Using machine learning to predict suicide in the 30 days after discharge from psychiatric hospital in Denmark

Tammy Jiang, Anthony J. Rosellini, Erzsébet Horváth-Puhó, Brian Shiner, Amy E. Street, Timothy L. Lash, Henrik T. Sørensen and Jaimie L. Gradus

Background

Suicide risk is high in the 30 days after discharge from psychiatric hospital, but knowledge of the profiles of high-risk patients remains limited.

Aims

To examine sex-specific risk profiles for suicide in the 30 days after discharge from psychiatric hospital, using machine learning and Danish registry data.

Method

We conducted a case–cohort study capturing all suicide cases occurring in the 30 days after psychiatric hospital discharge in Denmark from 1 January 1995 to 31 December 2015 ($n = 1205$). The comparison subcohort was a 5% random sample of all persons born or residing in Denmark on 1 January 1995, and who had a first psychiatric hospital admission between 1995 and 2015 ($n = 24\,559$). Predictors included diagnoses, surgeries, prescribed medications and demographic information. The outcome was suicide death recorded in the Danish Cause of Death Registry.

Results

For men, prescriptions for anxiolytics and drugs used in addictive disorders interacted with other characteristics in the risk profiles (e.g. alcohol-related disorders, hypnotics and sedatives) that led to higher risk of postdischarge suicide. In women, there was interaction between recurrent major depression and other characteristics (e.g. poisoning, low income) that led to increased risk of suicide. Random forests identified important suicide predictors: alcohol-related disorders and nicotine dependence in men and poisoning in women.

Conclusions

Our findings suggest that accurate prediction of suicide during the high-risk period immediately after psychiatric hospital discharge may require a complex evaluation of multiple factors for men and women.

Keywords

Suicide; machine learning; psychiatric hospital; postdischarge suicide; suicide prediction.

Suicide is a severe public health problem, with approximately 800 000 people dying by suicide each year globally.[1] The World Health Organization estimates that the global suicide rate is 10.6 per 100 000 person-years.[2] The immediate time after hospital discharge for a mental health condition is a critical period during which suicide risk is high.[3,4] The rate of suicide in the first month after psychiatric hospital discharge is reported to be 2060 per 100 000 person-years.[4] The suicide rate among persons in the first month after discharge from a psychiatric hospital is over 190 times the global suicide rate.[2,4] The immediate period after discharge represents a unique opportunity for prevention to reduce suicide deaths among this vulnerable population.

Suicide prediction

Our understanding of the factors that predict heightened risk of suicide after psychiatric hospital discharge remains limited. Previous studies documented increased postdischarge suicide risk among persons who have a history of self-harm, affective disorders, relationship problems, job loss and are living alone.[5,6] Although many risk factors for suicide have been documented, a 2017 meta-analysis of the past 50 years of research on suicidal thoughts and behaviours found that we still cannot accurately predict who will die by suicide.[7] Most studies examined a small number of risk factors, but accurate suicide prediction likely requires examination of hundreds of risk factors and their interactions. However, conventional parametric statistical techniques are not designed to examine large, highly correlated sets of predictors, or to detect interactions among predictors without *a priori* specification. Machine learning methods can detect complex patterns and return useful algorithms for predicting suicide, thus enabling the development of novel suicide risk profiles that include constellations of predictors. Furthermore, the development of prediction models in high-risk groups, such as persons who have been admitted to psychiatric hospital, is likely to improve relevance and acceptability to clinicians who work with this population.[8]

Kessler and colleagues developed algorithms to estimate suicide risk among USA army soldiers after a stay in hospital, and in the USA Veterans Health Administration system.[9,10] However, despite this contribution, several gaps remain. First, few studies have focused on predicting suicide in the relatively short 30-day window of interest to clinicians.[7] One reason for this is the lack of an adequate sample size, given that suicide is rare. Second, findings from USA army members and veterans may not be generalisable to the broader population of people who are admitted to a psychiatric hospital. Third, little is known if risk profiles of suicide after a psychiatric hospital admission differ among men and women. Previous research predicting postdischarge suicides by using machine learning did not examine gender differences, which is likely because of the smaller proportion of women in the military than men. However, there may be different risk profiles of suicide in men and women, since there are well-established gender differences[11] in suicide risk and risk factor distributions.[12]

Aims of study

The purpose of this study was to identify novel interactions and variables that predict suicide in the 30 days after discharge from a

psychiatric hospital, in men and women. We leveraged Danish registry data captured over a 20-year period, encapsulating all diagnoses, surgeries, medication prescriptions and demographic/social register information. We used machine learning methods, including classification trees and random forests, to achieve our dual goals of characterising interactions between variables and elucidating novel predictors of suicide after discharge from a psychiatric hospital.

Method

Study sample

The source population was all individuals born or residing in Denmark as of 1 January 1995. The start of the study period coincides with the switch from the ICD-8 to the ICD-10 in 1994, and the start of reporting of all hospital out-patient clinic visits to the Danish National Patient Registry, covering all Danish hospitals in 1995.[13] We implemented a case–cohort design because it is an efficient approach for studying rare outcomes.[14] We did not match cases and subcohort members, to allow for maximum variability in the predictors in the analysis. Cases were all individuals who died by suicide and had been hospitalised for a psychiatric disorder within 30 days before their death between 1 January 1995 and 31 December 2015 in Denmark ($n = 1205$). The comparison subcohort was a 5% random sample of individuals in Denmark on 1 January 1995, and who had a first hospital admission for a psychiatric disorder between 1 January 1995 and 31 December 2015 in Denmark ($n = 24\,559$). We included persons who were hospitalised for the following disorders, as recorded by two-digit ICD-10 codes from the Danish Psychiatric Central Research Register[15] and Danish National Patient Registry:[13] mental disorders due to known physiological conditions (F01–F09); substance use disorders (F10–F19); schizophrenia (F20–F29); mood disorders (F30–F39); anxiety, dissociative, stress-related and somatoform disorders (F40–F48); behavioural syndromes associated with physiological disturbances and physical factors (F50–F59); personality and behaviour disorders (F60–F69); behavioural and emotional disorders with onset usually occurring in childhood and adolescence (F90–F98) and unspecified mental disorders (F99). We used central personal registry numbers that are unique, individual-level identifiers assigned to all Danish residents to link data across Danish administrative and medical registries.[16] We used the Danish Civil Registration System to randomly select comparison subcohort members.[17]

