

Public Health surveillance from emergency call center data: visualization dashboard and NLP of call reports

Alexandre Naprous^{1,*}, Marta Avalos-Fernandez^{1,2,*}, Catherine Pradeau^{1,3}, Emmanuel Lagarde¹, Cédric Gil-Jardiné^{1,4}

¹University of Bordeaux, Bordeaux Population Health Research Center, UMR U1219, INSERM, F-33000, Bordeaux, France

²SISTM team Inria BSO, F-33405, Talence, France

³University Hospital of Bordeaux, CRRA 15 - SAMU 33 - SMUR, F-33000, Bordeaux, France

⁴University Hospital of Bordeaux, Pole of Emergency Medicine, F-33000, Bordeaux, France

*Equal contributors

{first name.last name}@u-bordeaux.fr

Abstract

By focusing on symptoms and not diagnoses, the so-called syndromic surveillance system gains in immediacy what it loses in specificity with respect to other more traditional options for public health surveillance. Reports of calls to emergency medical communication centers (EMCC) supplemented by the data collected by the rescue workers who arrived on the scene constitute a cost-effective and rich source of information. Unfortunately, EMCC data are infrequently used and their utility has not been demonstrated.

The aim of this study was to explore the usefulness for public health surveillance of EMCC data when analyzed using text mining and visualization tools. Transformer-based deep learning architectures were used to classify call reports according to the reason for the call. We also extracted indicators that could serve as proxy measures using a keyword-search algorithm. We then developed a dashboard visualization tool to enable dynamic and spatial exploratory analyses. Finally, we explored the potential of this tool for two applications. While the tool proved unable to detect Covid-19 outbreaks, it appeared to be promising for a better understanding of territorial inequalities in healthcare access.

Introduction

Syndromic surveillance is an approach in which automatic data recording procedures allow the provision of data based on signs, symptoms or preliminary diagnoses for near-real-time tracking unexpected health events, monitoring expected trends, and conducting health impact assessment of infectious or environmental hazards. Syndromic surveillance can supplement traditional Public Health surveillance by analyzing less specific data from clinical and non-clinical sources. Non-clinical sources include social Medias, web searches, telephone helplines, mobile phone applications, pharmacy sales, etc. Clinical sources include data collected from calls to emergency medical communication centers (EMCC) and hospital emergency room (ER) data. While ER data are widely used for syndromic surveillance, data collected from the other emergency medical services are less frequent.

Yet, the data recorded during calls to EMCC possibly enhanced with data transmitted by the ambulance person-

nel when the dispatch of an ambulance is necessary, constitute a cost-effective and rich source of information including temporal, spatial and standardized/unstandardized clinical data. These data have the potential of providing help in early detection of health events in the general population or subpopulations (Duijster et al. 2020). This could cover information about less severe health conditions, not requiring emergency transfers and uncovered by ER data but that could be useful for Public Health surveillance. However, because the primary goal of the EMCC is not diagnostic accuracy but resource allocation, the specificity of the data may be low. The few studies that have used EMCC data have mainly focused on the surveillance of infectious diseases (Duijster et al. 2020), though there are also examples of studies on alcohol and drug intoxication (Holzer et al. 2012; Manca et al. 2021), mental health (Duncan et al. 2019), stroke (Seo et al. 2014) and trends (Man Lo et al. 2012; Gil-Jardiné et al. 2021b). Its utility in Public Health surveillance has not been fully demonstrated.

In France, EMCC, the so-called SAMU-Center 15 (emergency medical call center - phone number 15), are free public call centers, which operate 24/7 and respond to all health emergency calls for a defined area. They are the French equivalent of 911 in the USA and 112 in Europe. In France, dialing 112 will re-direct to the SAMU-Center 15 or the fire department, depending on the department. A call is first received by a medical assistant, and then an emergency physician or a general practitioner (depending on the severity of the case) decides on the appropriate response, from medical advice to the dispatch of an ambulance or a mobile intensive care unit (MICU).

