analysis of an urban, public emergency department's frequent users. *Acad Emerg Med* 2000;7:637–46.
51. Mann EG, Johnson A, VanDenKerkhof EG. Frequency and characteristics of healthcare visits associated with chronic pain: results from a population-based Canadian study. *Can J Anaesth* 2016;63:411–41.
52. McMahon CG, Power Foley M, Robinson D, *et al*. High prevalence of frequent attendance in the over 65s. *Eur J Emerg Med* 2018;25:1.
53. Milani SA, Crooke H, Cottler LB, *et al*. Sex differences in frequent ED use among those with multimorbid chronic diseases. *Am J Emerg Med* 2016;34:2127–31.
54. Mueller EL, Hall M, Carroll AE, *et al*. Frequent Emergency Department Utilizers Among Children with Cancer. *Pediatr Blood Cancer* 2016;63:859–64.
55. Nambiar D, Stoové M, Dietze P. Frequent emergency department presentations among people who inject drugs: A record linkage study. *Int J Drug Policy* 2017;44:115–20.
56. Neufeld E, Viau KA, Hirdes JP, *et al*. Predictors of frequent emergency department visits among rural older adults in Ontario using the Resident Assessment Instrument-Home Care. *Aust J Rural Health* 2016;24:115–22.
57. Norman C, Mello M, Choi B. Identifying Frequent Users of an Urban Emergency Medical Service Using Descriptive Statistics and Regression Analyses. *West J Emerg Med* 2016;17:39–45.
58. Palmer E, Leblanc-Duchin D, Murray J, *et al*. Emergency department use: is frequent use associated with a lack of primary care provider? *Can Fam Physician* 2014;60:e223–e9.
59. Panopalis P, Gillis JZ, Yazdany J, *et al*. Frequent use of the emergency department among persons with systemic lupus erythematosus. *Arthritis Care Res* 2010;62:401–8.
60. Paul P, Heng BH, Seow E, *et al*. Predictors of frequent attenders of emergency department at an acute general hospital in Singapore. *Emerg Med J* 2010;27:843–8.
61. Pereira M, Singh V, Hon CP, Greg McKelvey T, Sushmita S, De Cock M, *et al*. eds. *Predicting future frequent users of emergency departments in California state2016*: Association for Computing Machinery, Inc.
62. Pines JM, Buford K. Predictors of frequent emergency department utilization in Southeastern Pennsylvania. *J Asthma* 2006;43:219–23.
63. Quilty S, Shannon G, Yao A, *et al*. Factors contributing to frequent attendance to the emergency department of a remote Northern Territory hospital. *Med J Aust* 2016;204:111–7.
64. Samuels-Kalow ME, Bryan MW, Shaw KN, *et al*. Low-Acuity Utilization of the Pediatric Emergency Department. *Acad Pediatr* 2017;17:256–60.
65. Schmoll S, Boyer L, Henry JM, *et al*. [Frequent visitors to psychiatric emergency service: Demographical and clinical analysis]. *Encephale* 2015;41:123–9.
66. Soler JJ, Sánchez L, Román P, *et al*. Risk factors of emergency care and admissions in COPD patients with high consumption of health resources. *Respir Med* 2004;98:318–29.
67. Sun BC, Burstin HR, Brennan TA. Predictors and outcomes of frequent emergency department users. *Acad Emerg Med* 2003;10:320–8.
68. Tangherlini N, Pletcher MJ, Covec MA, *et al*. Frequent use of emergency medical services by the elderly: a case-control study using paramedic records. *Prehosp Disaster Med* 2010;25:258–64.
69. Thakarar K, Morgan JR, Gaeta JM, *et al*. Predictors of Frequent Emergency Room Visits among a Homeless Population. *PLoS One* 2015;10:e0124552.
70. Vandyk AD, VanDenKerkhof EG, Graham ID, *et al*. Profiling frequent presenters to the emergency department for mental health complaints: socio-demographic, clinical, and service use characteristics. *Arch Psychiatr Nurs* 2014;28:420–5.
71. Vinton DT, Capp R, Rooks SP, *et al*. Frequent users of US emergency departments: characteristics and opportunities for intervention. *Emerg Med J* 2014;31:526–32.
72. Vu F, Daeppen JB, Hugli O, *et al*. Screening of mental health and substance users in frequent users of a general Swiss emergency department. *BMC Emerg Med* 2015;15:27.