Outcome

We obtained suicide cases by ICD-10 codes X60–X84, as recorded in the Danish Cause of Death Registry.[18] This registry records data on age of death, manner of death (e.g. natural, suicide), place of death and autopsy results.[18] A validation study confirmed suicide as the cause of death for 92% of the deaths recorded as suicides.[19]

Predictors

We examined the following variables as predictors in the machine learning models: age, marital status, immigration status, citizenship, family suicide history (parent or spouse), employment, income, mental disorders, physical health disorders, surgeries, prescription drugs and psychotherapy. We used the Danish Civil Registration System[17] to obtain data on age, marital status, immigration status, generation of citizenship and family suicide history. We used the Integrated Database for Labor Market Research[20] and Income Statistics Register to obtain baseline data on employment and income.[21] We ascertained psychiatric disorder diagnoses by using two-digit ICD-10 codes from the Danish Psychiatric Central

Research Register[15] and Danish National Patient Registry.[13] We also used the Danish National Patient Registry to obtain physical health diagnoses, as recorded by second-level ICD-10 groupings. Surgery procedure codes from the Danish National Patient Registry were examined according to the body system. We obtained data on prescription drugs from the Danish National Prescription Registry.[22] Prescription drugs for this study were coded according to level three Anatomical Therapeutic Classification codes. All codes analysed are in the Supplementary Appendix, available at https://doi.org/10.1192/bjp.2021.19.

Statistical analyses

The time-varying predictors for both cases and the comparison subcohort members were defined 30 days after discharge from psychiatric hospital. Cases were persons who died by suicide in the 30 days after discharge. For cases, we dummy-coded variables to create time-varying predictors with intervals of 0–6, 0–12, 0–24 and 0–48 months before the date of suicide. To compute the prevalence of each predictor in the person-time that gave rise to cases for the comparison subcohort, we used the date 30 days after discharge to calculate the prevalence of predictors in the 0–6, 0–12, 0–24 and 0–48 months before that date. For example, for a member of the comparison subcohort who was discharged from a psychiatric hospital on 1 December 2010, we calculated the prevalence of predictors 0–6, 0–12, 0–24 and 0–48 months before 31 December 2010 (i.e. 31 December 2010 is 30 days after the discharge date). Time intervals were chosen to be consistent with intervals used in previous research that used machine learning for suicide prediction.[9,23–25] Age, immigration status, employment and income at baseline were not coded as time-varying predictors. Predictors from all periods were evaluated simultaneously in the models.

We conducted data reduction to avoid overfitting. Overfitting arises when a model finds patterns that are unique to a specific data-set, but are not generalisable to external samples.[26] We performed data reduction for men and women separately, by removing rare predictors that had fewer than ten observations in any cell of a 2×2 contingency table of the predictor and suicide,[9] and removing predictors with negligible associations with suicide (unadjusted odds ratio between 0.9 and 1.1). We removed emergency department diagnoses because of their low positive predictive values.[27,28] The initial analytic data-set contained 2563 predictors. After data reduction, the final number of included predictors was 509 for men and 422 for women. The Supplementary Appendix provides the considered and retained predictors.

Given our interests in identifying novel predictors and interactions that accurately predict suicide after discharge from psychiatric hospital, we used recursive partitioning methods that automatise detection of interactions and provide metrics of predictor importance.[29] First, we used classification trees, which are a nonparametric method that builds a decision tree based on predictors and their combinations that result in the highest probability of differentiating cases from non-cases. Classification trees can elucidate interactions among large sets of predictors without *a priori* specification, and provide a visual depiction of risk factor constellations that predict suicide. However, classification trees are more vulnerable to overfitting than random forests.[30] To decrease the risk of overfitting, we used ten-fold cross-validation of classification trees. To increase visual interpretability, we set the maximum tree depth and minimum number of observations in any node to five. To address class imbalance, we used equal priors.[31] The risk of suicide was calculated for each identified combination of predictors. We used the R package *rpart*[32] to implement classification trees.

Second, we implemented random forests, which are a recursive partitioning method that comprises a set of decision trees generated

with bootstrapped samples of the data. Each forest was built with 1000 trees, and a minimum of ten observations were needed to attempt a split. The number of variables sampled as split candidates at each node were 23 for men and 21 for women (i.e. square root of the total number of predictors for men and women; R package *randomForest* default). To address class imbalance, each individual tree was built using all suicide observations and an equally sized number of randomly selected non-suicide observations, using the sampsize tuning parameter. We used two-fold cross-validation to generate individual-level random forests predicted values. We calculated the mean decrease in accuracy of each variable, which represents the reduction in accuracy if a predictor were permuted.[29] The larger the mean decrease in accuracy of a predictor, the more important it is for accurate prediction of suicide. We used the R package *randomForest*.[33] Although random forests provide metrics of predictor importance, they do not provide a visualisation of interactions between variables as classification trees do. Thus, we leveraged the strengths of both classification trees and random forests to best serve our dual interests in identifying novel predictors and interactions that predict suicide after discharge from psychiatric hospital.

We evaluated prediction accuracy by receiver operating characteristics curve analysis conducted in 1000 bootstrap replicates, to estimate the area under the curve (AUC) and its 95% confidence interval. We also examined the sensitivity and specificity of the classification trees and random forests in detecting suicide. The analyses were conducted separately for men and women. We conducted gender-stratified analyses instead of including gender as a predictor in an analysis of the entire sample, because the latter approach would not reveal gender differences in random forest variable importance, and in classification trees, it would only display separate patterns of risk in men and women at the point that gender is chosen as a splitting variable, but not earlier in the tree. Analyses were conducted in SAS, version 9.4 (SAS Institute) for Windows[34] and R, version 3.5.2 in Windows (R Core Team, Vienna Austria, https://www.R-project.org/).[35] This study was determined to be exempt from review by the Boston University Institutional Review Board, approved by the Danish Data Protection Agency (record number 2015-57-0002). Use of data from Danish registries does not require informed consent according to Danish law.

