RESEARCH ARTICLE

Developing and validating a prediction model for frequent attenders at a Swedish emergency department using an electronic medical record system, a retrospective observational study [version 1; peer review: awaiting peer review]

Lis Abazi1,2, Elin Lindqvist1,2, Gunnar Edman1,3, Magnus Norberg4, Jan Bergman1, Ingmar Zachrisson1, Sune Forsberg1,2

1Norrtälje hospital, Tiohundra AB, Norrtälje, Sweden
2Department of Clinical Science and Education, Karolinska Institutet, Södersjukhuset, Stockholm, Sweden
3Department of Clinical Sciences, Karolinska Institutet, Danderyd Hospital, Danderyd, Sweden
4IT Department, Tiohundra AB, Norrtälje, Sweden

First published: 10 Sep 2021, 10:909
https://doi.org/10.12688/f1000research.53193.1

Abstract

Background:
Frequent attenders (FA) account for a significant number of emergency department (ED) visits but to date there is no prediction model to identify patients at risk of becoming a FA.
The aim of this research was to identify and describe FA using readily available data provided by electronic medical records and create a prediction model to identify future FA

Method:
Adults ≥18 years that visited the ED during 2015 were included. Patients with ≥4 visits were defined as FA, and patients with ≤3 visits were placed in the control group. Numerous variables were analyzed and differences between the groups compared. Logistic regression analysis was used to determine the predictor variables and the model validated using Receiver Operating Characteristic (ROC) on an independent sample.

Results:
6635 patients were included in developing the model: 15.3 (n=1012) were classified as FA and 15.4 (n=1011) as the control group. Variables associated with at risk of becoming a FA were the following: age above 60 years OR 1.52 [CI 1.27 – 1.82], ED arrival by ambulance or helicopter OR 1.31 [CI 1.08 – 1.58], sheltered living OR 3.82 [CI 2.37 – 6.17], previous contact with psychiatric department OR 1.52 [CI 1.23 – 1.89], 10 outpatient care visits or more OR 4.81 [CI 3.81 – 6.08] and 10

Open Peer Review

Reviewer Status  AWAITING PEER REVIEW

Any reports and responses or comments on the article can be found at the end of the article.
outpatient care physician visits or more OR 3.94 [CI 3.25 – 4.78]. The ROC in the validation set had an area under the curve of 0.85 [CI 0.84 – 0.86].

**Conclusion:**
Data from electronic medical record software can be used to create and validate the risk of becoming a FA in the ED. We found that age over 60 years, ED arrival by ambulance or helicopter, sheltered living, previous contact with psychiatric departments, and frequent visits at outpatient care together predict the risk of becoming a FA.

**Keywords**
Emergency department, Frequent attenders
Introduction
Several reports and studies report a steadily increasing demand for services at emergency departments (ED).\(^1\)\(^-\)\(^4\) A small percentage of patients (1-8\%) account for a large percentage of ED visits per year (8.5-28\%).\(^1\)\(^-\)\(^4\)\(^,\)\(^6\) These patients, who are commonly referred to as “frequent attenders” (FA) are defined by the number of their visits and definitions range from 3-12 ED visits per year.\(^1\)^\(^-\)\(^4\)\(^,\)\(^6\) Although this patient group was identified and described decades ago,\(^1\) to date there are no effective interventions that reduce the frequency of their ED visits.\(^6\)

FA consist of a heterogenic group of patients. Their demographics, their diseases, and their social conditions vary significantly. Thus, the group is not easily described.\(^1\) While it is claimed that some FA visit the ED for non-emergency care, several studies reveal FA are a vulnerable group with a verifiably increased risk of mortality.\(^1\) Previous studies have described FA as somewhat older and more likely to seek ED care in non-daytime hours than non-FA.\(^5\),\(^10\),\(^12\) Poverty, homelessness, chronic disease, mental disorders, and drug and alcohol misuse are characteristics more often associated with FA than with non-FA.\(^8\),\(^10\),\(^12\),\(^13\)

A general assumption is that health care services do not sufficiently meet the medical and social needs of FA.\(^1\)\(^,\)\(^4\),\(^6\) Given this situation, plus the increasing pressure on limited ED resources from frequent attendance, a solution is needed that can address the needs of this group and the demands on the health care system.\(^1\),\(^4\),\(^6\),\(^10\),\(^11\),\(^13\),\(^14\) Numerous attempts have been made to solve this problem. These attempts include interventions such as case management, patient education, management care plans, social interventions, and health care centres. However, these attempts have not had a major effect in reducing the number of ED visits by FA.\(^4\),\(^6\),\(^12\) One explanation for this failure may be that these interventions are only initiated after ED patients are identified as FA. Possibly the interventions might be more successful if patients were identified as likely FA earlier in the identification phase.

