Ionela MANIU
Research and Telemedicine Center in Neurological Diseases in Children,
Pediatric Clinical Hospital Sibiu,
"Lucian Blaga" University, Sibiu, Romania
Alexander WANDSCHNEIDER
UPU-SMURD Sibiu,
"Lucian Blaga" University student, Sibiu, Romania
Bogdan NEAMTU
Research and Telemedicine Center in Neurological Diseases in Children,
Pediatric Clinical Hospital Sibiu,
"Lucian Blaga" University, Sibiu, Romania

PRACTICAL RECOMMENDATIONS OF DATA PREPROCESSING AND GEOSPATIAL MEASURES FOR OPTIMIZING THE NEUROLOGICAL AND OTHER PEDIATRIC EMERGENCIES MANAGEMENT

Case Study

Keywords
Prehospital care,
Pediatric,
Neurologic emergencies,
Geospatial location,
Emergency services

Abstract

Time management, optimal and timed determination of emergency severity as well as optimizing the use of available human and material resources are crucial areas of emergency services. A starting point for achieving these optimizations can be considered the analysis and preprocess of real data from the emergency services. The benefits of performing this method consist in exposing more useful structures to data modelling algorithms which consequently will reduce overfitting and improves accuracy. This paper aims to offer practical recommendations for data preprocessing measures including feature selection and discretization of numeric attributes regarding age, duration of the case, season, period, week period (workday, weekend) and geospatial location of neurological and other pediatric emergencies. An analytical, retrospective study was conducted on a sample consisting of 933 pediatric cases, from UPU-SMURD Sibiu, 01.01.2014 – 27.02.2017 period.

INTRODUCTION

Very few studies have addressed the issue of prehospital care of pediatric emergencies, especially in Romania.

Emergency medical services (EMS), as well as any other community service, require constant planning, re-evaluation, and optimization efforts. For this purpose, preprocessing and analysis of real data is a must.

This article describes retrospectively the data that define different types of problems encountered by the Mobile Emergency Service for Resuscitation and Extrication - UPU-SMURD - from Sibiu service area (UPUS), over a consecutive 3-years period (01.01.2014 – 27.02.2017, 112 calls).

MATERIALS AND METHODS

During the study period, 933 pre-hospital computerized records from UPUS were reviewed, with an emphasis on pediatric patients' age, gender, environment, main complaints, date and time of the announcement, duration of intervention, vital functions and forms of management/procedures.

In the process of data preprocessing, as a first step, we identified the cases of inconsistency or data entry errors and then we created new variables from the existing data.

After a detailed data-analysis, we considered the following exclusion criteria: non-pediatric cases - persons older than 18 years, unknown age, more people per case (which does not allow accurate identification of vital functions and maneuvers), cases where different time recordings were not known (time of arrival, time of departure to hospital, etc.).These cases were excluded from analysis.

Age is an extremely important element in pediatric pathology and therefore, it has been analysed in detail. Creating a scatterplot plot, an overview of age was obtained, allowing the visual display of age groups by aggregating the cases in the upper and lower part of the graph. According to literature, two new age groups were considered. AG1 divides cases into newborn/infant (0-1 years), ante-pre-schooler (2-3 years), pre-schooler (4-7 years), schooler (8-15 years) and adolescent (16-18 years). These age groups were also identified in puericulture course, in the General Medicine Specialisation, Year 5.

Starting from the date of case recording, new variables were created: "Season", "Weekdays", "WDG". In the "Season" variable there are obvious entities: Spring, Summer, Winter, Autumn. "Weekdays" are days of the week, and "WDG" is a variable that divides weekdays on working days and weekend.

Processing the time of the announcement we defined two new variables. The medical literature uses specific values: 0-8 night shift, 8-16 - morning shift and 16-24 afternoon shift. This division was stored in the "HI1" variable. Regarding the way the nurses and doctors' tours were made at the UPUS, there are only two shifts: day time shift (8-20 o'clock) and night time shift (20-8 o'clock). This division is found in the variable "HI2".

