HAL open science

Trends in reasons for emergency calls during the COVID-19 crisis in the department of Gironde, France using automatic classification.

Cédric Gil-Jardiné, Gabrielle Chenais, Catherine Pradeau, Eric Tentillier, Philipe Revel, Xavier Combes, Michel Galinski, Eric Tellier, Emmanuel Lagarde

To cite this version:

Cédric Gil-Jardiné, Gabrielle Chenais, Catherine Pradeau, Eric Tentillier, Philipe Revel, et al.. Trends in reasons for emergency calls during the COVID-19 crisis in the department of Gironde, France using automatic classification.. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, In press, 10.21203/rs.3.rs-106403/v1 . hal-03203146

HAL Id: hal-03203146

https://hal.science/hal-03203146v1

Submitted on 20 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Trends in reasons for emergency calls during the COVID-19 crisis in the department of Gironde, France using automatic classification.

Gil-Jardiné Cédric[1,2], Chenais Gabrielle[1], Pradeau Catherine[2], Tentillier Eric[2], Revel Philipe[1,2], Combes Xavier[1,2], Galinski Michel[1,2], Tellier Eric[1,2], Lagarde Emmanuel[1*].

Affiliations

1 Inserm, ISPED, Bordeaux Population Health Research Center Inserm U1219 "Injury Epidemiology Transport Occupation" team, Bordeaux cedex, France.

2 University Hospital of Bordeaux, Pole of Emergency Medicine, Inserm, ISPED, Bordeaux Population Health Research Center Inserm U1219 "Injury Epidemiology Transport Occupation" team, Bordeaux cedex, France.

Corresponding author: Emmanuel Lagarde. Emmauel.lagarde@u-bordeaux.fr

Abstract

Objectives

During periods such as the COVID-19 crisis, there is a need for responsive public health surveillance indicators related to the epidemic and to preventative measures such as lockdown. The automatic classification of the content of calls to emergency medical communication centers could provide relevant and responsive indicators.

Methods

We retrieved all 796,209 free-text call reports from the emergency medical communication center of the Gironde department, France, between 2018 and 2020. We trained a natural language processing neural network model with a mixed unsupervised/supervised method to classify all reasons for calls in 2020. Validation and parameter adjustment were performed using a sample of 20,000 manually-coded free-text reports.

Results

The number of daily calls for flu-like symptoms began to increase from February 21, 2020 and reached an unprecedented level by February 28, 2020 and peaked on March 14, 2020, 3 days before lockdown. It was strongly correlated with daily emergency room admissions, with a delay of 14 days. Calls for chest pain, stress, but also those mentioning dyspnea, ageusia and anosmia peaked 12 days later. Calls for malaises with loss of consciousness, non-voluntary injuries and alcohol intoxications sharply decreased, starting one month before lockdown.

Discussion

This example of the COVID-19 crisis shows how the availability of reliable and unbiased surveillance platforms can be useful for a timely and relevant monitoring of all events with public health consequences. The use of an automatic classification system using artificial intelligence makes it possible to free itself from the context that could influence a human coder, especially in a crisis situation.

Conclusion

The content of calls to emergency medical communication centers is an efficient epidemiological surveillance data source that provides insights into the societal upheavals induced by a health crisis.

Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Introduction

Coronavirus 2019 Disease (COVID-19) caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was detected for the first time in December 2019 in China[1]. Since then, as of first November 2020, more than 40 million cases and more than 1,100,000 deaths have been reported worldwide[2].

Health-related surveillance data are indispensable for the adjustment of public health policies aimed at curbing the spread of the virus and protecting vulnerable populations, for adapting the means of caring for the most seriously ill, and for assessing the indirect consequences of the pandemic and of the measures implemented to control it: lockdown, movement restriction, social distancing and quarantine for some countries.

Many countries have relied on an extrapolation of classic infection-control and public-health measures to contain the COVID-19 pandemic, similar to those used for SARS in 2003 and H1N1 in 2009. These range from extreme quarantine measures in China to painstaking detailed contact tracing with hundreds of contact tracers (e.g., Singapore, Hong Kong, South Korea). However, these measures may not be effective in 2020 for tackling the scale of COVID-19[3]. Monitoring, surveillance, detection and prevention of COVID-19 have launched a new challenge on a global scale and digital technologies have so far failed to be the driving force behind it with a few notable exceptions[4].

The number of people who test positives by PCR or chest CT images with characteristic lung damage is the most reliable indicator of the number of people with the virus. It is, however, heavily dependent on the screening strategy, which varies greatly from one country to another. The number of people entering an ER with symptoms suggesting SARS-CoV-2 infection was implemented nationally in France on March 16, 2020, with specific coding of the main diagnosis in ER summary reports centralized by Santé Publique France in the OSCOUR® Emergency Department Surveillance Network[5] set up in 2004. The French health authorities are also monitoring the number of COVID-19 patients hospitalized, admitted to intensive care units, and the number of deaths in hospital and in nursing homes (EHPAD).