Furthermore, very little scientific data exist on FA in Swedish emergency departments and the few publication that exist focus on primary health care.\(^1\)\(^5\)\(^,\)\(^10\) The Swedish health care and social system differs significantly from many of those countries studying FA (e.g. USA, United Kingdom etc.) and therefore it is of interest to investigate if the experiences of FA are similar in a Swedish emergency department compared to that reported internationally.

The aim of this study was to identify and characterise FA and to create and validate a prediction model to find FA by using electronic medical record software in an emergency department in Sweden.

Methods
Ethical approval
The Stockholm Regional Ethics Review Board approved this research (D. nr 2017/1695-31/2). There was no requirement to obtain patient consent set by the ethical committee.

Study population
The patients in the study were adults (≥ 18 years) who had visited the ED at Norrtälje Hospital, Sweden, during 2015. The patients’ initial visit during 2015 was used for inclusion and therefore labelled the “inclusion visit” [Figure 1]. This visit, which was used for patient selection only, was not included in the analysis. A study period of 12 months was chosen for each included patient. This individual 12-month observational period was defined by tracking a 12-month period before and after the inclusion visit in such a way that a maximum number of visits were included during a 12-month period [Figure 1]. Patients with four or more ED visits during this 12-month period were identified as FA and patients with three or fewer ED visits were identified as controls. The first visit (at the beginning of the 12-month observational period) was labelled the “first visit”, the second visit was labelled the “second visit”, etc. The purpose of this inclusion process was to observe and compare each individual patient in the same way instead of basing the observational period on calendar time only. The cut of value of four visits or more was chosen prospectively based on previous studies.

Design and data collection
This is a retrospective observational register study. The medical record system – Take Care (TC) at Norrtälje Hospital (Sweden) – was used to identify patients. We identified patients by their personal identity numbers (encrypted before the analysis). The data variables from the first visit were age, gender, number of ED visits in the 12-month observational period, number of outpatient care visits, sheltered living (care facility or not), and enrolment at a health care centre. The data variables from the first visit and (if applicable) from the second visit were ED visit date and time, ED transport, triage priority, main complaint, ED care unit (medicine/surgery/orthopaedic), medication number/type, and ICD-10 diagnosis at discharge (if hospitalised).

Physical findings recorded for the patients were blood pressure (mmHg), heart rate (beats per minute/bpm), temperature (°C), peripheral oxygen saturation (%), height (cm), weight (kg). Laboratory statistics recorded for the...
patients were haemoglobin g/L (Hb), thrombocyte particle concentration $\times 10^9$/L (TPK), leukocyte particle concentration $\times 10^{12}$/L (LPK), creatinine (μmol/L), glucose (mmol/L), C-reactive protein (mg/L) (CRP), sodium (mmol/L), potassium (mmol/L), bilirubin (μmole/L), aspartate aminotransferase (μkat/L) (ASAT), alanine aminotransferase (μkat/L) (ALAT), alkaline phosphatase (μkat/L) (ALP), lactate dehydrogenase (μkat/L) (LD), albumin (g/L). Information on previous contact with psychiatric departments and death within one year following the first visit was also collected. These data were chosen because they were easily available to extract from the electronic medical record software and were thought to be clinically relevant in the light of previous studies.

Norrtälje Municipality and Norrtälje Hospital

The Norrtälje Municipality is approximately seventy kilometres from Stockholm, Sweden. The Municipality has a total area of 6030.42 km² with a population of 61,864 inhabitants (2019). One-third of the Municipality’s inhabitants (20,721 inhabitants, 2018) live in the central town. The surrounding rural area is sparsely inhabited. The Municipality is a somewhat deprived socioeconomic area (compared to many other areas in Sweden). The average yearly income and the average education level are below national averages. Norrtälje Hospital is a small hospital that serves the Municipality and its environs. The Hospital, which has specialist clinics and four wards with 96 beds, is integrated with local primary health care in a way that the both the hospital and the primary health care are within the same organisation with the same management and board. The health care is publicly funded, requires no health insurance and the nearest next emergency department which also is of a higher level, is 60 km away. Attendance to the emergency department does not require referral and can be done at the initiative of the patient. The emergency department treats approximately 20,000 patients yearly.