Results

Table 1 shows the descriptive characteristics of the study sample. Among men, the mean ages were similar in the suicide cases and the comparison subcohort. Among women, suicide cases were, on average, younger than the comparison subcohort members (mean 52 [s.d. 16] *v.* 57 [s.d. 24]). Across sexes, suicide cases and subcohort members had similar proportions of immigrant status and persons in a married or registered partnership. Suicide cases were less likely to be in the lowest income quartile compared with subcohort members.

Classification trees

The highest risk of suicide was among men who were not prescribed antidepressants in the 48 months preceding admission to psychiatric hospital, were diagnosed with alcohol-related disorders in the preceding 6 months, were prescribed hypnotics and sedatives, had a poisoning diagnosis (poisoning by, adverse effect of and underdosing of drugs, medications and biological substances) in the preceding 48 months and were prescribed anxiolytics in the preceding 12 months (*n* = 20; suicide risk 91%). However, men with the same risk profile (i.e. same combination of variables) who were not prescribed anxiolytics in the preceding 12 months (*n* = 14) had a 0% risk of suicide. This result suggests that in this subgroup, there is an important interaction between this combination of variables and anxiolytics prescriptions. Another stark contrast in suicide risk is among men who were prescribed antidepressants in the prior 48 months and hypnotics and sedatives in the preceding 24 months, were diagnosed with cerebrovascular diseases in the preceding 12 months, and not diagnosed with poisoning in the preceding 48 months. Among men with this risk profile, those who were prescribed drugs used in addictive disorders in the preceding 6 months (*n* = 14) had a 72% risk of suicide. However, men with the same risk profile, but who were not prescribed drugs used in addictive disorders in the preceding 6 months (*n* = 112) had a 0% risk of suicide. Figure 1 shows other combinations of predictors that conferred elevated risk of suicide in men (AUC = 0.80, 95% CI 0.78–0.81).

Among women who were hospitalised for psychiatric disorders, the highest risk of suicide was in women who were prescribed antipsychotics and had a poisoning diagnosis in the preceding 48 months (*n* = 313, suicide risk 93%). The second highest risk group was women who were prescribed antipsychotics and anxiolytics and diagnosed with a specific personality disorder, but were not diagnosed with poisoning, in the preceding 48 months (*n* = 172, risk 91%). Another