Using Data Mining Techniques to Understand Patterns of Suicide and Reattempt Rates in Southern Brazil

Caibe A. Pereira[1][a], Rômulo C. R. Peixoto[2][b], Manuella P. Kaster[1][c], Mateus Grellert[3][d] and Jônata T. Carvalho[2][e]

[1]*Biochemistry Department, Federal University of Santa Catarina, Florianopolis, Brazil*
[2]*Informatics and Statistics Department, Federal University of Santa Catarina, Florianopolis, Brazil*
[3]*Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil*

Keywords: SUS, Health Infrastructure, Machine Learning.

Abstract: Suicide is a multifactorial, complex condition and one of the leading global causes of death, with suicide attempt as the main risk factor. To this day, studies have shown relevant indicators that help identify people with risk of committing suicide, but the literature still lacks comprehensive studies that evaluate how different risk factors interact and ultimately affects the suicide risk. In this paper, we aimed to identify patterns in data from the Brazilian Unified Health System – SUS, from 2009 to 2020, of individual reports of suicide attempts and suicide deaths in the Brazilian Southern States, integrating those with a database of the healthcare infrastructure. We framed the problem as a classification task for each micro-region to predict suicide and reattempt rate as low, moderate, or high. We developed a pipeline for integrating, cleaning, and selecting the data, and trained and compared three machine learning models: Decision Tree, Random Forest, and XGBoost, with approximately 97% accuracy. The most important features for predicting suicide rates were the number of mental health units and clinics, and for both suicide and reattempts were the number of physicians and nurses available. This novel result brings valuable knowledge on possible directions for governmental investments in order to reduce suicide rates.

1 INTRODUCTION

Suicide is one of the leading causes of death, with approximately 700,000 deaths annually. Suicides are preventable, and prevention programs can mediate this problem. To be more effective, the programs must be directed to the risk populations (WHO, 2014). Suicide attempts are one of the main risk factors, and each additional attempt increases the risk. Moreover, detecting the risk of suicide is an open challenge in mental health research (Gao et al., 2015).

The stress-diathesis model for suicidal behavior associates biological and psychiatric traits with environmental stressors (van Heeringen & Mann, 2014), increasing the complexity of this condition. The environment has a significant contribution to suicidal behavior during the spectrum from suicide ideation to completion (Turecki et al., 2019). Environmental factors, such as biological (e.g., noise and pollution) and social stressors (e.g., crime and harassment fear) and infrastructure (e.g., green areas, traffic) influence mental health (Johnson et al., 2023).

The studies that analyze suicide behavior or deaths usually use sociodemographic or clinical features, and not the environment context of the individual. Thus, in this study, we propose a data mining and machine learning workflow to identify patterns of healthcare infrastructure to predict suicide reattempt rate (SAR) and suicide rate (SR). We used aggregated data from individuals who attempted or committed suicide and healthcare infrastructure in the Southern region of Brazil to predict rates.

[a] https://orcid.org/0000-0002-9719-2832
[b] https://orcid.org/0000-0001-6319-9668
[c] https://orcid.org/0000-0003-0258-6204
[d] https://orcid.org/0000-0003-0600-7054
[e] https://orcid.org/0000-0001-9020-2076

Pereira, C., Peixoto, R., Kaster, M., Grellert, M. and Carvalho, J.
Using Data Mining Techniques to Understand Patterns of Suicide and Reattempt Rates in Southern Brazil.
DOI: 10.5220/0012350500003657

In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2024) - Volume 2*, pages 385-392
ISBN: 978-989-758-688-0; ISSN: 2184-4305

Data were collected from the Brazilian Mortality Information System (SIM) and Notifiable Diseases Information System (SINAN), part of the Unified Health System platform (DATASUS). In this paper, we selected a case study in the southern states of Brazil: Paraná (PR), Santa Catarina (SC), and Rio Grande do Sul (RS). The SIM form is performed *post-mortem*; thus, it is more prone to missing values, especially if the deceased relatives are not present.