With ~300 thousand calls per year, the SAMU of the Gironde department (1.6 million inhabitants) is the second largest EMCC of France (Gil-Jardiné et al. 2021a). Its information system allows exhaustive archiving of all the calls and their treatment since 2005. This database is very large in terms of the number of records archived, but also in terms of the scope of the information it contains. Indeed, for each call, the origin of the call, the profile of the caller, the place of intervention, the reason for the call, the circumstances, the actions triggered by the call and the fate of the patient are reported. These data could be either standardized or in the form of free texts. Finally, there is information allowing both geolocation and clinical assessment (vital signs of the

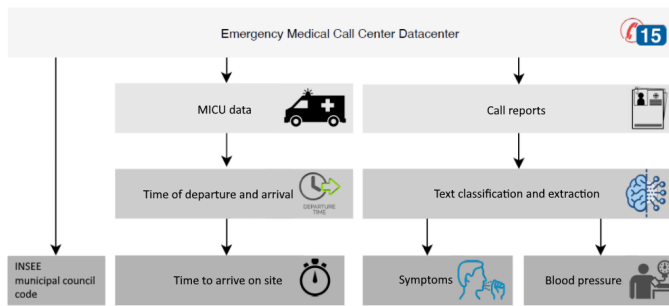


Figure 1: Data-flow diagram. Only authorized hospital staff can have access to the call reports which contain PHI information. Data used for the dashboard (last line of the data-flow) were de-identified.

patient) measured by the callers and the ambulance personnel.

The aim of this study was to explore the usefulness of EMCC data in Public Health surveillance with relevant tools. To do so, we used data from the Gironde department. From the call reports, we applied deep learning methods for knowledge extraction. Transformer-based architectures were used to classify call reports according to the reason for the call or the pathology (influenza-like symptoms, chest pain, stroke, violence, etc.). We also extracted indicators that could serve as proxy measures for access to healthcare from text reports using a keyword-search algorithm. We then developed a dashboard visualization tool to perform a dynamic and spatial exploratory analysis of these data. Finally, we explored the reliability of the tool through two applications: detection of Covid-19 outbreaks and highlighting of territorial disparities in healthcare access.

Methods

Data sources

For all cases handled by the Gironde’s EMCC, a clinical report is created in the form of a computerized free-text note in which the circumstances of the event at the origin of the call and the clinical observations are detailed. It is progressively updated by a medical assistant and a physician through the various telephone interactions with the patient/family/witnesses and, when applicable, with the medical or paramedical staff. Date and time are registered for every input. After removal of duplicate or empty clinical reports, ~4 million remained over the period 2005 to 2020. A portion of clinical reports are annotated during the calls using a standardized diagnosis code (~60%). We considered 13 broad categories for the main reasons of the call: chest pain, gastroenteritis and abdominal pain, flu-like symptoms and breathing difficulties, focal neurological deficit and stroke, road traffic crash, violence, suicide and self-harm, injury other than violence, self-harm and road crash, pregnancy and delivery problems, malaise with loss of consciousness, stress and anxiety, and “other”. For the validation purposes of the current study, a random sample of size ~40 thousand from the unlabeled clinical reports of

2019 was manually annotated by emergency nurses (single annotation).

During a call, the address of the patient is registered systematically. Authorized hospital staff linked addresses to the official geographic code, a numeric indexing code used by the French National Institute for Statistics and Economic Studies (INSEE) to identify municipal councils.

If a MICU dispatch is required, times of departure and arrival on site are sent and recorded. On the other hand, vital parameters such as blood pressure are routinely measured at the arrival on site and sent to be included in the clinical report.

Only authorized hospital staff can have access to the call reports which contain PHI information (e.g., address). For use in this research, reports were automatically de-identified. We used a grep-based text-search procedure that was applicable because of the standardized format of personal information inserted in the reports. Figure (1) shows data-flow diagram.

Text classification and text extraction

A Transformer-based model (Vaswani et al. 2017) was trained in two phases: a self-supervised phase using unlabeled examples followed by a supervised one using annotated examples. Text classification was initially performed on the 2016-2020 dataset. We used ~900 thousand call reports corresponding to the period 2016-2018 in the self-supervised phase. Of these, we used the ~700 thousand that were manually annotated live during calls with a standardized diagnostic code. The 117-million parameters version of the GPT-2 model (Wolf et al. 2019; Xu et al. 2020) was trained on a workstation with one Nvidia GeForce RTX Titan Graphic Processing Unit with 24GB of video random access memory. The sample of ~40 thousand manually annotated reports of the year 2019 allowed to validate the trained and validated models on off-line annotated data.