interesting combination of variables was among women who had a poisoning diagnosis in the preceding 48 months, were not prescribed antipsychotics in the preceding 48 months or anxiolytics in the preceding 6 months, and were not in the highest income quartile. Among women with this risk profile, those who had a recurrent major depressive disorder diagnosis in the preceding 6 months (*n* = 22) had an 86% risk of postdischarge suicide, whereas those who did not have recurrent depression (*n* = 254) had a 10% risk of postdischarge suicide. Figure 2 shows other combinations of predictors and their associated suicide risks among women (AUC = 0.83, 95% CI 0.80–0.86).

Table 1 Characteristics of the suicide cases and the subcohort, Denmark, 1 January 1995

	Men		Women	
Variable	Suicide cases (*n* = 762)	Comparison subcohort (*n* = 11 966)	Suicide cases (*n* = 443)	Comparison subcohort (*n* = 12 593)
Age, mean (s.d.), years	51 (16)	51 (21)	52 (16)	57 (24)
Marital status, *n* (%)				
Married or registered partnership	249 (33%)	4024 (34%)	129 (29%)	3578 (28%)
Divorced/widowed/never married/single	513 (67%)	7901 (66%)	314 (71%)	8979 (71%)
Immigrant, *n* (%)	30 (4%)	498 (4%)	16 (4%)	485 (4%)
Income quartile, *n* (%)				
<1	131 (17%)	3245 (27%)	95 (21%)	3223 (26%)
1 to <2	198 (26%)	2986 (25%)	88 (20%)	3193 (25%)
2 to <3	210 (28%)	2836 (24%)	111 (25%)	3095 (25%)
>3	223 (29%)	2899 (24%)	149 (34%)	3082 (25%)
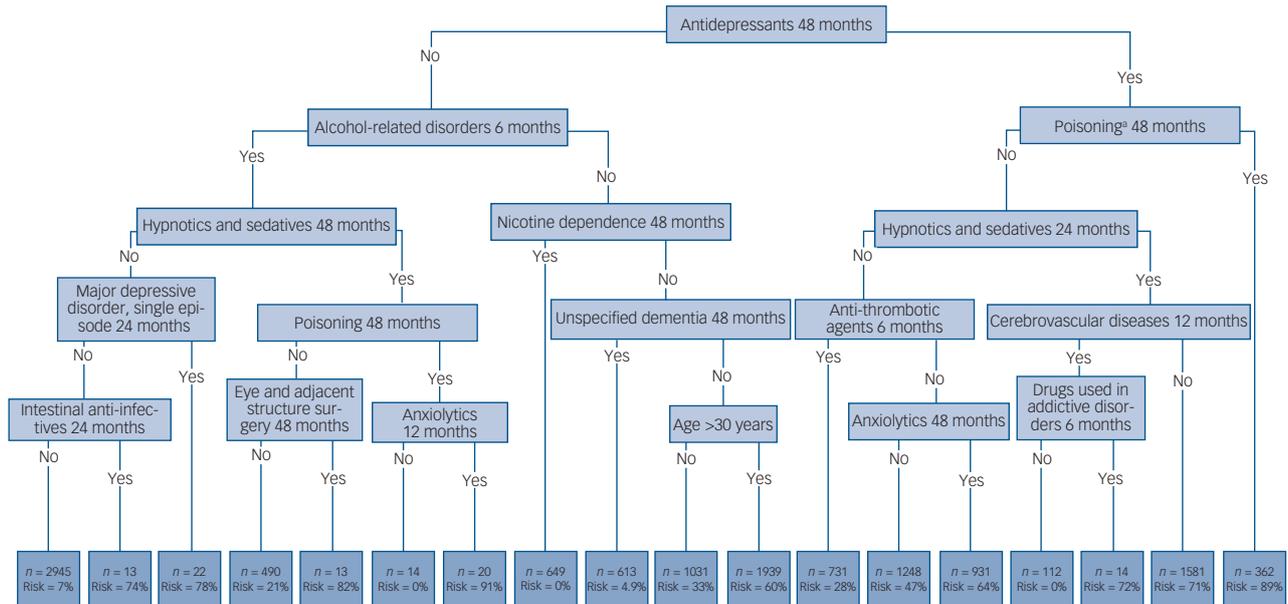
Fig. 1 Classification tree depicting suicide predictors among male patients hospitalised for psychiatric disorders in Denmark, 1995–2015. Each shaded rectangular bin at the bottom (terminal node) represents the group of people with the characteristic profile in the branches above. Within the rectangular bins, *n* indicates the number of people who had the characteristic profile, and risk indicates the proportion of people in that bin who died by suicide. ªPoisoning by, adverse effect of and underdosing of drugs, medications and biological substances.
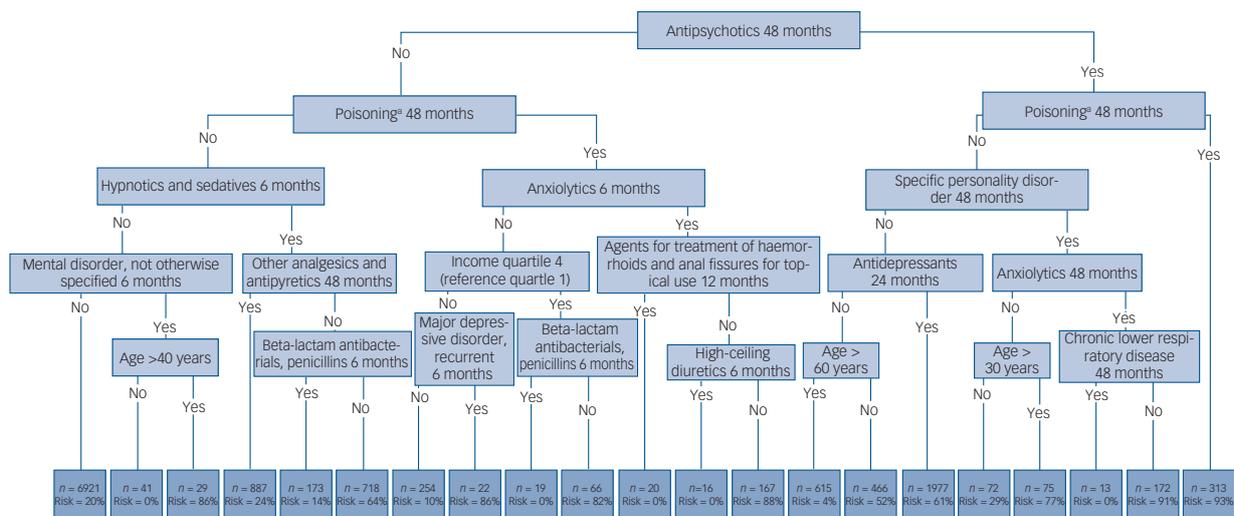
Fig. 2 Classification tree depicting suicide predictors among female patients hospitalised for psychiatric disorders in Denmark, 1995–2015. Each shaded rectangular bin at the bottom (terminal node) represents the group of people with the characteristic profile in the branches above. Within the rectangular bins, *n* indicates the number of people who had the characteristic profile, and risk indicates the proportion of people in that bin who died by suicide. ªPoisoning by, adverse effect of and underdosing of drugs, medications and biological substances.