This paper is organized as follows: Section 2 presents an overview of suicide, with a focus on Brazilian data and machine learning studies of suicide. Section 3 and 4 specifies the data collection, treatment, and analysis for this study, and the machine learning models used. Section 5 presents a discussion of the results and concludes the study. To the best of our knowledge, this is the first work that proposes a learning-based solution for suicide or attempted suicide in Brazil using large-scale public health data. We show here how important nurses' and physicians' availability and basic and mental health facilities are to classify suicide risk.

2 BACKGROUND

In Brazil, suicide attempt is a compulsory notification and every attempt is registered in the SINAN. The notification started to be compulsory in 2011, but in a gradual manner; thus, we chose more recent years to analyze. Suicide attempt notification occurs in the healthcare unit where the victim was admitted. SINAN form is used for this purpose. It is an individual form divided into ten domains, comprising sociodemographic data about the individual with details about the attempt (Ministry of Health, 2021).

All Brazilian healthcare-related infrastructure is sent to the National Registry of Healthcare Establishments (CNES) database, with different categories. CNES is monthly updated and we collected data from the southern states between 2009 and 2020. We gathered information on types of establishments, professional specialization, mental health services.

2.1 Suicide Data in Brazil

Brazil is a continental country divided into five macro-regions (North, Northeast, Central-West, Southeast, and South). SR and SAR differ between regions, with the highest rates in the southern region. Suicide in Brazil follow global trends, such as sex ratio or methods (Ministry of Health, 2021). Registry of suicide deaths in Brazil is compulsory and classified according to the International Classification of Diseases (ICD)-10 since 1996. Data are publicly available in the DATASUS. The system provides information on the death circumstance, location, and personal information.

2.2 Related Work

2.2.1 Brazilian Studies

The studies that evaluated Brazilian SR and SAR data were exploratory, analyzing the profile of the individuals but with no predictive power. For SR, data was collected from DATASUS; SAR data was collected from different databases. In the Ministry of Health (2021) report, from 2010 to 2019 SR showed an increase of 42% in Brazil. The Southern was the region with the highest SR. Moreover, attempt at younger ages was more prevalent and mainly occurred at home by self-poisoning.

Other studies focus on different aspects of suicide in Brazil. From 2000 to 2017, Brazilian SR increased significantly but differently by sex: male SR increased by 75% and female by 85% (Palma et al., 2021). Male-to-female ratio is even higher in older people, reaching 8.2-fold more male deaths (Martini et al., 2019). In adolescents, SR also increased over the years and was associated with social inequality and unemployment (Jaen-Varas et al., 2019).

In SC, a comparison between suicide attempters and non-suicidal self-injured adolescents and adults from 2014 to 2018 shows common traits between groups, such as the majority of females, White people, people with psychiatric disorders, and poisoning as the method used. Comparing the recurrence of the attempt, adults attempt more than adolescents and with a higher percentage of alcohol intoxication during the attempt (Pinheiro et al., 2021).

2.2.2 Machine Learning and Suicide

Electronic records were used to assess medical databases and predict suicide attempts from self-injured patients. The prediction was performed at different time points using the Random Forest (RF) model, with prediction accuracy decreasing with time (Walsh et al., 2017). RF was used to predict attempts in males and females with cocaine abuse; for females, psychiatric issues showed more importance, whereas for males, it was drug use (Roglio et al., 2020).

Different machine learning tests were used to predict suicide in patients with depressive disorder. Attempt, ideation, race, religion, and depressive severity were the main factors for an attempt (Nordin et al., 2021). In addition, socioeconomic variables (e.g., income, age, drinking, education) were the main variables to predict suicide attempt and ideation by XGBoost (XGB), support vector machine, or logistic regression models (Choi et al., 2021).

Men and women differed regarding risk factors with physical health more important for men. For both sexes, psychiatric disorders and medication, 48 months previous to death, were important risk factors (Gradus et al., 2020). Deep neural networks were used to stratify the risk of suicide attempt with clinical and demographic indicators. The model was able to separate the sample into four categories, from low to very high risk of attempt (Zheng et al., 2020). McCullough et al. (2023) constructed an interaction network of patients admitted to the hospital for suicide or self-harm risk, indicating the trajectories that led to referral to the emergency department.