The models were then applied to the ~250 thousand labeled and unlabeled reports of the year 2020, providing a classification of reasons for calls. The confidence of the method was assessed using a bootstrap procedure with 10 thousand random partitions of the validation sample. For each partition and each category, the cut point was selected using the first half of the partition sample such that precision and recall were equal. The difference between the manually determined prevalence of the given category and the predicted prevalence was then computed in the second half of the sample. Median bias and 95% confidence interval were derived. The applied text classification methodology is detailed in a previous work (Gil-Jardiné et al. 2021b).

Data concerning time of departure and arrival of MICU were directly extracted from standardized variables in the EMCC database for the period 2010-2020.

Blood pressure measures are available in the free text of clinical notes. We used a regular expression search method to extract these data for the period 2010-2020. Regular expression search was also applied to symptoms. The negative forms were excluded. Misspelling and syntax variations were taken into account using the same method. The procedure applied to symptoms was previously developed and

validated for a list of common symptoms on the 2005-2020 database (Gil-Jardiné et al. 2021a).

Dashboard development

We computed the mean number of calls for a given category (among the 13 defined above) at a given date by municipal council as the moving average number of calls over a period of user-defined length centered at the given date. We also considered the mean number of calls for a given category and a given municipal council among the mean number of calls for the given municipal council (being the mean number of calls a proxy of the dynamic population density). In order to compare two different categories, we also computed the ratio between mean number and proportion of calls. Finally, for these measures, we considered spatial correlation indices (global and local Moran's I statistics, with spatial weights matrix based on crow flies distances as well as driving time through the Open Source Routing Machine Application Programming Interface -OSRM API-).

To visualize these spatial indices and their evolution over time we developed a dashboard. The Shiny R package was used to build an interactive dashboard (Chang et al. 2021). Python was used for calls to APIs and processing of results as well as for the creation of analysis notebooks.

We used a choropleth map (a map fragmented into different polygons whose color changes according to the geographical data but not the size). The dashboard includes interactive features to select the indices (counts, proportions, etc.), the period of study, the date within this period, the length of the period, the call category. A reactive function allowed manipulating on-line data fast enough to keep the dynamic aspect of the dashboard when changing options. To make colors of the municipal councils change according to the quantitative indices, we used the *r-Leaflet* library (Cheng, Karambelkar, and Xie 2021), and downloaded the GeoJSON data of the municipal councils of Gironde on the *france-geojson* site. This site generates GeoJSON files of the French territory according to the scales and areas requested, with a JSON-like format describing various geometrical objects. We used a linear color scale to be able to compare municipal councils during epidemic (high number of calls everywhere) and non-epidemic periods. Interactive features also include a video export button that allows saving the changes in a video. The main purpose was to visualize the evolution of indices over time and to easily share it.

Applications

We explored the reliability of the tool through two applications: detection of Covid-19 outbreaks and highlighting of territorial disparities in healthcare access.

Detection of Covid-19 outbreaks

Since our aim was to create consistent surveillance tools for new influenza-like outbreaks we explored if Covid-specific features could have been observed in near-real time from EMCC data. We considered calls related to flu-like symptoms and difficulty breathing at the first Covid-19 wave in Europe (late February 2020) and between 2016-2019 (a reference period used for comparison and for model-building). We chose not to analyze the presence of unconventional but

Covid-specific symptoms, such as loss of taste and smell, because these symptoms took some time to emerge, so they may not have been reported early in the pandemic.

Highlighting territorial inequalities in access to healthcare

In France, territorial inequality in access to health care (infrastructure, medical personnel, but also concentration of unfavorable socio-economic conditions in certain areas) is a complex problem for which it is difficult to provide objective evidence of its consequences.

We explored the spatial distribution of two indicators that could (directly or indirectly) measure territorial inequalities in access to healthcare. First, we computed the time between the departure and arrival on site of MICU. Second, we calculated whether a patient under care had high blood pressure (HBP), defined as a systolic blood pressure greater than or equal to 140 mmHg. The large portion of burden of diseases is attributed to HBP in the developing countries. HBP usually has no warning signs or symptoms, and many people are unaware that they have it until it is incidentally discovered by a healthcare professional. Awareness, treatment and control of HBP has been associated with socio-economic inequalities (Palafox et al. 2016; Anstey, Christian, and Shimbo 2019). HBP served as a proxy measure for access to healthcare.