73. Watase H, Hagiwara Y, Chiba T, *et al*. Japanese Emergency Medicine Network Investigators. Multicentre observational study of adults with asthma exacerbations: who are the frequent users of the emergency department in Japan? *BMJ Open* 2015;5:e007435.
74. Woo JH, Grinspan Z, Shapiro J, *et al*. Frequent Users of Hospital Emergency Departments in Korea Characterized by Claims Data from the National Health Insurance: A Cross Sectional Study. *PLoS One* 2016;11:e0147450.
75. Wu J, Grannis SJ, Xu H, *et al*. A practical method for predicting frequent use of emergency department care using routinely available electronic registration data. *BMC Emerg Med* 2016;16:12.
76. Hasegawa K, Tsugawa Y, Tsai CL, *et al*. Frequent utilization of the emergency department for acute exacerbation of chronic obstructive pulmonary disease. *Respir Res* 2014;15:40.
77. Freitag FG, Kozma CM, Slaton T, *et al*. Characterization and prediction of emergency department use in chronic daily headache patients. *Headache* 2005;45:891–8.
78. Friedman BW, Serrano D, Reed M, *et al*. Use of the emergency department for severe headache. A population-based study. *Headache* 2009;49:21–30.
79. O'Toole TP, Pollini R, Gray P, *et al*. Factors identifying high-frequency and low-frequency health service utilization among substance-using adults. *J Subst Abuse Treat* 2007;33:51–9.
80. Pasic J, Russo J, Roy-Byrne P. High utilizers of psychiatric emergency services. *Psychiatr Serv* 2005;56:678–84.
81. Rask KJ, Williams MV, McNagny SE, *et al*. Ambulatory health care use by patients in a public hospital emergency department. *J Gen Intern Med* 1998;13:614–20.
82. Blonigen DM, Macia KS, Bi X, *et al*. Factors associated with emergency department useamong veteran psychiatric patients. *Psychiatr Q* 2017;88:721–32.
83. Batra P, Fridman M, Leng M, *et al*. Emergency Department Care in the Postpartum Period: California Births, 2009-2011. *Obstet Gynecol* 2017;130:1073–81.

84. Burner E, Ruiz A, Sanchez A, *et al*. 155 Insulin Use Predicts High Emergency Department Utilization Among Patients With Poorly Controlled Diabetes. *Ann Emerg Med* 2018;72:S65–S.
85. Flood C, Sheehan K, Crandall M. Predictors of Emergency Department Utilization Among Children in Vulnerable Families. *Pediatr Emerg Care* 2017;33:765–9.
86. Kanzaria HK, Niedzwiecki MJ, Montoy JC, *et al*. Persistent Frequent Emergency Department Use: Core Group Exhibits Extreme Levels Of Use For More Than A Decade. *Health Aff* 2017;36:1720–8.
87. Naseer M, Dahlberg L, Fagerström C. Health related quality of life and emergency department visits in adults of age ≥ 66 years: a prospective cohort study. *Health Qual Life Outcomes* 2018;16:144.
88. Samuels-Kalow M, Peltz A, Rodean J, *et al*. Predicting Low-Resource-Intensity Emergency Department Visits in Children. *Acad Pediatr* 2018;18:297–304.
89. Weidner TK, Kidwell JT, Etzioni DA, *et al*. Factors Associated with Emergency Department Utilization and Admission in Patients with Colorectal Cancer. *J Gastrointest Surg* 2018;22:913–20.
90. Zook HG, Kharbanda AB, Puumala SE, *et al*. Emergency Department Utilization by Native American Children. *Pediatr Emerg Care* 2018;34:802–9.
91. Ahn E, Kim J, Rahman K, *et al*. Development of a risk predictive scoring system to identify patients at risk of representation to emergency department: a retrospective population-based analysis in Australia. *BMJ Open* 2018;8:e021323.
92. Andrews CM, Westlake M, Wooten N. Availability of Outpatient Addiction Treatment and Use of Emergency Department Services Among Medicaid Enrollees. *Psychiatr Serv* 2018;69:729–32.
93. Gruneir A, Cigsar C, Wang X, *et al*. Repeat emergency department visits by nursing home residents: a cohort study using health administrative data. *BMC Geriatr* 2018;18:157.
94. Lee J, Lin J, Suter LG, *et al*. Persistently Frequent Emergency Department Utilization among Persons with Systemic Lupus Erythematosus. *Arthritis Care Res* 2018 (Epub 2018/10/09).
95. Mann EG, Johnson A, Gilron I, *et al*. Pain Management Strategies and Health Care Use in Community-Dwelling Individuals Living with Chronic Pain. *Pain Med* 2017;18:2267–79.