Different types of data (sociodemographic, clinical) and populations (countries/cultures or diagnoses), can affect suicide; thus, analyzing data from various conditions can increase our knowledge about the risk factors and assist preventive measures. Table 1 presents the studies that analyzed suicide data. The studies focused on demographic or clinical data. Our work uses demographic data to profile the individuals attempting and committing suicide and we aimed to identify patterns within the city's healthcare infrastructure data to predict suicide and suicide reattempt rates in Southern Brazil.

3 METHODS

The next sections describe the methods for data collection, the analysis of the suicide rates dataset, features and the models used for the predictions, and our results. Figure 1 shows the workflow of the study. The codes for the methods and results are deposited in the following Github repository: https://github.com/anonimo-SBBD/SuicideRates.git

3.1 Data Collection and Pre-Processing

DATASUS provides curated data for SIM, SINAN, and CNES data sets. To access the data from DATASUS, we used the 'download' function of the PySUS library (Coelho et al., 2021). We downloaded the data from the SIM from 2009 to 2020. For SINAN, we used data from 2016 to 2020. The files were filtered according to the southern states and the cause of death (as suicide) for SIM. For SINAN, we filtered the observations in which the ‘self-inflicted’ and the type of violence ‘other’ columns corresponded to ‘yes’ (Pinheiro et al., 2021). We filtered the data based on the ‘violence specification’ column, using different string values that indicated suicide attempt (“suic/ts/t.s/intox/enve/enfor” referring to Portuguese spellings).

We chose the southern region as a case study because the region has the highest SR and SAR and similar socioeconomic and cultural aspects between states. We performed an initial separated analysis of the three states but observed that the analyses were similar and decided on a single, unified analysis to increase sample size.

We removed columns not relevant, redundant, with more than 30% of missing values, or with an unidentifiable meaning. The following variables remained: age, sex, race/color, marital status, schooling, city, and date of occurrence for suicide completion plus the presence of disabilities or disorders and the method used in the attempt. We excluded columns with specific values or columns with values equivalent to the outcome.

We add new variables based on the date and age: month, year, season, day of the week, and range of age (in 10 years, starting with 10 years old). Some variables with missing data were converted to "ignored value" (i.e., a value was not imputed while filling the form) because this an available option in SIM and SINAN forms. The entries with missing values in the "age" or "date" variables were excluded. We used these data sets in the exploratory analysis. In the CNES data set, we use the columns for public professional, category of the professional (43 categories), type of health establishment (33 types), and mental health specialization.

Since in SIM, SINAN, and CNES data sets variables were mostly categorical, we separated each category using the One-hot encoding method. Date and city code were used to group and add the values with information on the number of entries for each category presented and merged SIM or SINAN to CNES data sets into a single dataframe (DF). For the machine learning predictions, we calculated the SR of each city, in the respective year and month, by 100,000 inhabitants and added a column with a classification of the SR rate based on WHO recommendation: low SR ≤ 5; $5 \leq$ moderate SR ≤ 10; high SR > 10 (WHO, 2014). For SAR, we classified based on the percentage of reattempt rate: low SAR $\leq$ 33%; 33% $<$ moderate SAR $\leq$ 66%; high SAR $>$ 66% of reattempt considering all the attempts. We also add a column with the city's estimated population and demographic density for the models not to compute the rates based on the population number.

Table 1: Comparison between studies analyzing suicide data with machine learning approaches. Model results show the representative of the best model the study performed (if more than one model). AUC, area under the curve; Acc, accuracy; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; SIM, Mortality Information System; SINAN, Notification of Interpersonal and Self-Inflicted Violence; CNES, National Registry of Healthcare Establishments.

