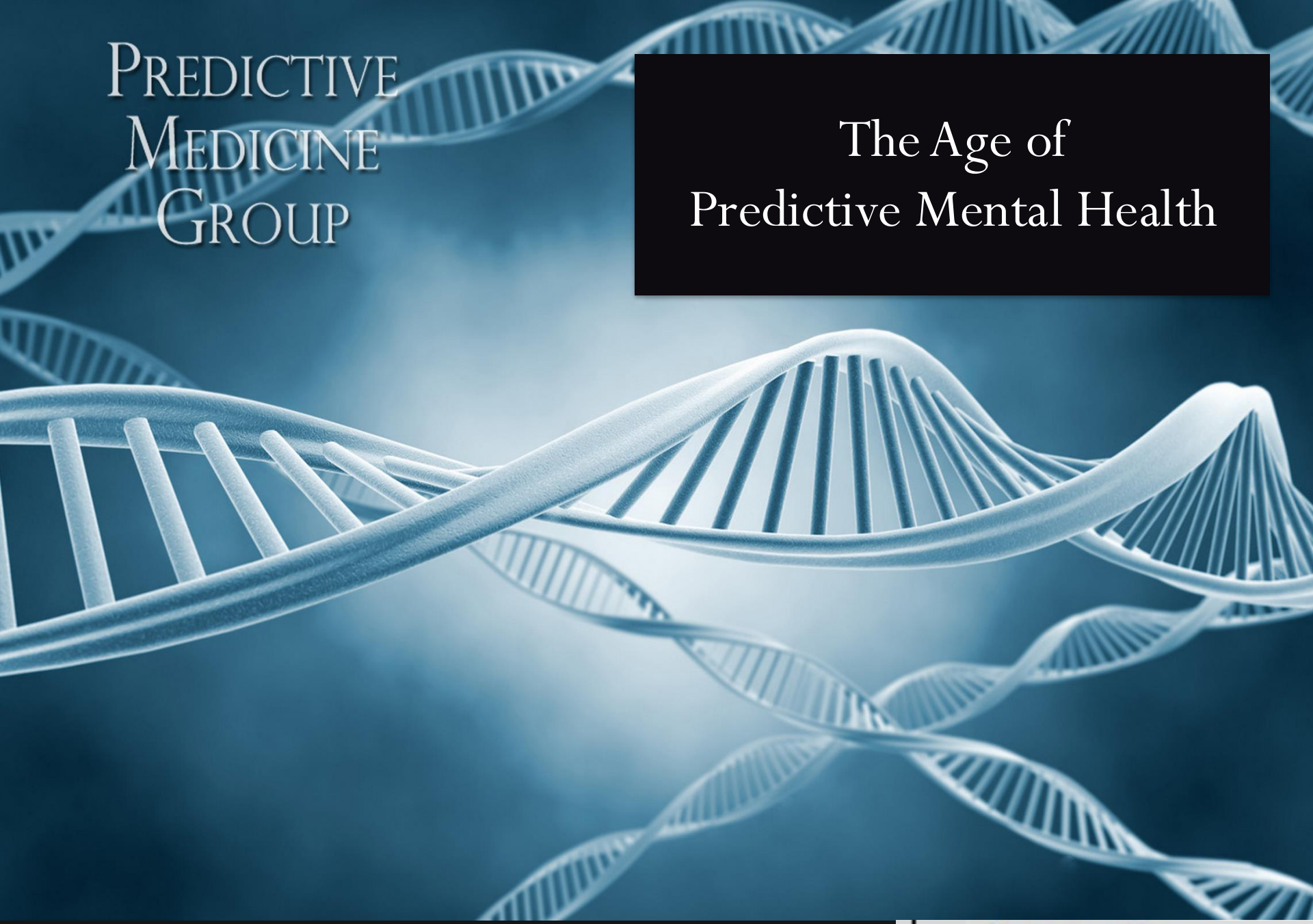


The Age of Predictive Mental Health

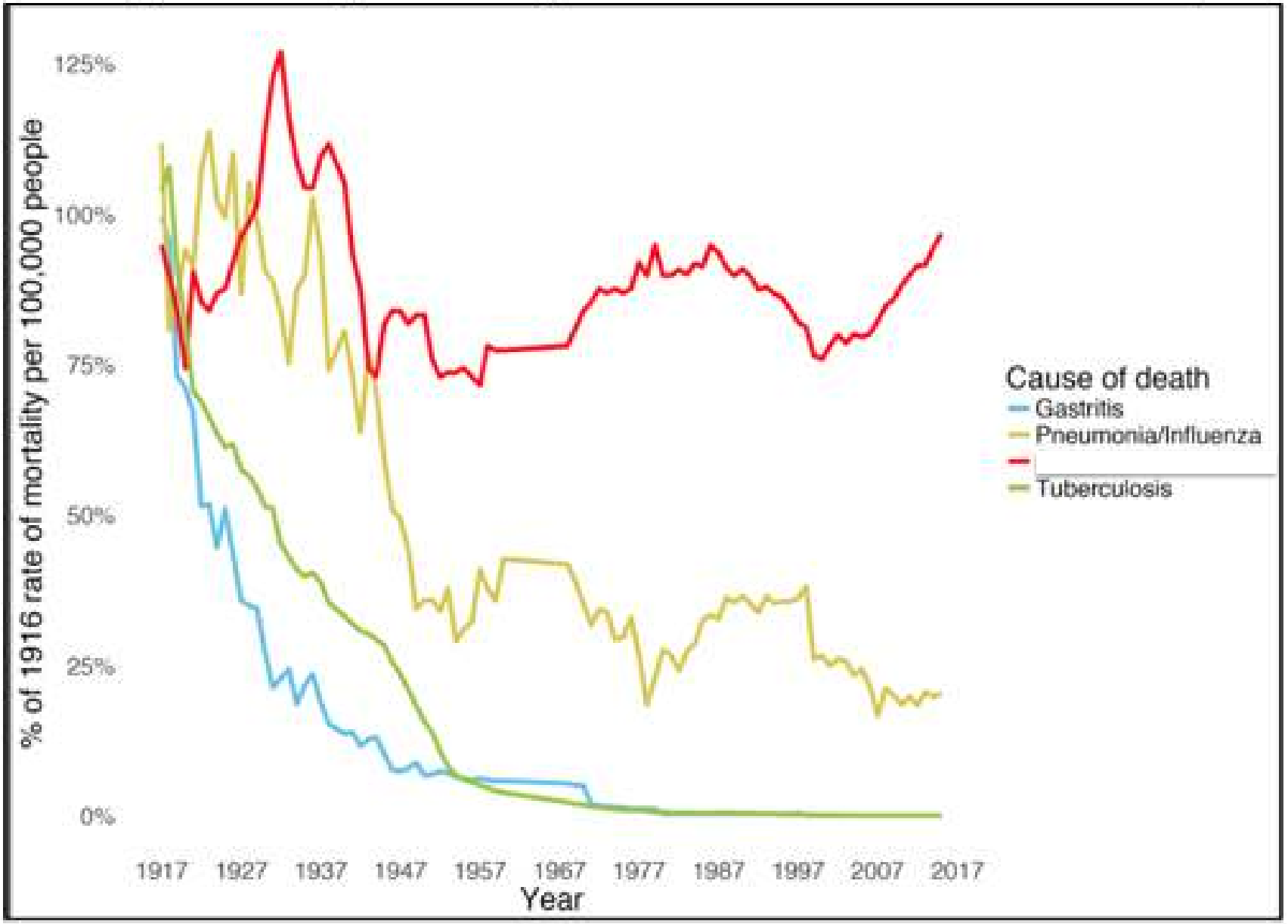
Ben Reis

chip
Boston Children's Hospital
HARVARD MEDICAL SCHOOL