Results

Text classification and text extraction

Deep-learning text classification achieved a Max-F1 score ranging from 0.48 to 0.80, and an area under ROC curve ranging from 0.79 to 0.99 on the validation dataset. The 3 classes achieving the best Max-F1 scores among the 13 classes were "pregnancy and delivery problems", "chest pain", and "road traffic crash" (scores of 0.80). The 3 classes worst performing were "malaise with loss of consciousness", "stress and anxiety", and "violence" (scores of 0.48, 0.49, and 0.64, respectively). Text classification results are detailed in a previous work (Gil-Jardiné et al. 2021b).

Keyword search applied to symptoms achieved a Cohen's kappa ranging from 59 to 75 for the following symptoms: agueusia, anosmia, cough, fever, muscle soreness, dyspnea. Text extraction results are detailed in a previous work (Gil-Jardiné et al. 2021a).

Dashboard

The current version of the dashboard is a pilot edition developed to explore the usefulness of EMCC data in Public Health surveillance. The Gironde department has 535 municipal councils that are represented in the interactive map dashboard. The number of calls over time is also represented by a time series graph. The dashboard contains 4 main sections detailed in Figure (2): (i) "Filter," (ii) "Variables," (iii) "Indices," and (iv) "Video." The Filter section allows the user to select inputs such as the period of study, a date within this period, the length of the period, the minimum number of calls of municipal councils to be represented, and the age range. In the Variable section, the user selects symptoms or call reasons to appear on graphs. In the Indices section, the

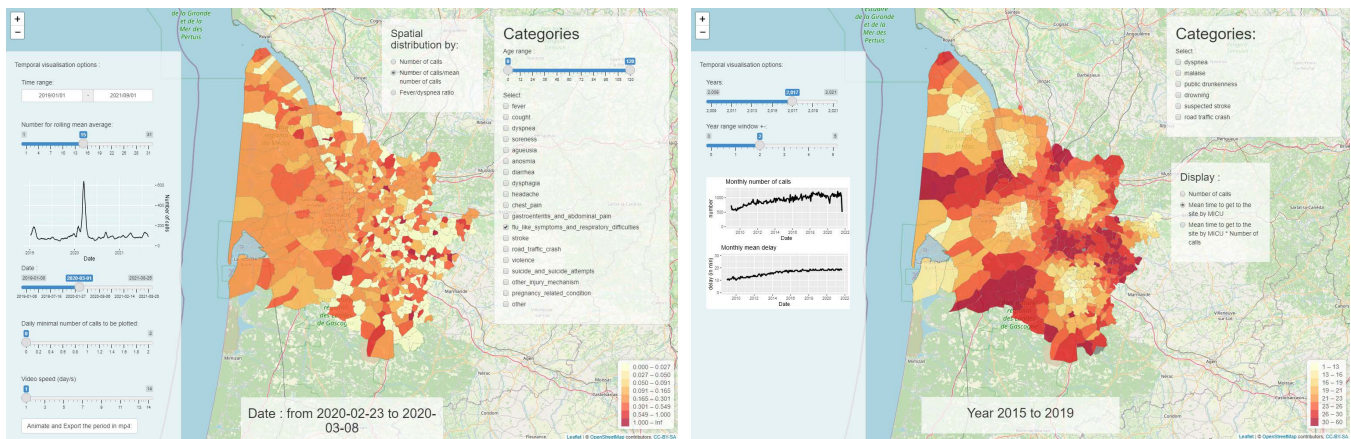


Figure 2: The dashboard to visualize spatial distribution and the number of calls over time with interactive features. Interactive features allows to select, for example, the indices to be spatially represented, the period of study, the date within this period, the length of the period, the minimum number of calls per municipal council, video export options, the call category and age range. Data correspond to flu-like symptoms and difficulty breathing calls during the first Covid-19 wave in Europe (left) and time (mean values over time per municipal council) to get to the site by mobile intensive care units (right).

indices to be spatially represented (counts, proportions, ratios, means, medians, etc) are selected. Finally, the Video page also lets the user select the video export options. Overall, the user therefore has a role in filtering the data and in deciding which information is represented.