Study	Database	Time range	Analysis Strategy	Models results	Country
McCullough et al., 2023	Clinical interactions of patients in the Hospital Emergency Department	-	Classify the interaction that led to referral to the Emergency Department	Bernoulli naive Bayes classifier Acc.= 0.82	Australia
Gradus et al., 2019	Clinical and demographic data	1995 - 2015	Suicide prediction and risk factor analysis	Random Forest Men – AUC=0.80, Spec.=0.97 Women – AUC=0.88, Spec.=0.96	Denmark
Zheng et al., 2020	Clinical and demographic records	2015 - 2017	Attempt prediction and risk factor analysis	Deep neural network AUC=0.77, PPV=0.10, relative risk=59.02	United States of America
Walsh et al., 2017	Synthetic derivative - Clinical and demographic data of previous attempt	1 week to 2 years	Recurrence prediction and risk factor analysis	Random Forests AUC=0.84, precision=0.79, recall=0.95	United States of America
Roglio et al., 2020	Clinical and demographic hospital data - cocaine abuse patients	2012 - 2018	Attempt prediction and risk factor analysis	Random Forest Men - AUC=0.68, Acc.=0.66, Sens.=0.82, Spec.=0.50, PPV=0.47, NPV=0.84 Women - AUC=0.73, other=0.71	Brazil
Nordin et al., 2021	Clinical and demographic hospital data - depressive disorder patients	-	Attempt prediction and risk factor analysis	Bagging Decision tree AUC=0.87, Acc.=0.92, Sens.=0.92, Spec.=0.53, PPV=0.89, NPV=0.76	Malasya
Choi et al., 2021	KNHANES dataset - suicide attempts and ideation	2007 - 2019	Attempt prediction and risk factor analysis	XGBoost AUC=0.99, Acc.=0,99, precision=0.98, recall=0.99, F1-score=0.98	South Korea
This study	**Demographic and healthcare infrastructure**	**2009 - 2020**	**Recurrence attempt rate and suicide rate prediction**	**XGBoost Acc.=0.98, precision=0.97, Sens./Spec.=0.94**	**Brazil**

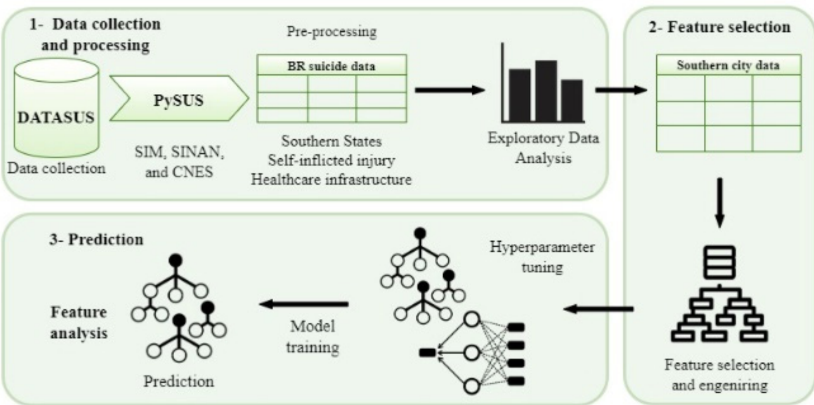

Figure 1: Workflow from data collection to outcome classification.

3.2 Machine Learning Models and Training Process

To classify SR or SAR, we used the SIM/CNES and SINAN/CNES merged DF. We divided the DF into two: train set (80%, SR - 17,007, SAR - 8,292 entries) and test set (20%); we set the same random state and divided the main outcome (SR or SAR). The models were tested with 10-fold cross-validation.

We tested three tree-based machine learning models: Decision Tree (DT), RF, and XGB. The models were chosen based on simplicity to execute, and analyze the results, do not require many data treatments, such as scaling, and perform well with categorical and numerical data. Feature importance was computed with the mean accumulation of impurity decrease within each tree.

4 RESULTS

4.1 Exploratory Data Analysis

4.1.1 Suicide Attempt

The mean age of the attempters was 30.7 years, and the absolute values decreased with the increase of age (27% from 10 to 19 decreasing to 4% from individuals over 60 years, approximately). For schooling, middle and high school education (26.5% and 32.0%) were the most represented (25.30% of the data were missing). More than 1/3 of the attempters had a disorder or disability (37.4%), and the most common were psychiatric (23.6%) and mood (12.9%) disorders. The main methods chosen for the attempt were intoxication (67.8%), cutting/stabbing (9.5%), and hanging (7.8%). Moreover, 18% of males and 8% of females were intoxicated with alcohol during the attempt. Interestingly, 38.5% of the individuals had already attempted suicide, whereas 41.7% were first-time attempters (19.8% of data was missing).