Random forests

Among men who were hospitalised for a psychiatric disorder, 64–67% (fold one to fold two) of the predictors had a mean decrease in accuracy above zero (mean 3.8, s.d. 3.4). Fifteen predictors were among the top 30 most important predictors in both folds (Fig. 3). The most important variables for predicting suicide included age >30 years, alcohol-related disorders, nicotine dependence, major depressive disorder, unspecified dementia, antidepressants, and reaction to severe stress and adjustment disorders. The AUC for the random forest across folds was 0.82 (95% CI 0.80–0.83).

Among women who were hospitalised for a psychiatric disorder, 62–64% (fold one to fold two) of the predictors had a mean

decrease in accuracy above zero (mean 2.9, s.d. 2.2). Twelve predictors were among the top 30 most important predictors in both folds (Fig. 4). The most important predictors of suicide among female patients hospitalised for psychiatric disorders included progestogens and oestrogens in combinations, poisoning, age >60 years, receiving a state pension, antipsychotics, bipolar disorder and major depressive disorder. The AUC for the random forest across folds was 0.85 (95% CI 0.83–0.87).

Operating characteristics of high-risk thresholds

The predicted probabilities from cross-validated random forests were rank-ordered, and operating characteristics were computed
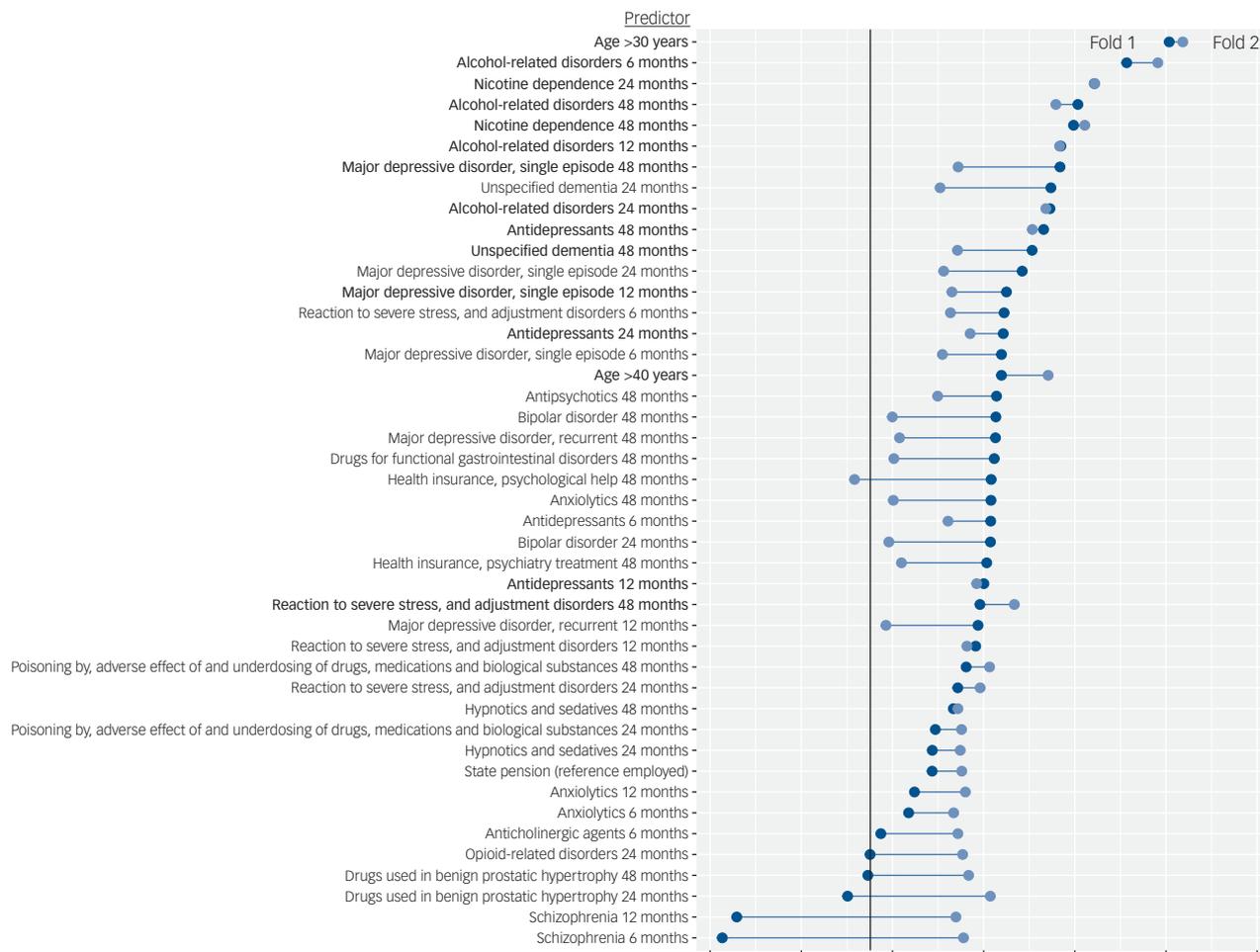
Fig. 3 Variable importance of suicide predictors among male patients hospitalised for psychiatric disorders in Denmark from split sample cross-validation, 1995–2015. The dark blue dots represent the mean decrease in accuracy (MDA) value in fold one, and the light blue dots represent the MDA value in fold two. The vertical line represents the average of the MDA values of all predictors with nonzero MDA values in folds one and two (3.8). The predictors shown in bold were in the top 30 predictors in folds 1 and 2 for men.

among individuals in the top quintile of the predicted risk distribution. Men in the top 5%, 10% and 20% of predicted risk accounted for 23%, 38% and 59% of all suicide cases among men, respectively (specificity of 96%, 92% and 82%). Women in the top 5%, 10% and 20% of predicted risk accounted for 38%, 52% and 71% of all suicide cases among women, respectively (specificity of 96%, 92% and 82%).

Discussion

This study demonstrates the complexity of the patient profiles that carry a high risk of suicide in the 30 days after discharge from psychiatric hospital, among men and women in a full population sample. Our findings build upon previous work demonstrating that accurate suicide prediction requires a complex combination of a large number of psychological, physical and social factors, many of which are time-varying.

For men, there were important interactions between specific risk profiles by anxiolytics and drugs used in addictive disorders. In women, we found an interaction between poisoning and recurrent major depressive disorder that elevates the risk of suicide after discharge from a psychiatric hospital. Poisonings may be capturing some non-fatal suicide attempts in these data, so these results may suggest that women with recurrent major depression and a

history of non-fatal suicide attempts are at high risk for post-discharge suicide. These novel interactions should be replicated in future studies and examined in conventional designs aimed at quantifying the causal joint effects of these variables.