All these tools, completed by a large set of ad-hoc surveys and specific surveillance systems have been mobilized to monitor the course of the epidemics and its consequences. For example, analysis of the registry of the Paris-Sudden Death Expertise Center observed a transient increase in cardiac arrest incidence[6]. A similar observation was reported in Lombardy, Italy, a region heavily impacted by the COVID-19 pandemic[7].

Reports of call content to emergency medical communication centers (EMCC) is a source of information that needs to be considered for health surveillance during such a pandemic. By dialing 15, French citizens can get medical advice and, if necessary, a medical mobile care unit can be sent to the scene. In addition, the health authorities had instructed the population to call 15 in the case of symptoms or concerns, and to avoid spontaneous uninvited visits to the emergency room (ER)[8].

We used an automatic classification tool based on an artificial neural network language model we recently adapted for free-text clinical notes[9] to identify the main reasons for calls to the Gironde EMCC, in order to monitor trends in the nature of these calls before, during and after the national lockdown established in France on March 17, 2020 and eased on May, 11, 2020.

Method

Setting

The Gironde department (1.6 million inhabitants) is served by a medicalized EMCC known as SAMU 33 (Service d'Aide Médicale Urgente de la Gironde) which answers calls to the French toll-free number dedicated to medical emergencies (the "15"). 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[5].

Clinical reports

For all cases handled, a clinical report is created in the form of a computerized free-text note and updated by a medical assistant and a physician through the various telephone interactions with the patient, family, witnesses, and then with the paramedics if applicable. These clinical notes contain all the elements that make it possible to know the circumstances of the event at the origin of the call and the clinical observations made by the protagonists, whether they are witnesses, the patients themselves, or the medical personnel involved on site or responding to the call.

Classification of clinical reports using the GPT-2 neural language model

Over the past 10 years, neural language models have progressively taken the largest share in the field of natural language, with applications like machine translation, document classification, text summarization and speech recognition.

New levels of performance have only recently been achieved with the use of models based on the concept of attention, which consists in learning dependencies between words in a sentence without regard to their distances. This mechanism has been implemented in a sequence to sequence neural network model, the Transformer architecture, proposed in 2017[12]. One of the latest examples is the GPT-2, published in February 2019 by OpenAI. GPT-2 is a large transformer-based language model trained on a dataset of 8 million web pages[13]. Beyond the capability to generate coherent texts, Transformer models have the potential to perform other tasks such as question answering and document classification with a limited number of training examples. The training of the model is performed in two distinct phases[14]: the first pre-training unsupervised phase (i.e. not needing a human classification), consists in exploitation of a text corpus. In our application, this consisted in the 312,367 clinical reports of the year 2018. This leads to the ability of automatic text generation (see Supplementary material 1). The relevance of these synthetic sentences suggests that the networks learn contextual semantic representations. The second training phase, this time a supervised one (i.e. using examples classified by humans), creates a system able to perform the specific classification tasks.

The 117-million parameters version of the GPT-2 model was trained on a workstation with one Nvidia GeForce RTX Titan Graphic Processing Unit with 24GB of VRAM. The computation time for the pre-training phase was 26 hours per epoch. On average, 12 samples were processed per second in the training phase.

Extraction of free-text digital calls reports from the EMCC of the Gironde department, France

We retrieved call reports from the digital medical record system of the EMCC of the Gironde department, France. Five EMCC datasets were extracted corresponding to the years 2016, 2017,2018, 2019 and 2020. The three 2016, 2017, 2018 datasets consisted of 888,469 calls with a free-text clinical report and were used both for the unsupervised pre-training of the GPT-2 model and for the building of the training sample based, when available, on standardized diagnosis coded during the calls. The 2019 dataset corresponded to the 302,925 free-text clinical call reports of the year 2019. A random sample of this dataset was manually coded by a trained pool of emergency nurses for validation. The trained and validated models were applied to the 2020 dataset which corresponded to the COVID-19 period, with 255,556 handled calls with a free-text report from January 1, 2020 to September 30, 2020. This provided a classification of reasons for calls before (January 1-March 16), during (March 17-May 10) and after (May 11-September 30) COVID-19 lockdown.

Selected groups of reasons for calls

A first model was built by grouping calls according to the main reasons into 12 broad categories (chest pain, gastroenteritis and abdominal pain, flu-like symptoms and breathing difficulties, focal neurologic 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). Many symptoms or situations did not fall into these broad categories and were grouped into a thirteenth "other" category.

Another distinct model was built to count calls for which an acute alcohol intoxication was involved.

Classification procedure

Classification models were built and validated using a five-phase procedure:

(i) A total of 888,469 records with call reports of 2016, 2017 and 2018 were used to pre-train the GPT-2 model in an unsupervised manner with 1 epochs.

(ii) A dataset of training examples for supervised training was built extracting the 690,103 (out of 888,469) call reports from the same 2016, 2017, 2018 years with a diagnosis code manually coded in real time by the people handling calls (medical regulation assistants, emergency physicians and general practitioners); The pre-trained model was then fine-tuned using these training labelled examples to build the model used to classified main reasons for call in 12 categories and a distinct model to identify calls for which an acute alcohol intoxication was involved.

(iii) These two models were validated using the manual coding of a 39,907 random sample from the 2019 dataset.