The following time moments are known related the total time of the intervention: 1. the moment when the crew mission was issued (at UPUS difference between receiving notice and departure is minimal with one minute during the day and a maximum of two minutes per night), 2. arrival at the place of intervention, 3. returning to the hospital, 4. arrival at the hospital, 5. receiveing of another case.

We calculated the differences of these times moments. Four time intervals have emerged: 1. A_SC representing the length of time from call receiving until the ambulance arrival on the scene(response time), 2 SC_PS as the time interval or the duration of the case (time spent on scene), 3. PS_SS was the time for transport to the hospital, 4. SS_E meant the duration of the patient's delivery, recovery of the used materials and return of the ambulance to the state of intervention. Two variables were further assigned to calculate the total time of a mission: 1. Total1, as the difference between the release date and time and the announce of date and time respectively, 2. Total2 , as the sum of all of the previously calculated time intervals. Then, both variables values were compared to detect any calculation and data entry errors.

Eventually, data preprocessing steps were performed using Microsoft Excel (Microsoft corp.) while descriptive data analysis was undertaken using SPSS Statistics (IBM). We used Google Maps API through the QGIS platform with the MMQGIS and OpenOSM extension to compute geospatial representation.

Continuous variables distribution was reported as a mean ± standard deviation, 95% confidence intervals (95% CI) and 25 percentile (P25), 50 percentile (P50, median), 75 percentile (P75). We even analysed the 95% CI in a case of response and transport time based on 100 bootstrap samples. On the other hand we described categorical variables distribution as a number of cases and percentages. To study the normal distribution for continuous variables we applied Kolmogorov-Smirnov test. To conclude, we employed student t-test, Kruskal-Wallis, Mann-Whitney tests for comparison, according to distribution type [Mocan 2005, Maniu, 2014].

RESULTS

Applying exclusion criteria only 926 out of the 959 cases remained (96.5%), spreaded over the last three years in the following pattern: 241 (26%) in 2014, 362 (39.1%) in 2015, 283 (30.6%) in 2016 and 40 (4.3%) in the first month of 2017. More than 50% of cases were medical cases, 23% trauma cases, 9% road accidents, 8% neurological cases, about 5% medical transport and intoxications and 5% other cases.

To achieve an image focused on medical cases, we considered three groups for further analysis: neurological cases (N = 77), medical cases (N = 464) and intoxications (N = 23). These patients totalized 564 cases (60.9%), 52% being females and 70% coming from urban areas.

The mean age value was M = 7.55 (SD = 6.37), with aproximately one year increase in the years 2015 (M=7.73, SD=6.40, P25=2, P50=6, P75=15) and 2016 (M=7.86, SD=6.31, P25=2, P50=6, P75=14) compared to 2014 (M = 6.68, SD = 6.40, P25 = 1, P50 = 4, P75 = 14). Statistically significant differences were noticed only for the pair 2015 vs 2016 with p=0.055(Mann Whitney test), while in the other two instances we could observe only a trend: 2014 vs. 2015 (Kruskal-Wallis test, p=0.112), 2014 vs 2016 (Mann Whitney test, p=0.077).

Unsurprisingly, intoxications have been encountered in higher average age cases than medical or neurological ones (figure no.1). Medical cases in infants were found in 39% of patients in 2014, 22% and 23% of patients in 2015 and 2016 respectively. Neurological cases in infants were referred in 40% of patients in 2015, 28% and 26% in 2014 and 2016 respectively. Ante-pre-school and pre-school medical cases were approximately constant during the three years, around 15% for each category (30% cumulative). Medical cases in school children were 22% in 2014, 30% and 26% in 2015 and 2016, while neurological cases were 47% in 2014, 15% and 26% in 2015 and 2016 respectively. In adolescents, there are medical and neurological cases below 15% for each category (figure no.2).

The number of neurological cases increased in weakend from 14.3% in 2014 to 30% in 2015 and 40% in 2016. The number of medical cases was constant over the last three years (about 30% on the weekend). Intoxications number ranged from 20% in 2014 and 2016 to 54% in 2015. In 2014 and 2015 there were more pediatric emergencies in spring and summer vs. autumn and winter, with neurological causes 53% in 2014 and 75% in 2015 while medical causes were slightly over 60%. In 2016 there were more cases in winter and spring (neurological 66%, medical about 60%).