Applications

Figure (2), left, shows the spatial distribution of the number of calls corresponding to flu-like symptoms and difficulty in breathing across all calls during the beginning of the first Covid-19 wave in Europe. The total number of calls is used as a proxy of the dynamic population density. Despite the fact that the Gironde department was little affected (in terms of hospitalized patients, since Covid-19 tests were not yet available) during the first wave, the proportion of flu-like symptoms and difficulty breathing is high almost everywhere. Regardless of the indicator used and the comparison made with the 2016-2019 years, no particular feature was detected.

Figure (2), right, shows the spatial distribution of time (mean values over the period 2010-2020 per municipal council) to get to the site by MICU. This map suggests that the closer to the hospital centers, which are located in urban areas, the sooner the MICU arrive.

Figure (3) shows the proportion of HBP by municipal council for the period 2010-2020 (left) and the median yearly income in EUR by municipal council for the year 2015 (right). There is not a perfect match between both maps, we observe a tendency: the northern and eastern periphery generally show high proportions of HBP and low median incomes.

Conclusion

Disease surveillance in public health has received increased attention over the past two years. Early identification of communicable disease outbreaks can be beneficial

in addressing many aspects of public health. Public health surveillance based on EMCC call reports appears to be feasible. However, the sensitivity and specificity of the EMCC signal appear to be variable and the quality of the results depends on the health event of interest. The EMCC data used in this work has the potential to be beneficial in these efforts if used properly.

In the Covid-19 application, we were faced with detecting weak signals in intense noise. The exclusion of novel disease-specific symptoms, such as loss of taste and smell, when reviewing the initial Covid data is questionable: these disease-specific symptoms are likely to be valuable once established. This work was conducted as part of a Covid-19 epidemiologic surveillance research project. However, the same tools could be applied to analyze seasonal patterns of the same symptoms in relation to influenza prevalence. By leveraging the extensive data available over time, these results could be more representative in predicting baseline seasonal resource needs, and provide additional differentiation, in an observed versus predicted manner, for the incremental call volume attributable to the pandemic era compared with the pre-pandemic era.

Indicators that can directly measure or serve as proxy measures of access to care recorded over a long period of time have suggested territorial inequalities in access to care. Patient outcomes can obviously be impacted by the time it takes for help to arrive. In addition to territorial inequalities in access to care, other important elements (currency and volume of skills, cost-effectiveness, limited and expensive MICU resources) may help explain the observed differences. On the other hand, people living in low-income municipal councils may have a higher risk of being unaware that they have BPH, which may be related to poorer access to health-care professionals. The use of an isolated blood pressure during a period of heightened stress (experience of an event sufficient to contact the EMCC) should be considered a screen-