96. Colligan EM, Pines JM, Colantuoni E, *et al*. Factors Associated With Frequent Emergency Department Use in the Medicare Population. *Med Care Res Rev* 2017;74:311–27.
97. Cunningham A, Mautner D, Ku B, *et al*. Frequent emergency department visitors are frequent primary care visitors and report unmet primary care needs. *J Eval Clin Pract* 2017;23:567–73.
98. Kidane B, Jacob B, Gupta V, *et al*. Medium and long-term emergency department utilization after oesophagectomy: a population-based analysis. *Eur J Cardiothorac Surg* 2018;54:683–8.
99. Schlichting LE, Rogers ML, Gjelsvik A, *et al*. Pediatric Emergency Department Utilization and Reliance by Insurance Coverage in the United States. *Acad Emerg Med* 2017;24:1483–90.
100. Supat B, Brennan JJ, Vilke GM, *et al*. Characterizing pediatric high frequency users of California emergency departments. *Am J Emerg Med* 2018.
101. Peltz A, Samuels-Kalow ME, Rodean J, *et al*. Characteristics of Children Enrolled in Medicaid With High-Frequency Emergency Department Use. *Pediatrics* 2017;140:e20170962.
102. Hasegawa K, Tsugawa Y, Camargo CA, *et al*. Frequent utilization of the emergency department for acute heart failure syndrome: a population-based study. *Circ Cardiovasc Qual Outcomes* 2014;7:735–42.
103. Huynh C, Ferland F, Blanchette-Martin N, *et al*. Factors Influencing the Frequency of Emergency Department Utilization by Individuals with Substance Use Disorders. *Psychiatr Q* 2016;87:713–28.
104. Lin WC, Bharel M, Zhang J, *et al*. Frequent Emergency Department Visits and Hospitalizations Among Homeless People With Medicaid: Implications for Medicaid Expansion. *Am J Public Health* 2015;105:S716–22.
105. Beck A, Sanchez-Walker E, Evans LJ, *et al*. Characteristics of people who rapidly and frequently reattend the emergency department for mental health needs. *Eur J Emerg Med* 2016;23:351–5.
106. Nambiar D, Spelman T, Stoové M, *et al*. Are People Who Inject Drugs Frequent Users of Emergency Department Services? A Cohort Study (2008-2013). *Subst Use Misuse* 2018;53:457–65.
107. Christensen EW, Kharbanda AB, Velden HV, *et al*. Predicting Frequent Emergency Department Use by Pediatric Medicaid Patients. *Popul Health Manag* 2017;20:208–15.
108. Hardie TL, Polek C, Wheeler E, *et al*. Characterising emergency department high-frequency users in a rural hospital. *Emerg Med J* 2015;32:21–5.
109. Meyer JP, Qiu J, Chen NE, *et al*. Frequent emergency department use among released prisoners with human immunodeficiency virus: characterization including a novel multimorbidity index. *Acad Emerg Med* 2013;20:79–88.
110. Milbrett P, Halm M. Characteristics and predictors of frequent utilization of emergency services. *J Emerg Nurs* 2009;35:191–8.
111. Sacamano P, Krawczyk N, Latkin C. Emergency Department Visits in a Cohort of Persons with Substance Use: Incorporating the Role of Social Networks. *Subst Use Misuse* 2018;53:2265–9.
112. Blair M, Poots AJ, Lim V, *et al*. Preschool children who are frequent attenders in emergency departments: an observational study of associated demographics and clinical characteristics. *Arch Dis Child* 2018;103.
113. Genell Andrén K, Rosenqvist U. Heavy users of an emergency department--a two year follow-up study. *Soc Sci Med* 1987;25:825–31.
114. Girts TK, Crawford AG, Goldfarb NI, *et al*. Predicting High Utilization of Emergency Department Services Among Patients with a Diagnosis of Psychosis in a Medicaid Managed Care Organization. *Disease Management* 2002;5:189–96.
115. Rauch J, Hüsers J, Babitsch B, *et al*. Understanding the Characteristics of Frequent Users of Emergency Departments: What Role Do Medical Conditions Play? *Stud Health Technol Inform* 2018;253:175–9.
116. Wong TH, Lau ZY, Ong WS, *et al*. Cancer patients as frequent attenders in emergency departments: A national cohort study. *Cancer Med* 2018;7:4434–46.
117. Neuman MI, Alpern ER, Hall M, *et al*. Characteristics of recurrent utilization in pediatric emergency departments. *Pediatrics* 2014;134:e1025–e31.