Over 4.5% of the cases were between 10:00 AM and 23:00 PM, with a bimodal distribution having peaks at 12:00 AM and 19:00 PM. More than 60% of cases were in day time shift with 63.6% neurological, 60.3% medical and 56.5 % intoxication. From neurological cases, 49.4% were in the 8-16 hour range and 47.8% of the intoxications in 17-24 hour range.

The mean response time was M=8 minutes (SD=7.39, 95%CI [7.29; 8.65]). Mean transport time was M=9.48 minutes (SD=15.30, 95%CI [8.44; 10.75]) while mean on-scene recorded time was M=15.17 minutes (SD=10.65, 95%CI [14.28; 16.31]).

Geospatial representation of pediatric emergencies was performed using Google Maps API for address geocoding through the QGIS platform with the OpenLayer extension for the OpenStreetMap base map layer. The basic idea of Heatmap [Khan 2003, May 1950] representation is a matrix with lines and columns (representing latitude and longitude geographic coordinates) overlapped on a geographic map. All values in this matrix are initially set to zero and when an event is added - in this example a case - the value of the cell corresponding to the line and column would be incremented. At the end of the data set, a matrix would be obtained with values between zero and maximum of cases. There would be a number of "lines x columns" regions. A colour can be assigned to each region, or ideally a gradient of colours such as white, grayscale and black, to express these values. The white colour will show zero map events on a region while black the highest number of events met. All shades of grey show the intermediate case values[Wilkinson 2009].

Using the World Geodesic System 84 standard, EPSG: 4326, we determined the geographical location of pediatric emergencies, the pediatric emergency department and schools in the Sibiu area. Then, we representing heatmaps for distribution of cases for this three years period and for the three groups.

The pathologies of this study were colour coded into blue – intoxications, green – medical and purple – neurological, each one on a distinct layer. We should notice that by changing the layer blending mode to multiply, layers are not only seen on top of the base map, but also there's a colour addition. The occurrence of all entities in an area gives a darker colour (figure no.3).

We discovered that intoxications cases occurred in the South-West of Sibiu. Medical cases were distributed in a cross-shape with the highest occurrence centrally, while neurological ones were found to be commonly in the Southern area. Furthermore, in correlation with case occurrence, we found five schools in the South, two in the West and five schools in the Central area.

Case distribution examination by year brought up the following results: 2014 had the majority of cases in the South-Western area, while 2015 and 2016 had a distribution with a cross-shaped appearance. Overlapping layers produced a black colour in the map's centre, West, North and South (figure no. 4).

CONCLUSION AND DISCUSSION

Neurological pediatric emergencies represent 8% of the cases in this three-year study, with a percentage between 26% and 40% in infants and between 26% and 47% in schoolers. About 50% were in a time range between 8 AM and 16 PM with a mean total time spent on scene of M=55.83 minutes (SD=47.44, 95%CI [45.06; 66.60]).

There is an imperative necessity of nationally and internationally representative data and data analysis to reflect the pediatric emergency profile and needs and to guide the efforts for training (pediatric skills), equipping, planning, optimizing and improving the activity of emergency medical services providers.

Each of the variable discretization variants offers specific information, which supports the idea of using several variants of discretization of a variable. The disadvantage resides in the fact it requires more time for results computing and analysis.

Chief complained procedures are divided differently (by groups and subgroups) in various studies [Tsai 1987, Richard 2006, Shah 2008, Miller 2009, Li 2013, Houtekie 2015, Diggs 2016]. Homogenization in this regard would allow a better comparison between studies.

Geospatial data analysis allows highlighting the spatial patterns and the density of events / casuistic and also the correlation between densities and frequent places, being useful in designing optimized prevention measures. Particularly, cases occurred commonly near schools, and in the downtown area of Sibiu, while the Northeastern part of the city was spared. A surprise was the high frequency in the area around the city centre in which the pediatric emergency department location lies.