4.1.2 Suicide Completion

The three states had a cumulative rate of 6.02 (PR), 8.31 (SC), and 10.32 (RS) per 100,000 inhabitants. From 2009 to 2020 SR increased: PR 41.00%, SC 51%, and RS 27.67%. Males died 3.85-fold more than females and the mean age was 45 years old; most deaths occurred between ages 20 to 59 years, and most were White (87.38%). Schooling was divided by years studied: from four to seven years corresponded to 26.76%, and from eight to eleven years corresponded to 24.20% (24.42% of the data was missing). Most deaths occurred in the person's residence (66.45%) and were distributed similarly throughout the year.

4.2 Machine Learning Classification

Generalization of the models presented high accuracy, approximately 0.95. Table 2 presents the best parameters of the models, the evaluation metrics, and the variables most important to the models to predict SR and SAR. Low, moderate, and high classes were imbalanced: SR - 58,67%, 17,87%, and 23,46%; SAR - 15,67%, 76,60%, and 7,73%.

5 DISCUSSION

Combat of suicide deaths is a challenge to countries around the globe. Low- and middle-income countries are at particular risk, with almost 80% of global deaths. In some of those countries, death rate has been increasing in the past years, such as Brazil (WHO, 2021). Suicide deaths are preventable, and the World Health Organization general guidelines suggest restricting means (firearm and medications control) and increasing awareness for mental health (WHO, 2014). Moreover, understanding regional specificity helps developing more effective public policies.

From an economic perspective, suicide deaths may cost up to 3 million dollars to the government in direct and indirect costs per death (Kinchin & Doran, 2018; Shepard et al., 2016). Moreover, in Brazil, the average cost of a suicide attempt is approximately US$ 7,000 (Sgobin et al., 2015). Many costs occur after the attempt and include years lost due to trauma, post-treatment, and impact on work and life of family and acquaintances (Kinchin & Doran, 2018).

Observing that many attempts occur between ages 10 and 20, we recommend the development of more programs of mental health and suicide prevention presented during school years. Indeed, school-based preventive programs targeted at children and adolescents can be cost-effective (Ahern et al., 2018).

Our analysis of the profile of individuals committing suicide corroborates previous findings, suggesting that preventive programs must target specific populations. Suicide attempts are the main risk factors, and many suicide completers do not die on the first attempt. Non-pharmacological approaches after the attempt, such as active contact and follow-up and cognitive behavioral therapy, may be used along with prevention programs to increase cost-effectiveness and decrease SR and SAR (Martínez-Alés et al., 2021; Ross et al., 2021).

Table 2: Comparison between machine learning models results and most important features used by the models to classify reattempt and suicide rates in Southern Brazil. Important features specific to reattempt (*) or completion ([#]) rates (features are in no particular order).

Model	Parameters	Results	Important Features
Decision Tree	ccp alpha = 0.005	Accuracy = 0.95 Precision = 0.94 Sensitivity = 0.95 Specificity = 0.91	Number of physicians Demographic density Number of nurses[#] Number of psychosocial centers[#] Number of basic units*
Random Forest	max_depth = 15	Accuracy = 0.96 Precision = 0.98 Sensitivity = 0.93 Specificity = 0.94	Demographic density Number of physicians Number of nurses Number of clinics[#] Public professionals[#] Number of dentists*
XGBoost	learning rate = 0.1 n estimators = 3000 subsample = 0.5 colsample bytree= 0.5	Accuracy = 0.98 Precision = 0.97 Sensitivity = 0.94 Specificity = 0.94	Number of physicians[#] Number of nurses[#] Number of clinics[#] Public professionals[#] Number of neonatal units* Number of short stay hospital*

Infrastructure of the environment in which the individual is inserted is also important. For example, green areas within the pathways used by the population and spatial inaccessibility to psychiatrists or psychotherapists was shown to interfere in suicide rates in the region (Shen & Lung, 2018; Tadmon & Bearman, 2023). In addition, since a substantial proportion of suicide completers use the healthcare system the year before their death (Ahmedani et al., 2014), we sought to understand the impact that the city's healthcare infrastructure has on SR and SAR in a case study of the Brazilian Southern region.

The Brazilian health care system is decentralized, with more focus on primary care and outpatient specialized services than in hospitalization. It has private and public domains, and more than 70% of the population relies on public services. The services are distributed based on the demographical density and economic indicators. Thus, smaller cities have primary but may lack specialized services. Larger cities usually work as an epicenter, providing specialized health services (Paim et al., 2011).