118. Ngamini-Ngui A, Fleury MJ, Moisan J, *et al*. High users of emergency departments in Quebec among patients with both schizophrenia and a substance use disorder. *Psychiatr Serv* 2014;65:1389–91.
119. Samuels-Kalow ME, Faridi MK, Espinola JA, *et al*. Comparing Statewide and Single-center Data to Predict High-frequency Emergency Department Utilization Among Patients With Asthma Exacerbation. *Acad Emerg Med* 2018;25:657–67.
120. Altman DG. *Practical statistics for medical research*. London: CRC press, 1990.
121. Moe J, Bailey AL, Oland R, *et al*. Defining, quantifying, and characterizing adult frequent users of a suburban Canadian emergency department. *CJEM* 2013;15:214–26.
122. Wajnberg A, Hwang U, Torres L, *et al*. Characteristics of frequent geriatric users of an urban emergency department. *J Emerg Med* 2012;43:376–81.
123. Street M, Berry D, Considine J. Frequent use of emergency departments by older people: a comparative cohort study of characteristics and outcomes. *Int J Qual Health Care* 2018;30:624–9.
124. Birmingham LE, Cochran T, Frey JA, *et al*. Emergency department use and barriers to wellness: a survey of emergency department frequent users. *BMC Emerg Med* 2017;17:16.
125. Kim JJ, Kwok ESH, Cook OG, *et al*. Characterizing Highly Frequent Users of a Large Canadian Urban Emergency Department. *West J Emerg Med* 2018;19:926–33.
126. Klein LR, Martel ML, Driver BE, *et al*. Emergency Department Frequent Users for Acute Alcohol Intoxication. *West J Emerg Med* 2018;19:398–402.
127. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med* 2001;23:89–109.
128. Hu X, Barnes S, Bjarnadóttir M, *et al*. Intelligent selection of frequent emergency department patients for case management: A machine learning framework based on claims data. *IISE Trans Healthc Syst Eng* 2017;7:130–43.
129. Liao S-H, Chu P-H, Hsiao P-Y. Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Syst Appl* 2012;39:11303–11.
130. Wang S, Summers RM. Machine learning and radiology. *Med Image Anal* 2012;16:933–51.
131. Kourou K, Exarchos TP, Exarchos KP, *et al*. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J* 2015;13:8–17.
132. Ramos-Pollán R, Guevara-López MA, Suárez-Ortega C, *et al*. Discovering mammography-based machine learning classifiers for breast cancer diagnosis. *J Med Syst* 2012;36:2259–69.
133. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* 2013;309:1351–2.
134. Churpek MM, Yuen TC, Winslow C, *et al*. Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards. *Crit Care Med* 2016;44:368–74.

135. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. New York: Springer, 2009.
136. Sanchez-Pinto LN, Luo Y, Churpek MM. Big Data and Data Science in Critical Care. *Chest* 2018;154:1239–48.
137. Johnson KW, Torres Soto J, Glicksberg BS, *et al*. Artificial Intelligence in Cardiology. *J Am Coll Cardiol* 2018;71:2668–79.
138. Taylor RA, Pare JR, Venkatesh AK, *et al*. Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach. *Acad Emerg Med* 2016;23:269–78.
139. Jolliffe IT. *Principal Component Analysis and Factor Analysis. Principal component analysis*: Springer, 1986:115–28.
140. Burgel PR, Paillasseur JL, Caillaud D, *et al*. Initiatives BPCO Scientific Committee. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. *Eur Respir J* 2010;36:531–9.
141. Gordon DB, Polomano RC, Pellino TA, *et al*. Revised American Pain Society Patient Outcome Questionnaire (APS-POQ-R) for quality improvement of pain management in hospitalized adults: preliminary psychometric evaluation. *J Pain* 2010;11:1172–86.
142. Gasquet I, Villeminot S, Estaquio C, *et al*. Construction of a questionnaire measuring outpatients' opinion of quality of hospital consultation departments. *Health Qual Life Outcomes* 2004;2:43.
143. Koenker R. *Quantile regression*: Cambridge university press, 2005.
144. Ding R, McCarthy ML, Desmond JS, *et al*. Characterizing waiting room time, treatment time, and boarding time in the emergency department using quantile regression. *Acad Emerg Med* 2010;17:813–23.
145. Bottai M, Cai B, McKeown RE. Logistic quantile regression for bounded outcomes. *Stat Med* 2010;29:309–17.