A surprising predictor of postdischarge suicide among men in the random forests was nicotine dependence. Nicotine dependence may serve as a proxy for smoking and health problems that are linked with chronic hypoxia and risk-taking, which in turn are associated with an increased risk of suicide.[36,37] Furthermore, smoking may be a form of self-medication for emotional distress, which in itself is a risk factor for suicide.[38] Nicotine dependence may represent a potentially transdiagnostic suicide prevention factor because it appeared among the top most important predictors of postdischarge suicide above many psychiatric disorders. This finding is worthy of additional research to quantify its effect on suicide risk. It is also noteworthy that alcohol-related disorders were important for accurate prediction of suicide in men, but less so in women, according to the random forests and classification trees. Alcohol-related disorders improved prediction accuracy to a small extent in women, but they were not among the top 30 predictors in both folds in cross-validation of the random forests and they did not appear in the classification tree for women. Previous work suggests that male suicide decedents may be more likely to have had alcohol use disorders than female suicide decedents, which may explain our findings.[39] A novel predictor identified among
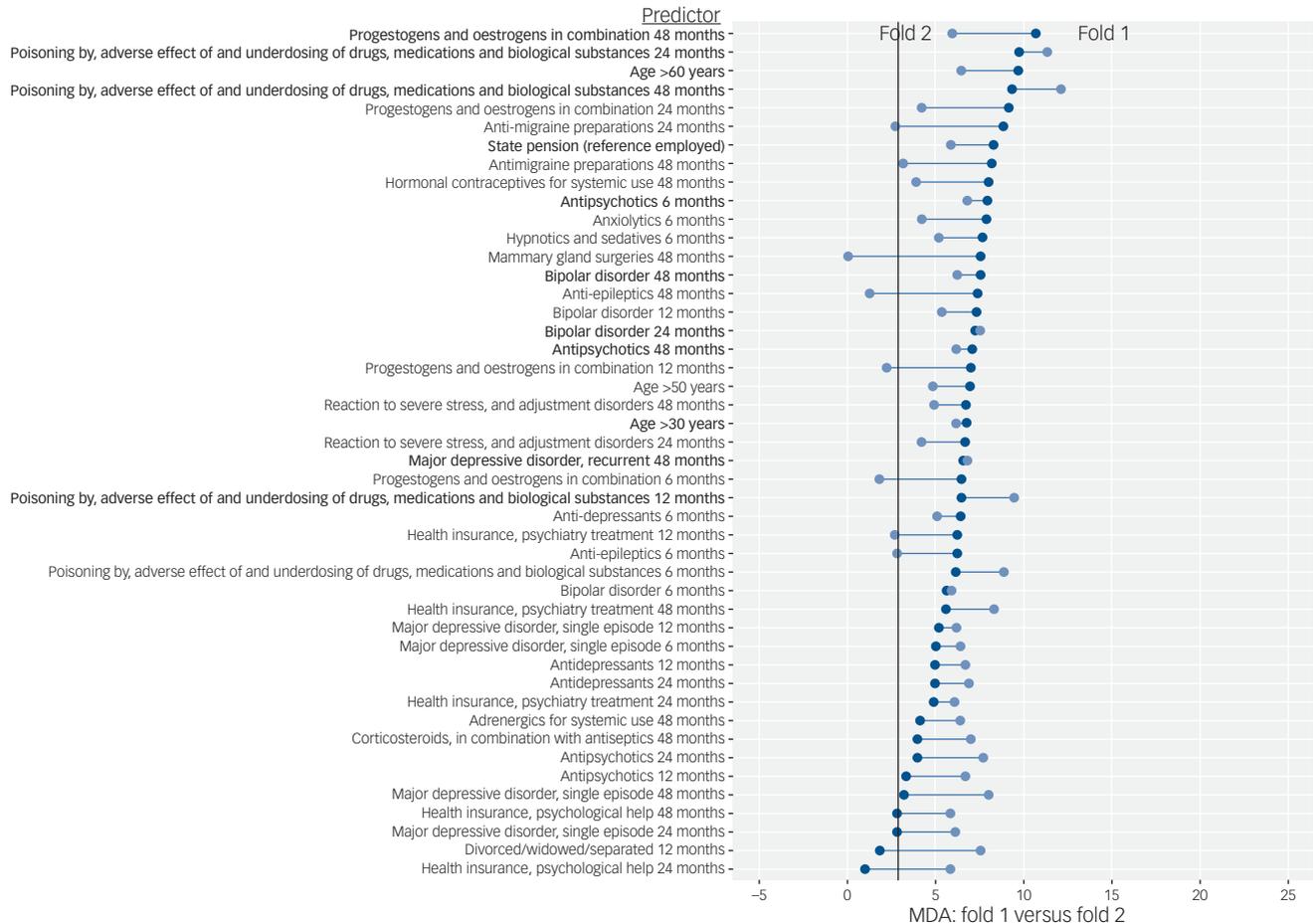
Fig. 4 Variable importance of suicide predictors among female patients hospitalised for psychiatric disorders in Denmark from split sample cross-validation, 1995–2015. The dark blue dots represent the mean decrease in accuracy (MDA) value in fold one, and the light blue dots represent the MDA value in fold two. The vertical line represents the average of the MDA values of all predictors with nonzero MDA values in folds one and two (2.9). The predictors shown in bold were in the top 30 predictors in folds 1 and 2 for women.

women was prescriptions for progestogens and oestrogens in combination. An earlier study found lower suicide mortality in those receiving oestrogen compared with those who were not receiving oestrogen.[40] However, this association may be explained by the selective prescription of postmenopausal oestrogens to healthier women.[41] It is noteworthy that these predictors should be interpreted as risk markers and not causal risk factors, given that our analyses were not intended to quantify the causal effect of any of these predictors, but rather to examine their contribution to accurate prediction of postdischarge suicide.

Our random forests' operating characteristics are comparable with those of a previous study that used machine learning to predict postdischarge suicide among veterans,[10] which found that the 5% of patients with the highest predicted risk accounted for 32% of suicides in the 1 month after psychiatric hospital discharge. In our study, persons in the highest 5% of predicted suicide risk accounted for 23% of all suicide deaths among men and 38% among women 1 month postdischarge. This suggests that a prevention program delivered to only 5% of hospitalised patients with the highest predicted risk could capture a large proportion of patients who would otherwise die by suicide.[10]

This study has several limitations. We were unable to observe more detailed social data. For many patients, the hospital may represent a respite from intolerably stressful situations, but upon discharge, patients are re-exposed to the same stressors that preceded their hospital stay, such as social isolation, financial difficulties, problematic relationships, dependent care responsibilities and other stressors.[42] We lacked data on important suicide risk factors, such as trauma exposure, sexual minority status and homelessness. Adequate representation of social conditions and acute emotional states in registry data remains an ongoing challenge. A second limitation is that there may be measurement error of variables, which may decrease model performance and distort variable importance in random forests.[43] For example, a diagnosis of poisoning by, adverse effect of and underdosing of drugs, medications and biological substances that may include non-fatal drug overdose, including accidental, intentional self-harm, assault or undetermined intent, as well as poisoning owing to adverse effects or underdosing. The broadness of this indicator makes it susceptible to measurement error. We were unable to conduct probabilistic quantitative bias analyses to examine the effect of measurement error because of the computational capacity limitations of the analytic server. A third limitation is that although we were able to perform ten-fold cross-validation for the classification trees, we were unable to do so for random forests. We instead performed two-fold cross-validation of the random forests because the analytic server was unable to conduct ten-fold cross-validation of 1000 trees. The generalisability of these results to the USA remains unclear, but our results are generally consistent with existing USA-based suicide findings.

Our results illustrate the complexity of the interactions between risk factors that elevate suicide risk in the immediate period after

psychiatric hospital discharge, and the ways that they differ between men and women. We also highlight surprising, novel factors that emerged as important predictors for accurate classification of post-discharge suicide that are worthy of additional research.

Tammy Jiang (ID), MPH, Department of Epidemiology, Boston University School of Public Health, Massachusetts, USA; **Anthony J. Rosellini**, PhD, Department of Psychological and Brain Sciences, Boston University, Massachusetts, USA; **Erzsébet Horváth-Puhó**, PhD, MS, Department of Clinical Epidemiology, Aarhus University Hospital, Denmark; **Brian Shiner**, MD, MPH, National Center for PTSD, White River Junction Veterans Affairs Medical Center, Vermont, USA; **Amy E. Street**, PhD, Women's Health Sciences Division, National Center for PTSD, VA Boston Healthcare System, Massachusetts, USA; **Timothy L. Lash**, DSc, MPH, Department of Epidemiology, Rollins School of Public Health, Emory University, Georgia, USA; **Henrik T. Sørensen**, MD, PhD, DMSc, Department of Clinical Epidemiology, Aarhus University, Denmark; **Jaimie L. Gradus** (ID), DMSc, DSc, MPH, Department of Epidemiology, Boston University School of Public Health, Massachusetts, USA

Correspondence: Tammy Jiang. Email: tjiang1@bu.edu

First received 4 Aug 2020, final revision 18 Jan 2021, accepted 30 Jan 2021

Supplementary material

Data availability

The data used for this study contain sensitive personal information and therefore cannot be made publicly available according to Danish regulations. Requests for data can be made to the Department of Clinical Epidemiology at Aarhus University Hospital.