Furthermore, this study highlights the heterogeneity of case distribution in Sibiu. By analysing the conditions of cases, it is possible to find place-based social determinants of health like shown in other studies [DA Dworkis 2017], but therefore, additional data like neighbourhoods, income and other demographic information is needed. Growing this „digital infrastructure" in Romania and Sibiu is an important future challenge, which won't just aid the emergency department.

Acknowledgement

This work has been conducted in the Pediatric Clinic Hospital Sibiu, within Research and Telemedicine Center in Neurological Diseases in Children - CEFORATEN project (ID 928 SMIS-CSNR 13605) financed by ANCSI with the grant number 432 / 21. 12. 2012 thru the Sectoral Operational Programme „Increase of Economic Competitiveness”.

REFERENCES

[1] L.A. Diggs, M. Sheth-Chandra, G. DeLeo (2016) Epidemiology of pediatric prehospital basic life support care in United States, Prehospital Emergency Care Journal, 20(2):230-8. doi: 10.3109/10903127.2015.1076099

[2] Dworkis DA, Taylor LA, Peak DA, Bearnot B (2017) Geospatial analysis of emergency department visits for targeting community-based responses to the opioid epidemic. PLoS ONE 12(3): e0175115. https://doi.org/10.1371/journal.pone.0175115

[3] L. Houtekie, P. Meert, F. Thys, V. Gui-Viterbo, C. Stephan (2015) Prehospital paediatric emergencies in Belgium: an epidemiologic study, European Journal of Emergency Medicine, 10.1097/MEJ.0000000000000116

[4] O.A. Khan (2003) Geographic Information Systems and Health Applications. Hershey, PA: Idea Group Pub., 2003. pg 242-259.

[5] S.-T. Li, N.-C. Chiu, W.-C. Kung, J.-C. Chen, (2013) Factors Affecting Length of Stay in the Pediatric Emergency Department, *Pediatrics & Neonatology*, 54, 179-187

[6] I. Maniu (2014) Data analysis techniques: statistics, Ed. Univ. „Lucian Blaga” Sibiu, ISBN 978 – 606 – 12 – 0891 – 3

[7] J.M. May (1950) Medical Geography: Its Methods and Objectives, Geographical Review 40.1 (1950): 9. Web

[8] K.M. Miller, D. Dowd, M.C. Gratton, J. Cai, S.D. Simon (2009) Pediatric Out-of-hospital Emergency Medical Services Utilization in Kansas City, Missouri, Academic Emergency Medicine Journal, 16(6), 526-531

[9] I. Mocan (2005) SPSS Introduction to data analysis, Ed. Univ. „Lucian Blaga” Sibiu, ISBN 973 – 739 – 189 – 6, pg. 10-96

[10] J. Richard, M.Osmond, L. Nesbitt, Stiell (2006) Management and outcomes of pediatric patients transported by emergency medical services in a Canadian prehospital system, Canadian Journal of Emergency Medicine, doi: 10.1017/S14818003500013312

[11] M.N. Shah, JT Cushman, CO Davis, JJ Bazarian, P Auinger, B Friedman (2008) The epidemiology of emergency medical services use by children: an analysis of the National Hospital Ambulatory Medical Care Survey, Prehospital Emergency Care Journal, 12(3):269-76

[12] A. Tsai, G. Kallsen (1987) Epidemiology of pediatric prehospital care, Annals of Emeregency Medicine,16(3):284-92

[13] Leland Wilkinson, Michael Friendly (2009) The History of the Cluster Heat Map, The American Statistician, 63:2, 179-184, DOI: 10.1198/tas.2009.0033.