(v) All 254,633 call reports of 2020 (January to September) were classified using the two models built in the previous steps;

Emergency room admissions with suspected COVID-19 and SARS-CoV2-2 PCR tests result

We retrieved daily aggregated data of ER admission with suspected COVID-19 and the daily number of SARS-Cov-2 PCR positive tests in the department of Gironde from the national public health agency “Santé Publique France”. All data are available on the Geodes website (https://geodes.santepubliquefrance.fr/).

Analysis

For all call reasons, a probability cutpoint was determined in order to optimize the accuracy of the number of calls estimated by the model. This was done by selecting the cutpoint for which precision and recall are equal in the validation sample (i.e. when the number of false positive predictions equals the number of false negative predictions). The confidence of the method was assessed using a bootstrap procedure with 10,000 1:1 random partitions of the validation sample. For each partition (and each reasons), the cutpoint is selected using the first half of this partition sample and is the one for which precision equals recall. The difference between the manually determined prevalence of the given reasons and the predicted prevalence (using the selected cutpoint) is then computed in the second half of the sample. Median bias and 95% confidence interval were derived from the distribution of these 10,000 bias measures. Estimated daily counts of calls to EMCC were plotted against time with a 7-days moving average window to smooth the one-week periodicity in calls, which were systematically 25% more frequent on Saturday and Sunday.

Results

Model training, parametrization and validation

The unsupervised training phase led to a model capable of generating artificial texts with the same structure as the learning texts, containing essentially the same clinical notions, but whose coherence is often inconsistent (see examples in the appendix). The content of these synthetic reports however suggests that the model has learned contextual semantic representations.

The training sample included from 1,829 (road traffic crash) to 412,218 (other reasons) samples per group (see Table 1). Performance statistics computed using the 39,907 manually-coded samples are shown in Table 1. Area under the ROC curve ranged from 0.785 to 0.990.

The optimal cutpoints to minimize bias in prevalence estimates were chosen using the same manually-coded sample of 39,907 call reports. Bootstrap 95% confidence interval assessment showed that the method provided narrow confidence intervals (Table1).

Reasons for calls during the Covid-19 crisis

In 2020, the median daily number of calls to the EMCC with a report was 892 with a total number of 254,633 and a peak of 1,926 on March 14. The distribution of the estimated number of calls per group of reasons and according to the lockdown period is shown in Table 2.

Three groups exhibited a peak at the onset of the lockdown period. Figure 1 shows that calls for flu-like symptoms and breathing difficulties began to rise on February 21, 20 days before the rise in ER admission for COVID-19, and peaked on March 14 with 928 calls out of a total of 1,926 calls that day. By the 16th day after February 21, the number of calls in this group had reached 403 calls per day, which is higher than all levels over the past 5 years. Starting from March, 1, 2020, the correlation between daily calls for flu-like symptoms and daily ER admissions with a delay of 14 days was 0.80 ($p < 10^{-15}$; Pearson's test). Two other reasons, chest pain and stress and anxiety, were found to peak 12 days after March 14, with a curve of the same shape as the number of ER visits for COVID-19.

Figure 2 shows that calls for alcohol intoxication, road traffic crash and malaise with loss of consciousness experienced a sharp decline around the onset of lockdown and that these trends started several days before the first day of lockdown. In a lesser extent, the latter curves were parallel with those for violence and injuries other than violence and road crash. Finally, no noticeable trends in relation to lockdown was found for other groups of reasons including gastroenteritis and abdominal pain, stroke, suicide and self-harm, pregnancy and delivery problems.

Discussion

The number of calls to EMCC showed a significant spike centered on the March 14, the day before the closure of all non-essential public places in France, three days before lockdown. The main reasons were related to flu-like symptoms (cough and/or fever), followed 14 days later by a peak in ER admissions, in calls for chest pain and calls for stress and anxiety. Calls for road traffic crashes, malaises with loss of consciousness, non-voluntary injuries and alcohol intoxication fell by 59%, 38%, 24% and 23% respectively during lockdown.

Probably the most interesting finding of our study is the delay we observed between the rise in calls for flu-like symptoms (mainly cough and fever) and the rise in ER visits for suspected COVID-19. Thus, the curve began to rise 20 days before the increase in ER visits. By the 16th day after the start of this rise, the levels reached a level higher than any levels in the past 5 years. One could hypothesize that the peak of calls recorded around March 14 was due by the concern, if not anxiety, caused by the announcement on television of the closure of public places by the French President on that day. However, in a more affected part of the country, the Ile-de-France region, the peak was reached much earlier (10 days earlier)[15], suggesting that most of the calls we recorded were more motivated by symptoms than by concerns raised by communication by the authorities. Further, a spike in calls for stress and anxiety was measure 14 days later. EMCC call content is therefore probably the most predictive early indicator of the start of the epidemic, as recently shown by Riou and colleagues who found in the Ile-de-France region a strong correlation between calls regarding suspected COVID-19 and the number of patients in intensive care, with a delay of 23 days[15]. This is why this is considered for the monitoring of a potential relapses in the epidemic [16]. Finally, while the number of calls for flu-like symptoms proved to be an early and relevant signal, its intensity was probably increased by the request of the authorities not to go directly to the ER and to contact instead the EMCC.