We used three models to classify SR and SAR: DT, RF, and XGB. The models had similar results, with predictive values of approximately 95%. As the data are unbalanced, other metrics are also important to be analyzed, and as presented in Table 2, F1-score averaged approximately 0.94. As expected, XGB, having a correction of errors during training, had a better performance, although all models showed satisfactory results in classifying SR and SAR.

Understanding which factors are decisive may help focus on specific targets to change or improve. Our data shows that Psychosocial Care Centers (CAPS), a type of healthcare facility focused on treatment and improvement of psychiatric conditions in the population, are one of the main features. Additionally, since the majority of the Brazilian population relies on the National Health System, increasing the system with more clinicians and nurses, especially in cities with high SR and SAR, may help decrease the rates. Interestingly, some models also considered the dental professionals, showing that general well-being is important.

CAPS are divided into six classes and for people suffering psychological conditions, and focused on adults, or children and adolescents, or substance abusers, and may be open 24 hours per day (with hospital beds available). The simplest CAPS require 15,000 inhabitants, and the most completes up to 150,000 (Ministry of Health, 2023). Brazil has 5 565 cities, from which only 677 have more than 50 000 inhabitants. Thus, the majority of the cities only have CAPS I available. Considering that psychiatric conditions and substance abuse are some of the main factors of suicide, and that children and adolescents showed high rates of suicide attempt, rates would probably decrease with more centers in more cities.

A limitation of our study is that in each question of the SIM/SINAN forms, the evaluator may mark ‘ignored’. The correct filing of the form with answers that have semantic values can increase the specificity of the analysis, and if the forms were more similar, we could group the forms and analyze simultaneously. Also, for the SINAN form, the presence of an anonymous identifier would help distinguish the profile of the individual after the first attempt. In addition, we were not able to distinguish the different specializations of the healthcare professionals, such as psychiatrists from general physicians, or the teams present at the different facilities, which could improve the models' performance and point to more direct improvements. Lastly, underreporting plays a crucial role, especially in smaller regions, where suicide is more stigmatized.

6 CONCLUSIONS AND PERSPECTIVES

In this study, we focused on extracting and interpreting patterns from suicide completion and reattempt rates in Brazil. This is the first study using the Brazilian healthcare infrastructure to classify rates. Our models achieved a high predictive performance of up to 97% accuracy in predicting suicide death or reattempt. Compared to other studies, we focused on the environment in which the population is inserted, trying to use the model in a descriptive manner, to identify and better understand the patterns arising from models' application. This approach showed the importance of Psychosocial Care Centers and the number of physicians and nurses in impacting deaths and suicide reattempts. Future studies could use a similar approach with other city infrastructures, such as those related to industrialization, employment, education, and sanitation to decrease these preventable deaths.

ACKNOWLEDGEMENTS

We would like to thank CAPES and FAPESC for the financial support.

REFERENCES

Ahern, S., Burke, L.-A., McElroy, B., Corcoran, P., McMahon, E. M., Keeley, H., Carli, V., Wasserman, C., Hoven, C. W., Sarchiapone, M., Apter, A., Balazs, J., Banzer, R., Bobes, J., Brunner, R., Cosman, D., Haring, C., Kaess, M., Kahn, J.-P., … Wasserman, D. (2018). A cost-effectiveness analysis of school-based suicide prevention programmes. *European Child & Adolescent Psychiatry*, *27*(10), 1295–1304. https://doi.org/10.1007/s00787-018-1120-5

Ahmedani, B. K., Simon, G. E., Stewart, C., Beck, A., Waitzfelder, B. E., Rossom, R., Lynch, F., Owen-Smith, A., Hunkeler, E. M., Whiteside, U., Operskalski, B. H., Coffey, M. J., & Solberg, L. I. (2014). Health care contacts in the year before suicide death. *Journal of General Internal Medicine*, *29*(6), 870–877. https://doi.org/10.1007/s11606-014-2767-3

Choi, J., Cho, S., Ko, I., & Han, S. (2021). Identification of risk factors for suicidal ideation and attempt based on machine learning algorithms: a longitudinal survey in Korea (2007–2019). *International Journal of Environmental Research and Public Health*, *18*(23), 12772. https://doi.org/10.3390/ijerph182312772