Statistical methods

The Statistical Package for Social Sciences, SPSS Version 25.0 (IBM SPSS Statistics, RRID:SCR_019096) was used for the statistical analysis. The data was randomly split into two equally large samples where the first was used for analysis (training set) and the second for validation (validation set).

All variables were summarized using standard descriptive statistics such as frequency, mean and standard deviation. Group differences between the categorical variables (e.g., gender, weekday of the first visit, death within one year after the first visit) were analysed using Pearson’s chi-square method. Group differences between the continuous variables (e.g., age) were analysed using Student’s t-test for independent groups. The significance level in all analyses was 5% (two-tailed). If a variable was severely skewed (skewness above 1.5 as for length of hospital stay and for the number of pharmaceutical drugs), a non-parametric Mann-Whitney U test was conducted. The model was created using logistic regression analysis on complete cases on the training set on the relationships between group (FA – “Yes”/”No”) and the identified risk variables. After the model was defined, a predicted risk score for each patient was recorded in both the training set and the validation set. The risk score of the training set and validation set was then used in a Receiver Operating Characteristic (ROC) analysis comparing these two and the sensitivity and specificity determined.

Results

In total 13,193 patients were included in the study and randomly split into two groups of 6635 patients (training set) and 6558 patients (validation set). In the training set 15.3% (n = 1012) were classified as FA and 15.4% (n = 1011) were placed...
in the control group as non-FA. Significant differences between the two groups were found in the following variables: age, sheltered living, number of outpatient care visits, number of outpatient care physician visits, ED arrival by ambulance or helicopter, triage priority, weekday of the first visit, length of hospital stay, previous contact with a psychiatric department, and one-year survival after the first visit [Table 1]. No significant differences were found between gender and time of visit [Table 1].

On average, FA compared to non-FA were older (65.3 years vs 54.5 years), made more outpatient care visits (27.2 visits vs 9.0 visits) and more physician visits (15.5 visits vs 4.9 visits), and had a lower survival rate within one year after the first visit (75.5% survival rate vs 92.2% survival rate). The ED arrival rate by ambulance or helicopter was higher for the FA than for the non-FA (28.3% of arrivals vs 17.5% of arrivals). The FA also had a higher triage priority and made more visits on a weekday (74.8% of visits vs 71.1% of visits). FA also had longer lengths of hospital stays (7.9 days vs 1.1 days) and had a higher rate of previous contact with psychiatric departments (21.8% of contacts vs 10.7% of contacts).

Group comparisons of laboratory and physical findings revealed significant patient group differences in creatinine, potassium, haemoglobin, and glucose [Extended data – Table 1]. Significant differences were found between the groups for systolic blood pressure, saturation and thrombocytes, but these differences, although significant, applied to only a few patients. Approximately 50% of the laboratory findings and approximately 20% of the physical findings were missing in the entire study population. We were therefore unable to use these variables in the analysis. Therefore, they are only briefly presented in the supplementary material [Extended data – Table 1].

Multivariable logistic regression found the following variables posed a risk to patients of becoming a FA [Figure 2]; age above 60 years OR 1.52 [CI 1.27 – 1.82], ED arrival by ambulance or helicopter on the first visit OR 1.31 [CI 1.08 – 1.58], sheltered living OR 3.82 [CI 2.37 – 6.17], previous contact with psychiatric department OR 1.52 [CI 1.23 – 1.89], 10 outpatient care visits or more in the observational period OR 4.81 [CI 3.81 – 6.08], 10 outpatient care physician visits or more in the observational period OR 3.94 [CI 3.25 – 4.78].