An important difference between the work of Riou and colleagues and this study is that our process was clearly agnostic to the COVID-19 epidemic or to lockdown, as the automatic classification used models trained using reports from previous years. This results in a procedure that remains independent of the COVID-19 epidemic and lockdown, which would have influenced human codification. The signal thus obtained depends less on the context and is more likely to be an indicator of the actual public health situation. More generally, the added value of using an automatic classification procedure based on a natural language processing model is that it frees us from the context in which the reported events are coded. For this reason, we did not use the coded diagnoses at the time of the call to observe trends. In addition, these diagnoses were absent from one-third of the reports.

In the context of the COVID-19 epidemic, several research teams have used a similar approach, attempting to investigate the internet or social media to build early indicators of the epidemic[11, 17]. However, no such signal could be found from a Google keyword search[18], as the peak for cough, fever, coronavirus or COVID-19 was not reached until the week of 15-21 March.

Contrary to what was observed in Paris[19], no increase in calls related to cardiac arrest was observed in our study. This observation supports the hypothesis that the transient twofold increase in out-of-hospital cardiac arrests observed in Paris and its suburbs could be due to COVID-19 infections and to pandemic-related health system issues in heavily impacted regions. This was clearly not the case in the Gironde department where EMCC and intensive care units have never been overwhelmed.

A very significant decrease in calls for malaise with loss of consciousness, and to a lesser extent for strokes, was observed, starting one week before lockdown. This paralleled the sudden drop in ER visits that was observed in many countries that issued a statewide stay-at-home order[20], raising concerns that patients who needed medical care were not presenting to the hospitals and, for example, that stroke patients arrived too late to receive tissue plasminogen activator. The actual overall public health impact of this phenomenon will have to be carefully assessed when we have enough hindsight to appreciate its medium-term health consequences.

The decrease in calls associated with interpersonal violence and alcohol intoxication is less surprising as it is probably due to the reduction in social interactions. Interestingly, the figures returned to normal levels by the end of the lockdown period. Early on, concerns were raised about the risk of domestic violence as a result of lockdown[21]. This was not confirmed here by calls to EMCC. Although, unfortunately, not all domestic violence is reported to EMCC, this is an interesting result because most statistics used during lockdown to estimate the incidence of intimate partner violence were derived from Police reports and not all violent events reported to EMCC are reported to the police.

In order to produce results in a time frame compatible with the health emergency related to the recent lifting of lockdown measures, we used the samples from 2016-2018 for which a diagnosis was coded during the call by the medical assistant in charge of handling it. The ideal procedure would have been to perform a manual coding of this training sample, which was done for a sample of 39,907 reports from 2019, but retained in this work as a validation

sample. Our previous work has shown that it takes about a thousand different examples to maximize the performance of the model[9]. This would have meant manually coding more than 100,000 notes, a task that was out of reach in a short period of time. The performances of the GPT-2 model measured with the manually coded validation samples were, however, very high and allowed us to derive a reason for call for all reports including the 22% of them with a missing value for the diagnosis.

Some limitations need to be acknowledged. First, although we have shown that an AI-based natural language model shows high performance in classifying free-text clinical reports, a small proportion of reports remains misclassified. Because our exercise here was to provide prevalence estimates, we adjusted the decision cutpoints so that precision equals recall. The bootstrap analysis showed that this was a very reliable strategy. Second, not all calls are handled by the EMCC, a proportion of them remains unanswered and this proportion increases during peak periods. It is therefore likely that around March 14 the number of attempted calls was higher than those handled. Finally, the study was done in Gironde, a department with a reportedly low rate of SARS-Cov-2 infection if compared to the Ile de France and the north-east regions of France. However, lockdown and fear of the epidemic affected all French people and the Gironde EMCC are the third largest in terms of the number of calls received in France, which has made it possible to build up a sufficiently large database.

Conclusion

Given the fast spread and severity of COVID-19, the availability of reliable surveillance platforms is crucial for timely monitoring and for responding with adequate control measures to the COVID-19 epidemic and others to come, together with other major events with public health consequences. The results of this study, carried out in the context of a period of lockdown related to the COVID-19 epidemic, illustrate the extent to which automated classification of the reasons for calling EMCC is a powerful epidemiological surveillance tool, provides insights into the societal upheavals induced by a health crisis and would be instrumental to better anticipate the needs of the health care system.

References

1. Zhu N et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*; 382: 727–33. DOI: 10.1056/NEJMoa2001017 (2020)

2. WHO. Coronavirus (COVID-19) events as they happen. https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed Nov 1, 2020) (2020).

3. Ting DSW et al. Digital technology and COVID-19. *Nat Med* **26**, 459–461. https://doi.org/10.1038/s41591-020-0824-5 (2020).