Coelho, F. C., Baron, B. C., de Castro Fonseca, G. M., Reck, P., & Palumbo, D. (2021). *AlertaDengue/PySUS: Vaccine*. Zenodo. https://doi.org/10.5281/zenodo.4883502

Gao, K., Wu, R., Wang, Z., Ren, M., Kemp, D. E., Chan, P. K., Conroy, C. M., Serrano, M. B., Ganocy, S. J., & Calabrese, J. R. (2015). Disagreement between self-reported and clinician-ascertained suicidal ideation and its correlation with depression and anxiety severity in patients with major depressive disorder or bipolar disorder. *Journal of Psychiatric Research*, *60*, 117–124. https://doi.org/10.1016/j.jpsychires.2014.09.011

Gradus, J. L., Rosellini, A. J., Horváth-Puhó, E., Street, A. E., Galatzer-Levy, I., Jiang, T., Lash, T. L., & Sørensen, H. T. (2020). Prediction of sex-specific suicide risk using machine learning and single-payer health care registry data from Denmark. *JAMA Psychiatry*, *77*(1), 25. https://doi.org/10.1001/jamapsychiatry.2019.2905

Jaen-Varas, D., Mari, J. J., Asevedo, E., Borschmann, R., Diniz, E., Ziebold, C., & Gadelha, A. (2019). The association between adolescent suicide rates and socioeconomic indicators in Brazil: a 10-year retrospective ecological study. *Brazilian Journal of Psychiatry*, *41*(5), 389–395. https://doi.org/10.1590/1516-4446-2018-0223

Johnson, T., Kanjo, E., & Woodward, K. (2023). DigitalExposome: quantifying impact of urban environment on wellbeing using sensor fusion and deep learning. *Computational Urban Science*, *3*(1), 14. https://doi.org/10.1007/s43762-023-00088-9

Kinchin, I., & Doran, C. (2018). The cost of youth suicide in Australia. *International Journal of Environmental Research and Public Health*, *15*(4), 672. https://doi.org/10.3390/ijerph15040672

Martínez-Alés, G., Cruz Rodríguez, J. B., Lázaro, P., Domingo-Relloso, A., Barrigón, M. L., Angora, R., Rodríguez-Vega, B., Jiménez-Sola, E., Sánchez-Castro, P., Román-Mazuecos, E., Villoria, L., Ortega, A. J., Navío, M., Stanley, B., Rosenheck, R., Baca-García, E., & Bravo-Ortiz, M. F. (2021). Cost-effectiveness of a contact intervention and a psychotherapeutic program for post-discharge suicide prevention. *The Canadian Journal of Psychiatry*, *66*(8), 737–746. https://doi.org/10.1177/0706743720980135

Martini, M., da Fonseca, R. C., de Sousa, M. H., de Azambuja Farias, C., Cardoso, T. de A., Kunz, M., Longaray, V. K., & Magalhães, P. V. da S. (2019). Age and sex trends for suicide in Brazil between 2000 and 2016. *Social Psychiatry and Psychiatric Epidemiology*, *54*(7), 857–860. https://doi.org/10.1007/s00127-019-01689-8

Mccullough, M. H., Small, M., Jayawardena, B., & Hood, S. (2023). Mapping patient interactions in psychiatric presentations to a tertiary emergency department. *MedRxiv Preprint*, 1–23. https://doi.org/10.1101/2023.05.17.23290083

Ministry of Health. (2021). Mortalidade por suicídio e notificações de lesões autoprovocadas no Brasil. *Boletim Epidemiológico Secretaria de Vigilância Em Saúde − Ministério Da Saúde*, *52*(33), 1–10. https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2021/boletim_epidemiologico_svs_33_final.pdf

Ministry of Health. (2023). *Centros de Atenção Psicosocial*. https://www.gov.br/saude/pt-br/composicao/saes/desme/raps/caps

Nordin, N., Zainol, Z., Mohd Noor, M. H., & Lai Fong, C. (2021). A comparative study of machine learning techniques for suicide attempts predictive model. *Health Informatics Journal*, *27*(1), 146045822198939. https://doi.org/10.1177/1460458221989395