Table 1. Baseline data.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control</th>
<th>Frequent attender</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age M/Md (SD)</td>
<td>54.5/56 (20.25)</td>
<td>65.3/70 (20.06)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Women % (n)</td>
<td>49.9 (2,806)</td>
<td>47.0 (476)</td>
<td>0.093</td>
</tr>
<tr>
<td>Survival, 12 mo’s % (n)</td>
<td>92.2 (5,187)</td>
<td>75.5 (764)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sheltered living % (n)</td>
<td>1.5 (85)</td>
<td>3.6 (36)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of outpatient care visits M/Md (SD)</td>
<td>9.0/5 (12.20)</td>
<td>27.2/21 (23.54)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of emergency visits M/Md (SD)</td>
<td>1.5/1 (0.65)</td>
<td>6.1/5 (3.44)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Outpatient care at physician M/Md (SD)</td>
<td>4.9/3 (5.37)</td>
<td>15.5/13 (11.04)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;10 visits % (n)</td>
<td>85.8 (4,825)</td>
<td>33.4 (338)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10- visits % (n)</td>
<td>14.2 (798)</td>
<td>66.6 (674)</td>
<td></td>
</tr>
<tr>
<td>Arrived by ambulance or helicopter % (n)</td>
<td>17.5 (982)</td>
<td>28.3 (286)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Triage priority</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (immediately) % (n)</td>
<td>1.3 (71)</td>
<td>1.3 (13)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2 (≤15 min’s) % (n)</td>
<td>6.8 (381)</td>
<td>8.6 (87)</td>
<td></td>
</tr>
<tr>
<td>3 (≤60 min’s) % (n)</td>
<td>50.7 (2,836)</td>
<td>57.0 (575)</td>
<td></td>
</tr>
<tr>
<td>4 (≤120 min’s) % (n)</td>
<td>36.7 (2,054)</td>
<td>30.4 (307)</td>
<td></td>
</tr>
<tr>
<td>5 (≤240 min’s) % (n)</td>
<td>4.5 (254)</td>
<td>2.7 (27)</td>
<td></td>
</tr>
<tr>
<td>Time at visit M/Md (SD)</td>
<td>13.33/13.28 (5:13)</td>
<td>13.25/13.21 (5:14)</td>
<td>0.473</td>
</tr>
<tr>
<td>Weekday % (n)</td>
<td>71.1 (3,997)</td>
<td>74.8 (757)</td>
<td>0.016</td>
</tr>
<tr>
<td>Length of hospital stay M/Md (SD)</td>
<td>1.10 (3.53)</td>
<td>7.9/3 (12.23)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>n of pharmaceutical drugs M/Md (SD)</td>
<td>4.9/3 (5.49)</td>
<td>6.1/5.0 (4.52)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Psychiatric contact % (n)</td>
<td>10.7 (604)</td>
<td>21.8 (221)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Data from the multivariable logistic regression was used to perform a ROC curve analysis which resulted in an AUC of 0.841 [CI 0.828-0.854] for the training set and an AUC of 0.849 [CI 0.836-0.861] for the validation set [Figure 3] indicating a good model. The ROC curve was then used to find the optimal cut of value and resulted in a sensitivity of 44.5% and a specificity of 91.9% with a positive predictive value of 44.5% and negative predictive value of 91.9%.

Discussion
To our knowledge this is the first time readily available electronic medical record data has been used to identify and characterize FA and to create a FA prediction model. Our study shows that these data, which can be obtained technologically using electronic medical record software, can be used to predict ED patients’ risk of becoming FA.

We found that the variables that most closely correlated with FA status were the following: age above 60 years, ED arrival by ambulance or helicopter, sheltered living, previous contact with psychiatric department, 10 or more outpatient care visits and 10 or more outpatient care visits at physician. This model resulted in an AUC of 0.849 [CI 0.836-0.861] when validated in an independent sample.

The sensitivity and specificity of the model can be adjusted according to the needs the purpose of the model. In our case we aim to use the model for early identification of potential FA and target them with an intervention with low risk of harm. In such case false negative cases are of less importance. However, too many false cases will result in extra work and make the intervention more difficult to accomplish. Furthermore, it is of importance that we correctly identify enough future FA, so an intervention makes a clinical difference. Therefore, we finally settled with a sensitivity of 44.5% and a specificity of 91.9%. Although the sensitivity is low and many FA will be missed, there will be enough correctly identified for an intervention without including too many false positive.
Many studies have described ED FA as a heterogenous group. Studies also associate older age, poorer physical and mental health, and greater social vulnerability with these patients. These studies are performed in health care and social systems that differ from the Swedish. Still, our study confirms these findings. We found that the FA (compared to the control group) were older, were in a worse physical health state, had a higher mortality rate, arrived at the ED more often by ambulance or helicopter, had more deranged blood sample values and physical parameters, had been in contact with the psychiatric department more often, and had been hospitalized for more days.