4. Park S, Choi GJ, Ko H. Information Technology–Based Tracing Strategy in Response to COVID-19 in South Korea—Privacy Controversies. *JAMA* **323**(21):2129–2130. doi:10.1001/jama.2020.6602 } (2020)

5. Josseran L et al. Assessment of a syndromic surveillance system based on morbidity data: results from the Oscour network during a heat wave. *PLoS One* **5**(8):e11984. Published 2010 Aug 9. doi:10.1371/journal.pone.0011984 (2010).

6. Marijon, E et al. Out-of-hospital cardiac arrest during the COVID-19 pandemic in Paris, France: a population-based, observational study [published online ahead of print, 2020 May 27]. *Lancet Public Health*. S2468-2667(20)30117-1. doi:10.1016/S2468-2667(20)30117-1 (2020).

7. Baldi E et al. Out-of-hospital cardiac arrest during the Covid-19 outbreak in Italy. *N Engl J Med*; published online April 29. DOI:10.1056/NEJMc2010418 (2020).

8. Adnet F, Lapostolle F. International EMS systems: France. *Resuscitation*.;**63**(1):7-9. doi:10.1016/j.resuscitation.2004.04.001 (2004).

9. Xu, B. et al. Pre-Training a Neural Language Model Improves the Sample Efficiency of an Emergency Room Classification Model. *FLAIRS-**32*** - Thirty-Third International Flairs Conference, May 2020, North Miami Beach, United States. (2020).

10. Huang J, Osorio C, Wicent Sy L. An empirical evaluation of deep learning for icd-9 code assignment using mimic-iii clinical notes. *Computer Methods and Programs in Biomedicine*, **177**:141 – 153,. https://doi.org/10.1016/j.cmpb.2019.05.024 (2019).

11. Li C et al. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Euro Surveill* **25**: ES.2020.25.10.2000199. (2020).

12. Vaswani A**.** et al. Attention Is All You Need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems* **30**, pages 5998–6008. Curran Associates, Inc., NY, United States (2017).

13. Radford A et al. Language models are unsupervised multitask learners. *OpenAI Blog*, **1**(8), (2019).

14. Radford A, Narasimhan K, Salimans T, and Sutskever I. Improving language understanding by generative pre-training. *Open AI* (2018) https://s3-us-west-2. amazonaws. com/openai-assets/researchcovers/languageunsupervised/language understanding paper. pdf (2018).

15. The COVID-19 APHP-Universities-INRIA-INSERM Group, Riou, B. Emergency calls are early indicators of ICU bed requirement during the COVID-19 epidemic. *medRxiv 2020.06.02.20117499*; doi: https://doi.org/10.1101/2020.06.02.20117499 (2020).

16. Salje H et al. Estimation of the burden of SARS-CoV2 in France. *Science* **367**: [Epub ahead of print May 13:eabc3517] (2020).

17. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Niakan Kalhori RS. Predicting COVID-19 incidence through analysis of Google trends data in Iran: Data mining and deep learning pilot study. *JMIR Public Health Surveill* 2020; **6**: e18828.DOI: 10.2196/18828 (2020).

18. [Google Trends 2020]. https://trends.google.fr/trends/?geo=FR. (accessed june 3, 2020) (2020).

19. Lapostolle F, Agostinucci JM, Alhéritière A, Petrovic T, Adnet F. Collateral consequences of COVID-19 epidemic in Greater Paris. *Resuscitation*.;**151**:6-7. doi:10.1016/j.resuscitation.2020.04.010 (2020).

20. Wong LE, Hawkins JE, Langness S, Murrell KL, Iris P, Sammann A. Where are all the patients ? Addressing COVID-19 fear to encourage sick patients to seek emergency care. *NEJM Catalyst* May 14 (2020).

21. Mazza M, Marano G, Lai C, Janiri L, Sani G. Danger in danger: Interpersonal violence during COVID-19 quarantine [published online ahead of print, 2020 Apr 30]. Psychiatry Res. 289:113046. doi:10.1016/j.psychres. 113046 (2020).

Figure and Table legend

Table 1. Training sample size (from 2016-8 datasets) and validation from manually coded samples from 2019 dataset

Table 2. Reasons for calls to EMS in 2020 as determined by the GPT-2 model. Calls per day and per motive

Figure 1. Number of calls to the Gironde to emergency medical communication center in 2020 by reasons from GPT-2 classification with a significant increase around the onset of the lockdown period.

Figure 2. Number of calls to the Gironde to emergency medical communication center in 2020 by reasons for calls from GPT-2 classification with a significant decline around the onset of the lockdown period.

Supplementary material 1. Example of real and synthetic EMS clinical reports.

Confidentiality, ethics and data protection

No personal data were necessary for this work. All reports were automatically de-identified using a grep-based text-search procedure that was applicable because of the standardized format of personal information inserted in the reports. This work conforms in terms of the protection of personal health data and the protection of privacy to 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. It was approved by the Bordeaux Teaching Hospital committee for ethics and data protection.

Consent for publication

Not applicable

Availability of data and materials

A subset of the dataset analysed during the current study can be made available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The study was founded by the French Agence Nationale de la Recherche (ANR-20-COV1-0004-01).

Authors' contributions

CGJ and PC collected the data. GC and EL were involved in NLP models design. CGJ and EL analyzed the data and all authors were involved in results interpretation. EL and CGJ was a major contributor in writing the manuscript. All authors read and approved the final manuscript.