Author contributions

J.L.G., H.T.S., T.L.L. and E.H.-P. contributed to the acquisition of the data. T.J., J.L.G., A.J.R., B.S. and A.E.S. made substantial contributions to the conception or design of the work. T.J., A.J.R., E.H.-P. and J.L.G. were involved in the analysis. All authors were involved in the interpretation of data, drafting the work or revising it critically for important intellectual content, and provided final approval of the version to be published.

Funding

Declaration of interest

None.

References

1 World Health Organization. *Suicide Data*. World Health Organization, 2019 (http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/).

2 World Health Organization. *World Health Statistics Data Visualizations Dashboard: Suicide*. World Health Organization, 2018 (https://apps.who.int/gho/data/node.sdg.3-4-viz-2?lang=en).

3 Chung DT, Ryan CJ, Hadzi-Pavlovic D, Singh SP, Stanton C, Large MM. Suicide rates after discharge from psychiatric facilities. *JAMA Psychiatry* 2017; **74**(7): 694–702.

4 Chung D, Hadzi-Pavlovic D, Wang M, Swaraj S, Olfson M, Large M. Meta-analysis of suicide rates in the first week and the first month after psychiatric hospitalisation. *BMJ Open* 2019; **9**(3): e023883.

5 Qin P, Nordentoft M. Suicide risk in relation to psychiatric hospitalization: evidence based on longitudinal registers. *Arch Gen Psychiatry* 2005; **62**(4): 427–32.

6 Olfson M, Wall M, Wang S, Crystal S, Liu S-M, Gerhard T, et al. Short-term suicide risk after psychiatric hospital discharge. *JAMA Psychiatry* 2016; **73**(11): 1119–26.

7 Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol Bull* 2017; **143**(2): 187–232.

8 Fazel S, O'Reilly L. Machine learning for suicide research-can it improve risk factor identification? *JAMA Psychiatry* 2020; **77**: 13–4.

9 Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ, et al. Predicting suicides after psychiatric hospitalization in US army soldiers: the army study to assess risk and resilience in servicemembers (Army STARRS). *JAMA Psychiatry* 2015; **72**(1): 49–57.

10 Kessler RC, Bauer MS, Bishop TM, Demler OV, Dobscha SK, Gildea SM, et al. Using administrative data to predict suicide after psychiatric hospitalization in the Veterans Health Administration system. *Front Psychiatry* 2020; **11**: 390.

11 Weissman MM, Bland RC, Canino GJ, Greenwald S, Hwu HG, Joyce PR, et al. Prevalence of suicide ideation and suicide attempts in nine countries. *Psychol Med* 1999; **29**(1): 9–17.

12 Miranda-Mendizabal A, Castellví P, Parés-Badell O, Alayo I, Almenara J, Alonso I, et al. Gender differences in suicidal behavior in adolescents and young adults: systematic review and meta-analysis of longitudinal studies. *Int J Public Health* 2019; **64**(2): 265–83.

13 Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol* 2015; **7**: 449–90.

14 Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *J Clin Epidemiol* 1999; **52**(12): 1165–72.

15 Mors O, Perto GP, Mortensen PB. The Danish Psychiatric Central Research Register. *Scand J Public Health* 2011; **39**(7 suppl): 54–7.

16 Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol* 2014; **29**(8): 541–9.

17 Pedersen CB. The Danish Civil Registration System. *Scand J Public Health* 2011; **39**(7 suppl): 22–5.

18 Helweg-Larsen K. The Danish Register of Causes of Death. *Scand J Public Health* 2011; **39**(7 suppl): 26–9.

19 Tøllefsen IM, Helweg-Larsen K, Thiblin I, Hem E, Kastrup MC, Nyberg U, et al. Are suicide deaths under-reported? Nationwide re-evaluations of 1800 deaths in Scandinavia. *BMJ Open* 2015; **5**(11): e009120.

20 Timmermans B. The Danish Integrated Database for Labor Market Research: Towards Demystification for the English Speaking Audience. DRUID Working Papers, 2010 (https://ideas.repec.org/p/aal/abbswp/10-16.html).

21 Baadsgaard M, Quitzau J. Danish registers on personal income and transfer payments. *Scand J Public Health* 2011; **39**(7 suppl): 103–5.

22 Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. *Scand J Public Health* 2011; **39**(7 suppl): 38–41.

23 McCarthy JF, Bossarte RM, Katz IR, Thompson C, Kemp J, Hannemann CM, et al. Predictive modeling and concentration of the risk of suicide: implications for preventive interventions in the US Department of Veterans Affairs. *Am J Public Health* 2015; **105**(9): 1935–42.

24 Kessler RC, Hwang I, Hoffmire CA, McCarthy JF, Petukhova MV, Rosellini AJ, et al. Developing a practical suicide risk prediction model for targeting high-risk patients in the Veterans Health Administration. *Int J Methods Psychiatr Res* 2017; **26**(3): e1575.

25 Gradus JL, Rosellini AJ, Horváth-Puhó E, Street AE, Galatzer-Levy IR, Jiang T, et al. Prediction of sex-specific suicide risk using machine learning and single-payer health care registry data from Denmark. *JAMA Psychiatry* 2020; **77**(1): 25–34.

26 Hawkins DM. The problem of overfitting. *J Chem Inf Comput Sci* 2004; **44**(1): 1–12.

27 Lühdorf P, Overvad K, Schmidt EB, Johnsen SP, Bach FW. Predictive value of stroke discharge diagnoses in the Danish National Patient Register. *Scand J Public Health* 2017; **45**(6): 630–6.