Appendix

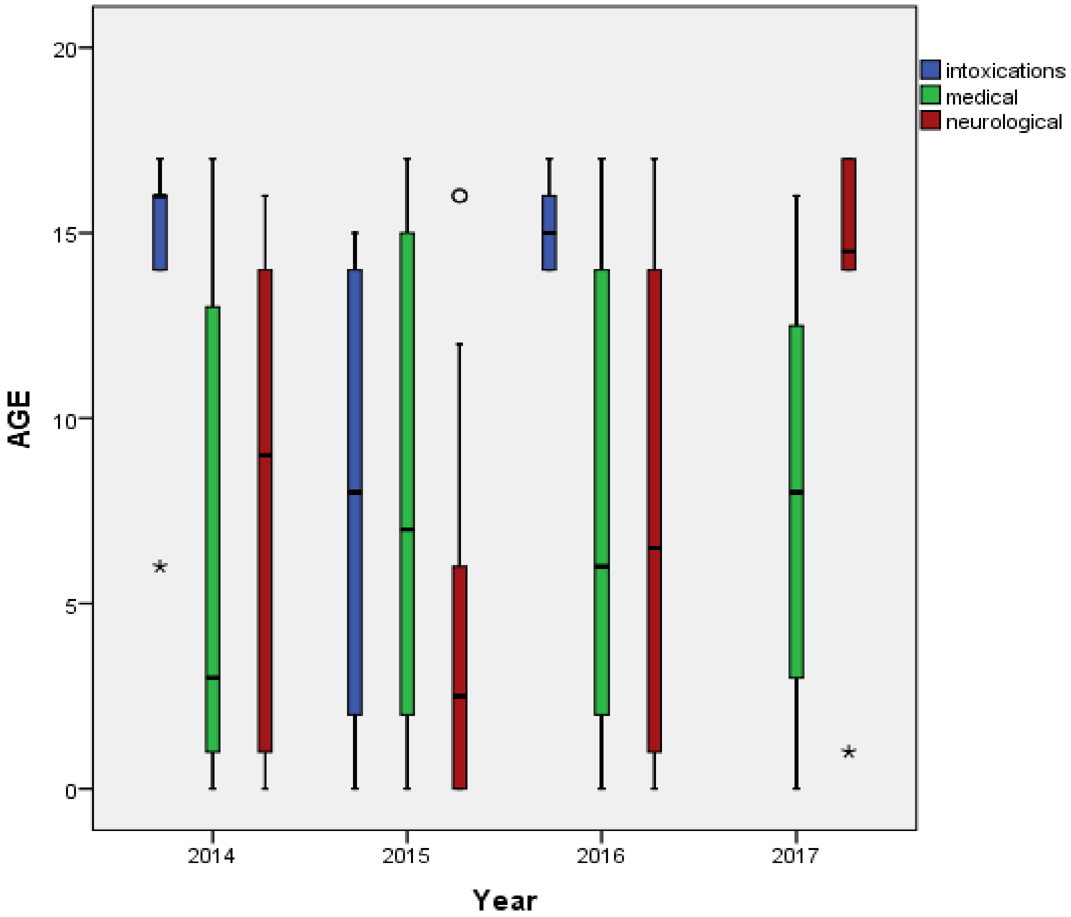

Figure No.1 Age distribution for the three pediatric emergency groups in case of each studied years