Acknowledgement

We thank the university hospital of Bordeaux for providing the logistical support that allowed us to access and analyse the data needed for the manuscript in such a short period. We are also grateful to Julien Anjoubault, Clarisse Marguinaud, Virginie Cocuelle, Delphine Vauthier, Alexandra Barbe, François Garreau, Quentin Bana, Claire Riou, Pauline Soubelet and Elisabeth Verbitskaya for their expertise, which allowed proper manual coding for validation and to Benjamin Contrand, Loïck Bourdois and Marie-Odile Coste for data management and administrative assistance. We also thank Sylviane Lafont for her help in the beginning of the project. BPH IETO Team activities are supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), University of Bordeaux, Ministère de l'Intérieur (Délégation à la Sécurité Routière).

Tables and Figures

Table 1. Training sample size (from 2016-8 datasets) and validation from manually coded samples from 2019 dataset

Reasons	Training sample size	Max F1	AUC	Accuracy	Proportion in 2020	95% confidence interval using bias estimated by bootstrap*
Main reasons for EMS calls						
Chest pain	29,310	0.800	0.987	0.978	0.057	0.0542 - 0.0598
Gastroenteritis and abdominal pain	63,446	0.710	0.958	0.946	0.082	0.0777 - 0.0863
Flu-like symptoms and breathing difficulties	72,323	0.683	0.958	0.929	0.151	0.1461 - 0.1559
Focal neurologic deficit, stroke	5,951	0.698	0.978	0.991	0.0135	0.0118 - 0.0152
Road traffic crash (RTC)	1,829	0.799	0.980	0.988	0.0233	0.0214 - 0.0252
Violence	3,158	0.636	0.984	0.991	0.011	0.0092 - 0.0128
Suicide and self-harm	5,904	0.654	0.969	0.988	0.015	0.0131 - 0.0169
Injury other than violence, self-harm and RTC	120,007	0.694	0.938	0.887	0.166	0.1596 - 0.1724
Pregnancy and delivery problems	6,222	0.804	0.990	0.994	0.013	0.0116 - 0.0144
Malaise with loss of consciousness	41,468	0.492	0.935	0.958	0.035	0.0313 - 0.0387
Stress and anxiety	12,198	0.479	0.877	0.956	0.046	0.042 - 0.05
Other reasons	412,218	0.673	0.785	0.746	0.385	0.3754 - 0.3946
Alcohol intoxication	8,934	0.712	0.982	0.979	0.033	0.0303 - 0.0357

* As estimated by bootstrapping (N=10,000) the validation sample

Table 2. Reasons for calls to EMS in 2020 as determined by the GPT-2 model. Calls per day and per motive

Motives	**Before lockdown** (calls per day)	**During lockdown** (calls per day))	**After lockdown** (calls per day)
Total	N=79,672	N=49,743	N=125,218
Main reasons for EMS calls			
Chest pain	5.1	7.2	5.5
Gastroenteritis and abdominal pain	8.9	7.3	8.2
Flu-like symptoms and breathing difficulties	20.7	22.7	8.7
Focal neurologic deficit, stroke	1.3	1.3	1.4
Road traffic crash (RTC)	2.2	0.9	3.0
Violence	0.10	0.08	0.13
Suicide and self-harm	1.3	1.4	1.6
Injury other than violence, self-harm and RTC	15.0	11.4	19.7
Pregnancy and delivery problems	1.2	1.3	1.4
Malaise with loss of consciousness	3.4	2.1	4.0
Stress and anxiety	3.8	5.5	4.8
Other reasons	35.4	36.6	41.3
Alcohol intoxication	3.0	2.3	3.9

Figure 1. Number of calls to the Gironde to emergency medical communication center in 2020 by reasons from GPT-2 classification with a significant increase around the onset of the lockdown period.