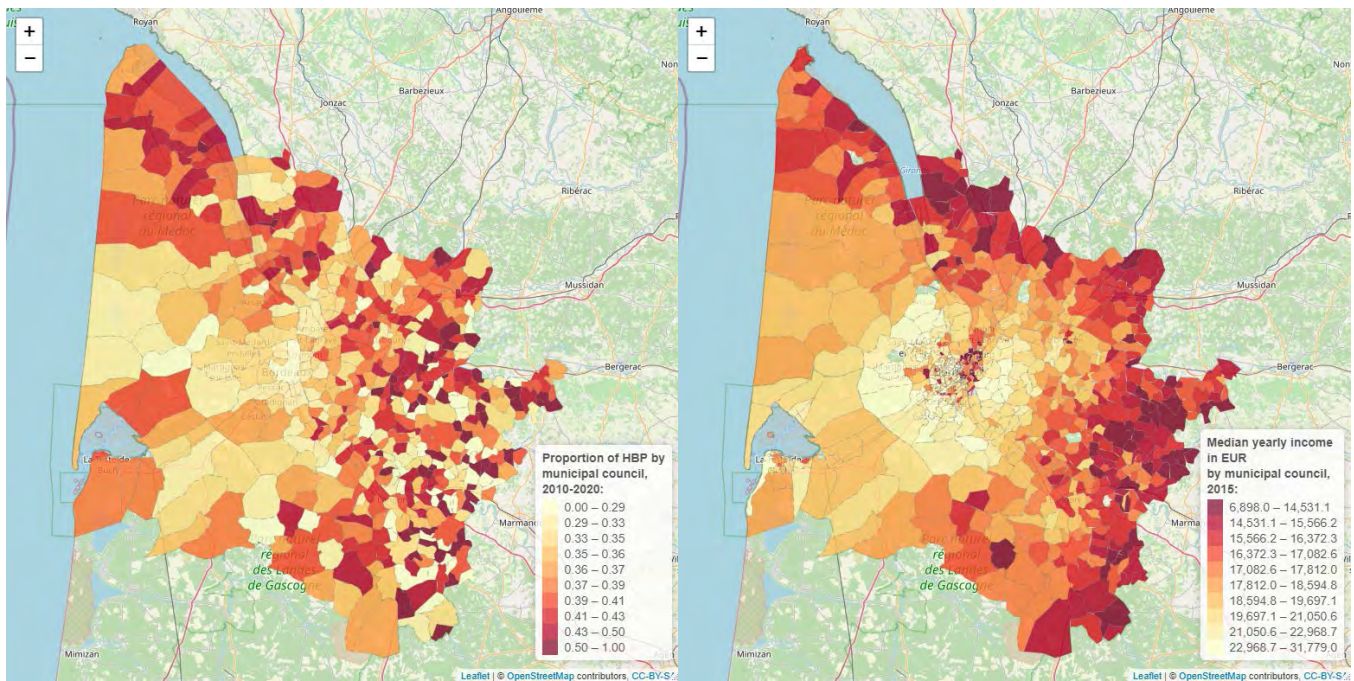


Figure 3: Proportion of HBP (number of HBP across all measured blood pressures) by municipal council for the period 2010-2020 (left) and median yearly income in EUR by municipal council for the year 2015 (right). Cut-off values were based on deciles.

ing rather than a diagnostic tool. Nevertheless, in the event that there is no association between suspected BPH and the municipal council, we should observe a similar distribution across the map.

This paper has illustrated the feasibility of implementing R Shiny to represent EMCC data to authorized end users. The current version of the dashboard is a pilot edition. It highlights the utility of EMCC data in public health surveillance (Yoon, Ising, and Gunn 2017). Visualizing EMCC data in near-real time through online dashboards is a pragmatic way to meet the epidemiology community's demand for up-to-date information. Scaling up such a tool to the national level will pose both technical and administrative challenges, particularly related to the current highly decentralized management of data from the French health centers.

Declarations

This work complies with the application framework provided by Article 65–2 of the amended French Data Protection Act and the General Regulation on the protection of personal data in terms of the protection of personal health data and the protection of privacy. It was approved by the head of the emergency department of the Bordeaux University Hospital.

Acknowledgments

We thank the anonymous reviewers for their insightful comments and suggestions.

Funding

This work was supported by the French National Research Agency (ANR) under the grant COSAM “Epidemiological surveillance of the COVID-19 pandemic period by real-time automatic classification of clinical notes from 15 emergency call centers using Transformer-based artificial neural networks” (project number ANR-20-COVI-0004-01). The authors' research teams had annual grants from the University of Bordeaux, INSERM U1219 and INRIA.

References

- Anstey, D.; Christian, J.; and Shimbo, D. 2019. Income inequality and hypertension control. *J Am Heart Assoc.* 8:e013636.
- Chang, W.; Cheng, J.; Allaire, J.; Sievert, C.; Schloerke, B.; Xie, Y.; Allen, J.; McPherson, J.; Dipert, A.; and Borges, B. 2021. *Shiny: Web Application Framework for R.*
- Cheng, J.; Karambelkar, B.; and Xie, Y. 2021. *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library.* R package version 2.0.4.1.
- Duijster, J.; Doreleijers, S.; Pilot, E.; Van der Hoek, W.; Kommer, G.; van der Sande, M.; Krafft, T.; and Van Asten, L. 2020. Utility of emergency call centre, dispatch and ambulance data for syndromic surveillance of infectious diseases: a scoping review. *Eur J Public Health.* 29:639–647.
- Duncan, E.; Best, C.; Dougall, N.; Skar, S.; Evans, J.; Corfield, A.; Fitzpatrick, D.; Goldie, I.; Maxwell, M.; Snooks, H.; Stark, C.; White, C.; and Wojcik, W. 2019. Epidemiology of emergency ambulance service calls related to men-