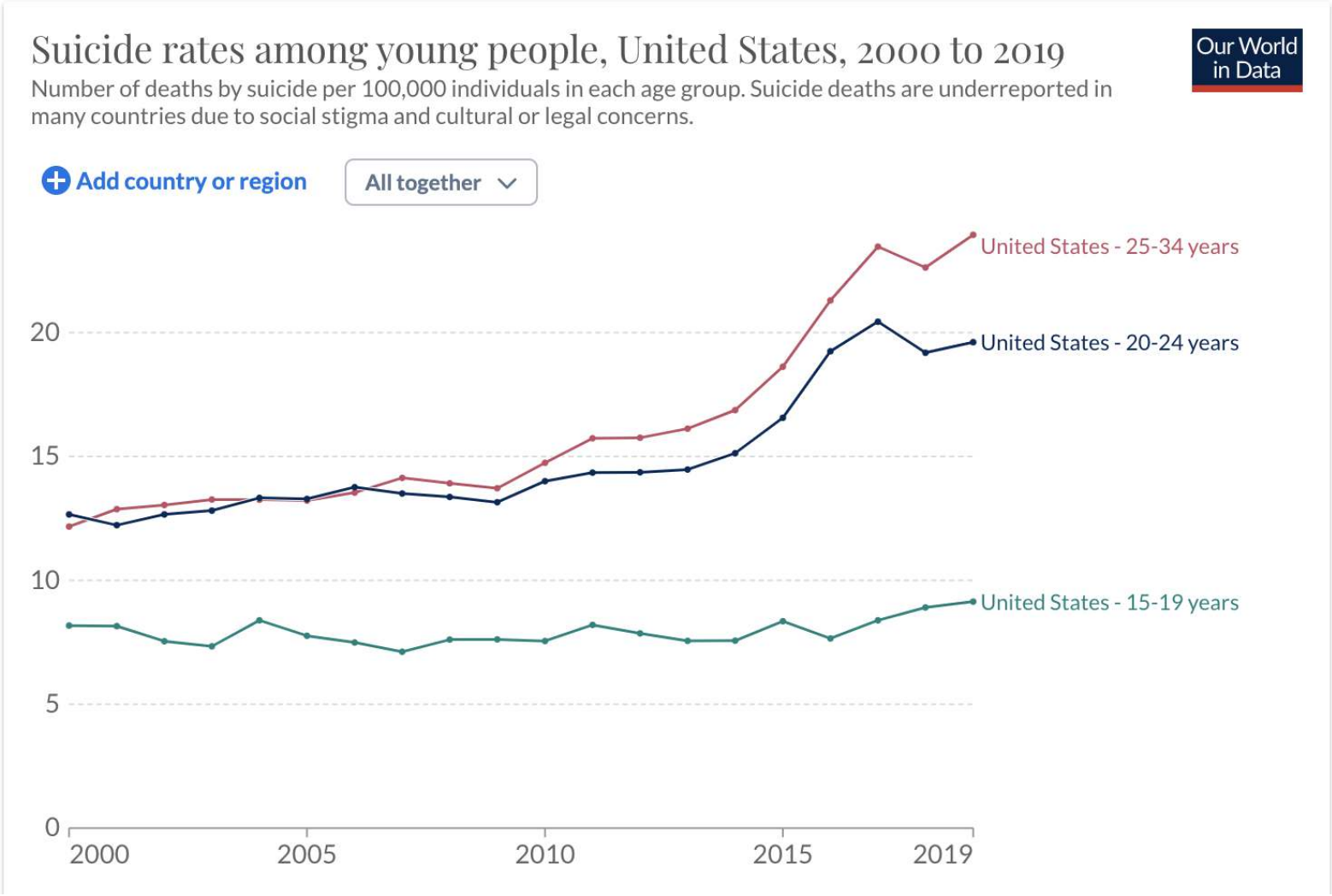


Figure 1: Change in Leading Causes of Death Over the Last Century




{SUICIDE}
IS 100% PREVENTABLE
SPEAK UP
REACH OUT

Suicide rates have increased sharply in the past 20 