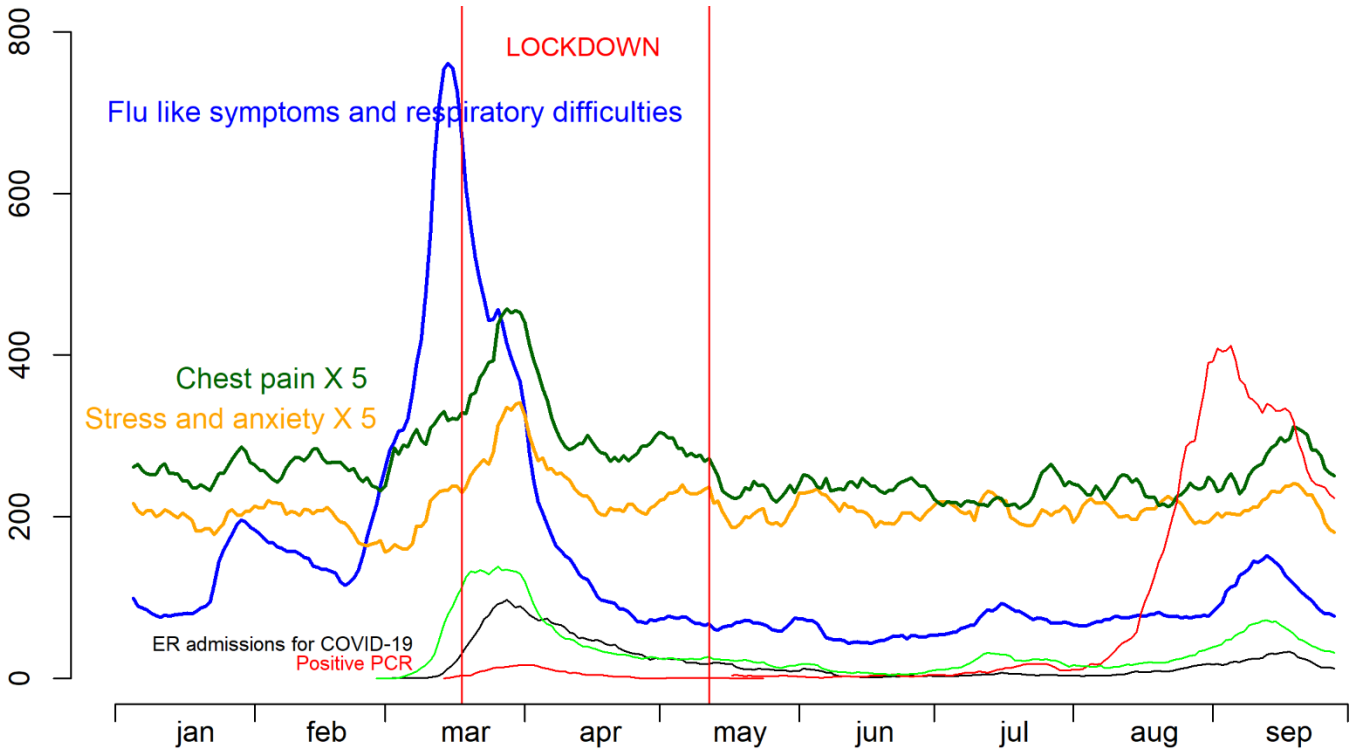
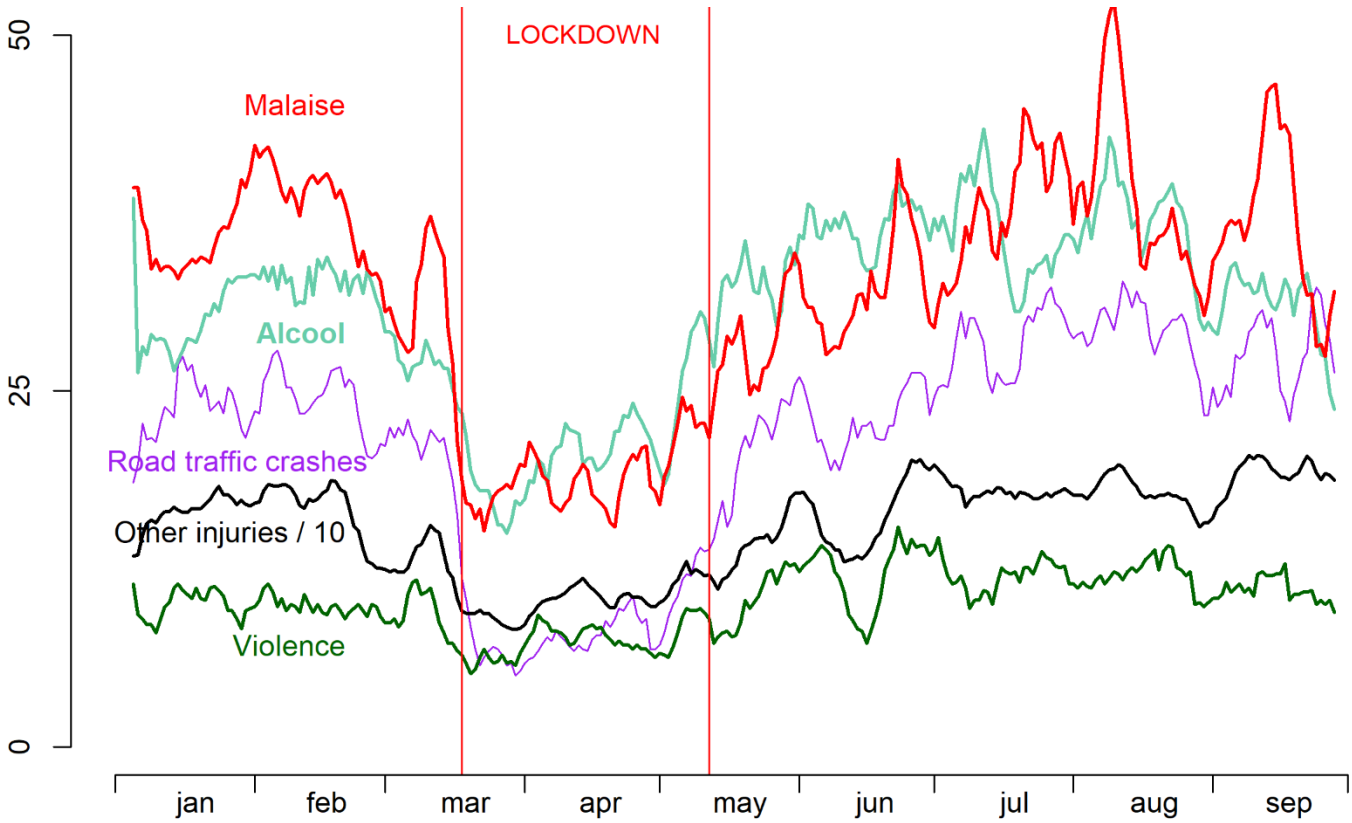