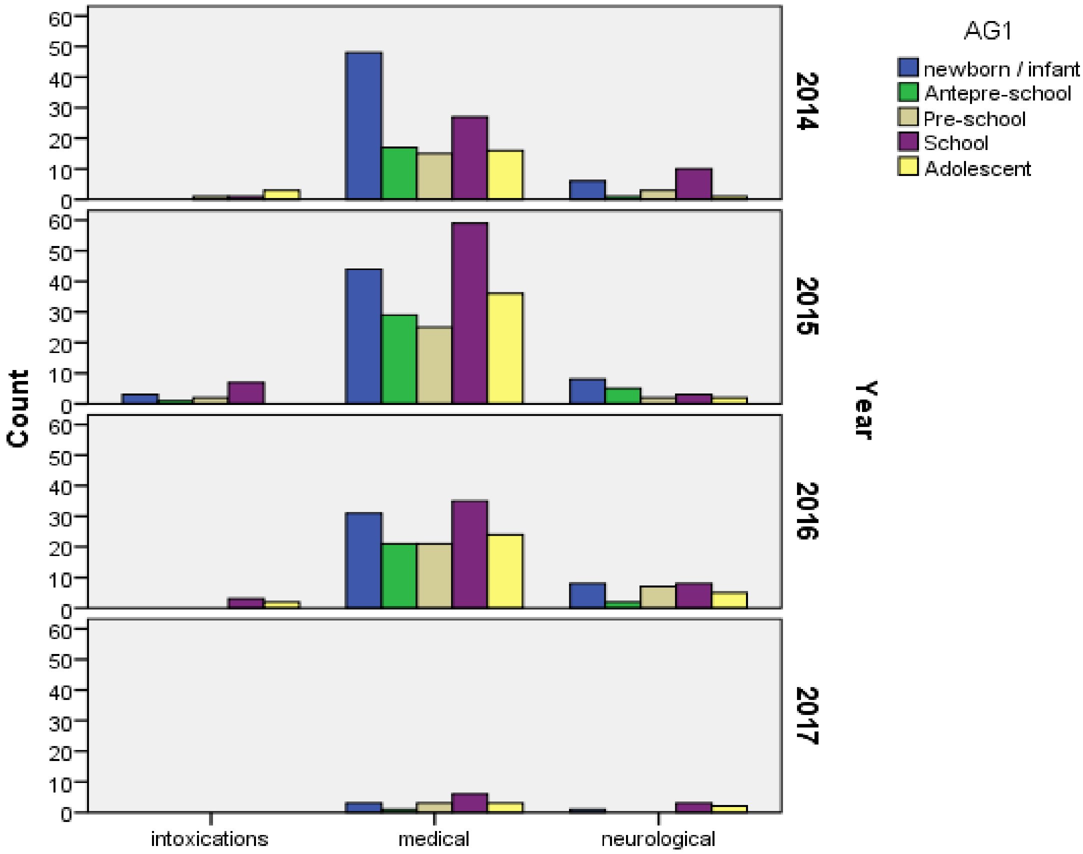

Figure No.2 Distribution by age groups for the three pediatric emergency groups in case of each studied year

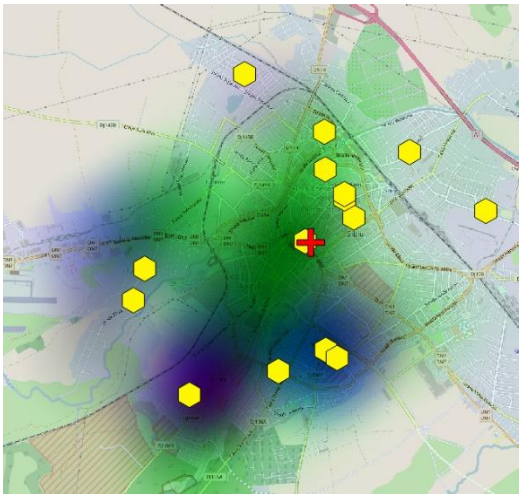

Figure No.3 Heatmap for the three pathologies. Blue – intoxications, Green – medical and Purple -neurological cases. Red Cross – Pediatric Emergency Department, Yellow Hexagon – Schools

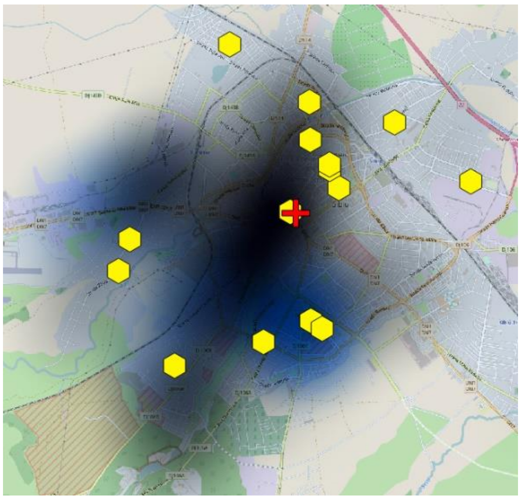

Figure No.4 Heatmap for the three years. Blue – 2014, Green – 2015 and 2016 - purple cases. Red Cross – Pediatric Emergency Department, Yellow Hexagon – Schools