Paim, J., Travassos, C., Almeida, C., Bahia, L., & Macinko, J. (2011). The Brazilian health system: history, advances, and challenges. *The Lancet*, *377*(9779), 1778–1797. https://doi.org/10.1016/S0140-6736(11)60054-8

Palma, D. C. de A., Oliveira, B. F. A. de, & Ignotti, E. (2021). Suicide rates between men and women in Brazil, 2000-2017. *Cadernos de Saúde Pública*, *37*(12), e00281020. https://doi.org/10.1590/0102-311x00281020

Pinheiro, T. de P., Warmling, D., & Coelho, E. B. S. (2021). Caracterização das tentativas de suicídio e automutilações por adolescentes e adultos notificadas em Santa Catarina, 2014-2018. *Epidemiologia e Serviços de Saúde*, *30*(4), e2021337. https://doi.org/10.1590/s1679-49742021000400026

Roglio, V. S., Borges, E. N., Rabelo-da-Ponte, F. D., Ornell, F., Scherer, J. N., Schuch, J. B., Passos, I. C., Sanvicente-Vieira, B., Grassi-Oliveira, R., von Diemen, L., Pechansky, F., & Kessler, F. H. P. (2020). Prediction of attempted suicide in men and women with crack-cocaine use disorder in Brazil. *PLOS ONE*, *15*(5), e0232242. https://doi.org/10.1371/journal.pone.0232242

Ross, E. L., Zuromski, K. L., Reis, B. Y., Nock, M. K., Kessler, R. C., & Smoller, J. W. (2021). Accuracy requirements for cost-effective suicide risk prediction among primary care patients in the US. *JAMA Psychiatry*, *78*(6), 642. https://doi.org/10.1001/jamapsychiatry.2021.0089

Sgobin, S. M. T., Traballi, A. L. M., Botega, N. J., & Coelho, O. R. (2015). Direct and indirect cost of attempted suicide in a general hospital: cost-of-illness study. *Sao Paulo Medical Journal*, *133*(3), 218–226. https://doi.org/10.1590/1516-3180.2014.8491808

Shen, Y.-S., & Lung, S.-C. C. (2018). Identifying critical green structure characteristics for reducing the suicide rate. *Urban Forestry & Urban Greening*, *34*, 147–153. https://doi.org/10.1016/j.ufug.2018.06.005

Shepard, D. S., Gurewich, D., Lwin, A. K., Reed, G. A., & Silverman, M. M. (2016). Suicide and suicidal attempts in the United States: costs and policy implications. *Suicide and Life-Threatening Behavior*, *46*(3), 352–362. https://doi.org/10.1111/sltb.12225

Tadmon, D., & Bearman, P. S. (2023). Differential spatial-social accessibility to mental health care and suicide. *Proceedings of the National Academy of Sciences*, *120*(19). https://doi.org/10.1073/pnas.2301304120

Turecki, G., Brent, D. A., Gunnell, D., O'Connor, R. C., Oquendo, M. A., Pirkis, J., & Stanley, B. H. (2019). Suicide and suicide risk. *Nature Reviews Disease Primers*, *5*(1), 74. https://doi.org/10.1038/s41572-019-0121-0

van Heeringen, K., & Mann, J. J. (2014). The neurobiology of suicide. *The Lancet Psychiatry*, *1*(1), 63–72. https://doi.org/10.1016/S2215-0366(14)70220-2

Walsh, C. G., Ribeiro, J. D., & Franklin, J. C. (2017). Predicting Risk of Suicide Attempts Over Time Through Machine Learning. *Clinical Psychological Science*, *5*(3), 457–469. https://doi.org/10.1177/2167702617691560

WHO (2014). *Preventing suicide: A global imperative*. World Health Organization.

WHO (2021). *Suicide worldwide in 2019: global health estimates*.

Zheng, L., Wang, O., Hao, S., Ye, C., Liu, M., Xia, M., Sabo, A. N., Markovic, L., Stearns, F., Kanov, L., Sylvester, K. G., Widen, E., McElhinney, D. B., Zhang, W., Liao, J., & Ling, X. B. (2020). Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health records. *Translational Psychiatry*, *10*(1), 72. https://doi.org/10.1038/s41398-020-0684-2