Figure 2. Number of calls to the Gironde to emergency medical communication center in 2020 by reasons for calls from GPT-2 classification with a significant decline around the onset of the lockdown period.



Supplementary material 1. Example of real and synthetic EMS clinical reports.

Real EMS reports
homme de 68 ans douleurs abdominales dernière selle ce matin pas de gaz vomissement il y a une heure atcd subocclusions sur brides(2) atcd dnid/irénale ap pour pole santé pas d'ambulance de garde sur le secteur tel ne repond pas sp carence h 68 a pas de pci cst fv 16 sat 94 aa puls 83 ta 146/66 bd vomissements
a raccroche avant conf gene respi je fais contre appel messagerie je rappelle a nouveau je reussi a avoir la patiente depuis vendredi grippe et gastro vomissements ne s alimente pas a fait malaise pas de gene respi vendredi médecin ttt / grippe et gastro vomissement ne s'alimente pas et ne boit pas non plus vertige douluer épigastrique >>>>sp en urgence f76a dl ventre et estomac depuis 3 jrs vu med vendredi gastro dl 8/10 sat 95 cste fv26 ta g 124/65 pas de gene respi a pas manger depuis 3 jrs
a bu et elle vomit consciente alcoolisée + toxique consciente >> ap pour hsa pas d'ap de garde sur secteur 8 sp carence f 16 a alcolisee + drogue puls70

ta 98/62 bd
sat 100 aa

ne mange pas / etourdi / diarrhée / s'urine dessus / ne veut pas aller a l'hopital
diarrhée
vertiges
toux
courbatures
atcds : 0
ttt : 0
>> sos
sos rappelle por demande de smur
cyanose
marbrures
pas d'atcd
appel pour asthénie / vertiges
8 de tension
cyanosé marbré
apyrétique
point appel pulmonaire
mis sous 02
==> mieux.
on va envoyer sp
h 69
sos med sur place
conscient
orienté
sos sur ll
pupilles sym
fv 26 diff sensation ettouff
parole poss
sat 84
cyanose
marbrures
fc115 reg
ta g 80/50
mhta
pas d'hospi
pas de tt
pas d'allergie
02 9l sat
sat a 93
refus des 4 pav et de libourne
demande exclusivement pellegrin
regul r evrard hsa

lieu de travail pense fatigue malaise pci ne reagit pas
mis en pls / insct respi /
f 34 a
dit avoir etait droguer
cst
puls 82
sat 100 aa
ta 127/87

Synthetic EMS reports
dl bras dt depuis dimanche / pas de fourmillement F 41a Douleur thoracique d'apparition brutale en se frottant les bras cette nuit sans faux mouvement Bouge bien sa jambe ni l'estomac Fumeuse ATCD : néant Douleur de novembre 2017 : RAS Pas de dyspnée pas de position antalgique >> SPazl, doliprane 1000 f 41 f 41 dl bras dt dl cou cervicales pas de céphalées a deja eu ca mais pas consulté : peut etre pire fc 57 ta 118/87 sat 99
toux fievre toux de nouveau febrile 39 toux seche pas de gene respi ni malaise ATCD: neant pas dautre symptome vomissements x 1 >> conseils >> rappel si besoin rappel de la mere toux souhaite mdg dans un premier temps virose selon la mere rhino toux rauque lavage de nez pas de gene respi maman sur LL toux rauque pas d autres signes de gravite mdg ok
chute / dl epaule dte / pas de pc / antcd prothese hanche dte chute a glissé douleur épaule adresser les sp Bilan SP : h 90, dl epaule conscient : oui ATCD : PLS : 72 TA : 132/67 Sat : 100 F Resp : 22
sage femme du bras elle dit qu'elle est triste dit qu'elle ne peut pas se deplacer

f 50 cste bien oriente pas de pc d apres sa femme de la dame a eu rale et fatigue antcd pbl respi + stents en 2017 p 64 sat 100 ta 142/67 refuse evac rappel sp : ne va pas bien dans cet etat pas de plainte est ok pour aller a l hopital
chute, plaie du front / assis par terre, assis sur le canapé, désorienté chute relevé par sa fille a du mal a marcher conscient ****pompiers pour pellegrin Bilan VSAV H 96 ans conscient chute de sa hauteur chariotale en bas pas de douleur a du mal a parler bien oriente FC 85 SAT 98 TA G 143/89 pas de trauma pas de saignement sous anticoagulant pas de d'anticoagulant