28 Tuckuviene R, Kristensen SR, Helgestad J, Christensen AL, Johnsen SP. Predictive value of pediatric thrombosis diagnoses in the Danish National Patient Registry. *Clin Epidemiol* 2010; **2**: 107–22.

29 Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychol Methods* 2009; **14**(4): 323–48.

30 James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning: with Applications in R*. Springer-Verlag, 2013.

31 Kuhn M, Johnson K. Remedies for severe class imbalance. In *Applied Predictive Modeling* (eds M Kuhn, K Johnson): 419–43. Springer Publishing, 2013.

32 Therneau T, Atkinson B, Ripley B. *rpart: Recursive Partitioning and Regression Trees.* The Comprehensive R Archive Network, 2019 (https://CRAN.R-project.org/package=rpart).

33 Liaw A, Wiener M. *Classification and Regression by randomForest. R News* 2002; **2**: 18–22 (https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf).

34 SAS Institute. *SAS/GRAPH 9.4*. SAS Institute, 2013 (www.sas.com).

35 R Development Core Team. *R: A Language and Environment for Statistical Computing*. The R Foundation, 2017 (https://www.R-project.org/).

36 Young SN. Elevated incidence of suicide in people living at altitude, smokers and patients with chronic obstructive pulmonary disease and asthma: possible

role of hypoxia causing decreased serotonin synthesis. *J Psychiatry Neurosci* 2013; **38**(6): 423–6.

37 Riblet NB, Gottlieb DJ, Hoyt JE, Watts BV, Shiner B. An analysis of the relationship between chronic obstructive pulmonary disease, smoking and depression in an integrated healthcare system. *Gen Hosp Psychiatry* 2020; **64**: 72–9.

38 Orlando M, Ellickson PL, Jinnett K. The temporal relationship between emotional distress and cigarette smoking during adolescence and young adulthood. *J Consult Clin Psychol* 2001; **69**(6): 959–70.

39 McGirr A, Séguin M, Renaud J, Benkelfat C, Alda M, Turecki G. Gender and risk factors for suicide: evidence for heterogeneity in predisposing mechanisms in a psychological autopsy study. *J Clin Psychiatry* 2006; **67**(10): 1612–7.

40 Petitti DB, Perlman JA, Sidney S. Noncontraceptive estrogens and mortality: long-term follow-up of women in the walnut creek study. *Obstet Gynecol* 1987; **70**(3): 289–93.

41 Schairer C, Adami H-O, Hoover R, Persson I. Cause-specific mortality in women receiving hormone replacement therapy. *Epidemiology* 1997; **8**(1): 59–65.

42 Owen-Smith A, Bennewith O, Donovan J, Evans J, Hawton K, Kapur N, et al. When you're in the hospital, you're in a sort of bubble." Understanding the high risk of self-harm and suicide following psychiatric discharge: a qualitative study. *Crisis* 2014; **35**(3): 154–60.

43 Jiang T, Gradus JL, Lash TL, Fox MP. Addressing measurement error in random forests using quantitative bias analysis. *Am J Epidemiol* [Epub ahead of print] 1 Feb 2021. Available from: https://doi.org/10.1093/aje/kwab010.

EXTRA CONTENT ONLINE

psychiatry in history

Meduna's Mixture: surreal ecstasy or perplexing abreaction

Jack Wellington (iD)**, Alexander Wellington** (iD) **and Grace Wellington**

Carbon dioxide therapy, a precursor of the modern-day provocation of seizures via electrical induction, became a widely accepted practice by both psychiatrists and related healthcare professionals to better treat neuroses and psychoses.

Dr Ladislas J. Meduna, neuropathologist and Professor of Psychiatry at the University of Illinois College of Medicine in Chicago, invented the concoction, later dubbed 'Meduna's Mixture', in 1921. Its formula consisted of 30% carbon dioxide and 70% oxygen (collectively termed 'carbogen'). However, carbon dioxide concentrations would vary from 1.5 to 50%, at the discretion of the practising psychiatrist or psychologist experimenting in the fields of psychosis and schizophrenia management. To induce abreaction, a psychoanalytical axiom of catharsis allowing individuals to become aware of repressed emotions or traumatic experiences, Meduna would administer 20–30 breaths of the gas mixture, resulting in an intentional unconscious, altered state leading to convulsion within 1–3 min.

Meduna described patients' sensory phenomena as unclassifiable, as 'the patient is not "sleeping" in the physiological sense, nor is he in the state of consciousness' concerning true hypnagogic or visual hallucinations.[1] Accounts demonstrating such experiences recall apparitions of a 'bright light, like the sun' after 20 respirations, others describing out-of-body sensations as if 'out in space,' 'wonderful' and 'marvellous'.[1] However, Meduna admitted to preconvulsive events that did not follow experiences of 'celestial beauty and serenity' but were 'horrifying beyond description' and these individuals were advised not to receive further psychotherapeutic interventions owing to fear of similar reactions. These experiences from carbogen inhalation relate to the consequences of a drastic elevation of blood acidity via excess carbonic acid production, and the preconvulsive events of either intense feelings of euphoria or sensations of impending horror are a direct result of such changes in the patient's biochemistry. Patients undergoing such treatment, be it challenging for some or pleasant for others, would then become climatised to future psychedelic therapies if required.

This form of convulsive therapy was thought to be useful for managing neuroses and relieving obsessive–compulsive disorder, allowing carbogen users to neglect repressed emotions and ideas while permitting a smoother transition to a more profound psychedelic state. Although Meduna paved the way for neuropsychiatrists and related clinicians regarding the development and practice of pharmacological convulsive therapy, administration of Meduna's Mixture was extensively abandoned owing to its ineffectiveness when compared with other forms of convulsive therapy. An example of such comparisons involves landmark work by Italian neurophysiologist Ugo Cerletti and psychiatrist Lucio Bini, who discovered electroconvulsive therapy in 1938.

The contributions Meduna has made to psychiatry have been historically significant and pivotal to the progression of managing difficult or treatment-resistant mental disorders via convulsive therapy. Although pharmacological forms of these therapies have been widely discarded in contemporary psychiatric practice, we must not forget how the application of Meduna's Mixture became instrumental in advancing psychiatric practice to this day.

Reference

1 Meduna LJ. *Carbon Dioxide Therapy*. Charles C. Thomas Publishing, 1950.