- tal health problems and self harm: a national record linkage study. *Scand J Trauma Resusc Emerg Med.* 27:34.
- Gil-Jardiné, C.; Chenais, G.; Pradeau, C.; Tentillier, E.; Revel, P.; Combes, X.; Galinski, M.; Tellier, E.; and Lagarde, E. 2021a. Surveillance of COVID-19 using a keyword search for symptoms in reports from emergency medical communication centers in Gironde, France: a 15 year retrospective cross-sectional study. *Intern Emerg Med.* 29:1–6.
- Gil-Jardiné, C.; Chenais, G.; Pradeau, C.; Tentillier, E.; Revel, P.; Combes, X.; Galinski, M.; Tellier, E.; and Lagarde, E. 2021b. Trends in reasons for emergency calls during the COVID-19 crisis in the department of Gironde, France, using artificial neural network for natural language classification. *Scand J Trauma Resusc Emerg Med.* 29:55.
- Holzer, B.; Minder, C.; Schätti, G.; Rosset, N.; Battegay, E.; Müller, S.; and Zimmerli, L. 2012. Ten-year trends in intoxications and requests for emergency ambulance service. *Prehosp Emerg Care.* 16:497–504.
- Man Lo, S.; Min Yu, Y.; Larry Lee, L.; Eliza Wong, M.; Ying Chair, S. J.; Kalinowski, E.; and Jimmy Chan, T. 2012. Overview of the shenzhen emergency medical service call pattern. *World J Emerg Med.* 3:251–6.
- Manca, F.; Lewsey, J.; Waterson, R.; Kernaghan, S.; Fitzpatrick, D.; Mackay, D.; Angus, C.; and Fitzgerald, N. 2021. Estimating the burden of alcohol on ambulance callouts through development and validation of an algorithm using electronic patient records. *Int J Environ Res Public Health.* 18:6363.
- Palafox, B.; McKee, M.; Balabanova, D.; AlHabib, K.; Avezum, A.; Bahonar, A.; Ismail, N.; Chifamba, J.; Chow, C.; Corsi, D.; Dagenais, G.; Diaz, R.; Gupta, R.; Iqbal, R.; Kaur, M.; Khatib, R.; Kruger, A.; Kruger, I.; Lanas, F.; Lopez-Jaramillo, P.; Minfan, F.; Mohan, V.; Mony, P.; Oguz, A.; Palileo-Villanueva, L.; Perel, P.; Poirier, P.; Rangarajan, S.; Rensheng, L.; Rosengren, A.; Soman, B.; Stuckler, D.; Subramanian, S.; Teo, K.; Tsolekile, L.; Wielgosz, A.; Yaguang, P.; Yeates, K.; Yongzhen, M.; Yusoff, K.; Yusuf, R.; Yusufali, A.; Zatońska, K.; and Yusuf, S. 2016. Wealth and cardiovascular health: a cross-sectional study of wealth-related inequalities in the awareness, treatment and control of hypertension in high-, middle- and low-income countries. *Int J Equity Health.* 15:199.
- Seo, M.; Begley, C.; Langabeer, J.; and DelliFraine, J. 2014. Barriers and disparities in emergency medical services 911 calls for stroke symptoms in the United States adult population: 2009 BRFSS survey. *West J Emerg Med.* 15:251–9.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L. u.; and Polosukhin, I. 2017. Attention is all you need. In Guyon, I.; Luxburg, U. V.; Bengio, S.; Wallach, H.; Fergus, R.; Vishwanathan, S.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.
- Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; and Funtowicz, M. and Brew, J. 2019. HuggingFace’s Transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.
- Xu, B.; Gil-Jardiné, C.; Thiessard, F.; Tellier, E.; Avalos, M.; and Lagarde, E. 2020. Pre-training a neural language model improves the sample efficiency of an emergency room classification model. In *Proceedings of the 33rd International FLAIRS Conference*, AAAI Press.
- Yoon, P.; Ising, A.; and Gunn, J. 2017. Using syndromic surveillance for all-hazards public health surveillance: Successes, challenges, and the future. *Public Health Rep.* 132(1 suppl):3S–6S.