In confirmation of other research, we found that the FA are recurrent visitors at most health care system levels. We found very high usage of outpatient care among the FA – 27 vs 9 outpatient care visits/year – and the highest OR for frequent attendance. With this descriptive model of FA, we can better identify them at early ED visits, provide better care for them by combining hospital and outpatient care resources, and perhaps reduce the number of their ED visits.

Several interventions have been tested that are designed to address the needs of FA. These interventions have had little success. The most promising intervention is perhaps “case management,” which has sometimes reduced the number of ED visits. We think the results from our study may be used to identify FA at early care stages and easier randomize FA to comparable interventions.

We acknowledge there is a difficulty in generalizing our findings (certainly beyond Sweden's national boundaries) because we use data only from one small Swedish hospital. In any case, it is challenging to draw generalizations in an international dimension regarding FA and to make comparisons among FA studies. Health care/social systems, emergency care, and populations vary greatly – country to country. The burden of disease (both physical and mental) in the ED also differ among countries. Nonetheless we believe that our method and findings can be of interest to the scientific community and add to the general knowledge.

**Limitations**

Our study sample was drawn from ED patients only at Norrtälje Hospital during a relatively brief time period (one year) which makes the results difficult to generalize. Nonetheless we believe that our method and findings can be of interest to the scientific community and add to the general knowledge. We may also have underestimated the number of FA as patients in the control group may have sought emergency health care at another caregiver which we would have missed. The nearest emergency department is 60 km away and requires therefore a certain effort from the patient to visit. We therefore estimate this to be a smaller group and not conflict significantly with the results. Furthermore, the definition of when a patient does his or her first visit at the ED is also very difficult to define. Our analysis was based on this first visit with the ambition of finding FA as early as possible. Even though we cannot say that we have correctly identified the first ED visit for the patients, we still believe our model of defining the first visit is a better way of finding FA at an earlier stage compared to only looking at a calendar year.

**Conclusion**

Our research finds that frequent attendance at the ED may be predicted by the following variables: age above 60 years, ED arrival by ambulance or helicopter, sheltered living, previous contact with psychiatric department, 10 or more outpatient care visits and 10 or more outpatient care visits at physician. These patient variables collectively observed in the first ED visit resulted in a good model with an AUC of 0.849 [CI 0.836-0.861] in the validation sample.

**Data availability statement**

**Underlying data**

We want to do our utmost to contribute to the scientific community by sharing data and being transparent. We are however unable to give our raw underlying data fully and unconditionally to anyone as we have asked and received permission to this data with condition of it being used for our publication purpose only and presented in an aggregated form. Moreover, our raw data contains several detailed variables and although the data is pseudonymized, if combined with other databases we see a risk of individuals being identified.

We therefore need to restrict our data sharing. However, if request is made for our data for scientific purpose in line with our research question, we commit to providing our raw data. This is done by contacting the corresponding author at lis.abazi@ki.se and a reply will be given at the latest within a month.

**Extended data**

Open Science Framework: Developing and validating a prediction model for frequent attenders at a Swedish emergency department using an electronic medical record system, a retrospective observational study. https://doi.org/10.17605/OSF.IO/BK6S4
This project contains the following underlying data:

- Extended data table 1.docx (supplemental table with laboratory results)

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

References

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text


   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

    PubMed Abstract | Publisher Full Text

    PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text | Free Full Text


   PubMed Abstract | Publisher Full Text

   PubMed Abstract | Publisher Full Text | Free Full Text

   PubMed Abstract | Publisher Full Text | Free Full Text

   PubMed Abstract | Free Full Text

   PubMed Abstract | Publisher Full Text | Free Full Text

23. Abazi L: Developing and validating a prediction model for frequent attenders at a Swedish emergency department using an electronic medical record system, a retrospective observational study. 2021, May 18. Publisher Full Text
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com