years, particularly in the younger age groups

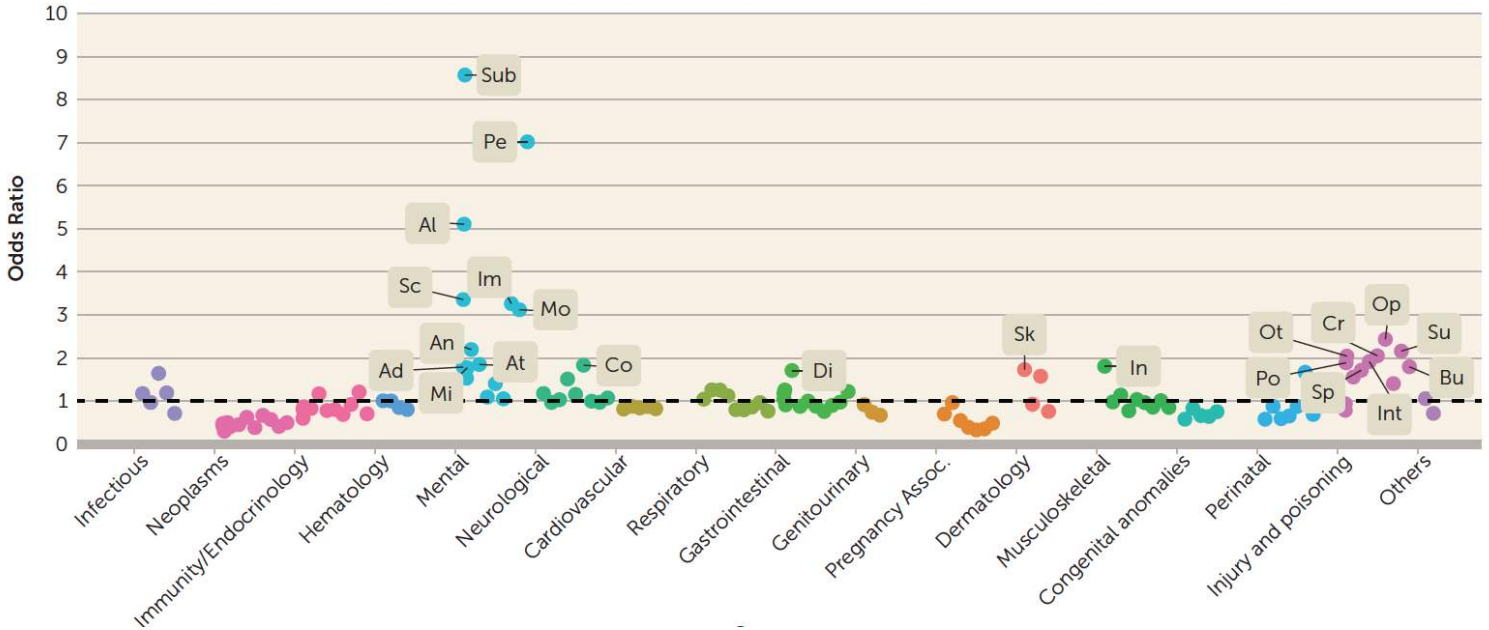


The American Journal of
Psychiatry

Predicting Suicidal Behavior From Longitudinal Electronic Health Records

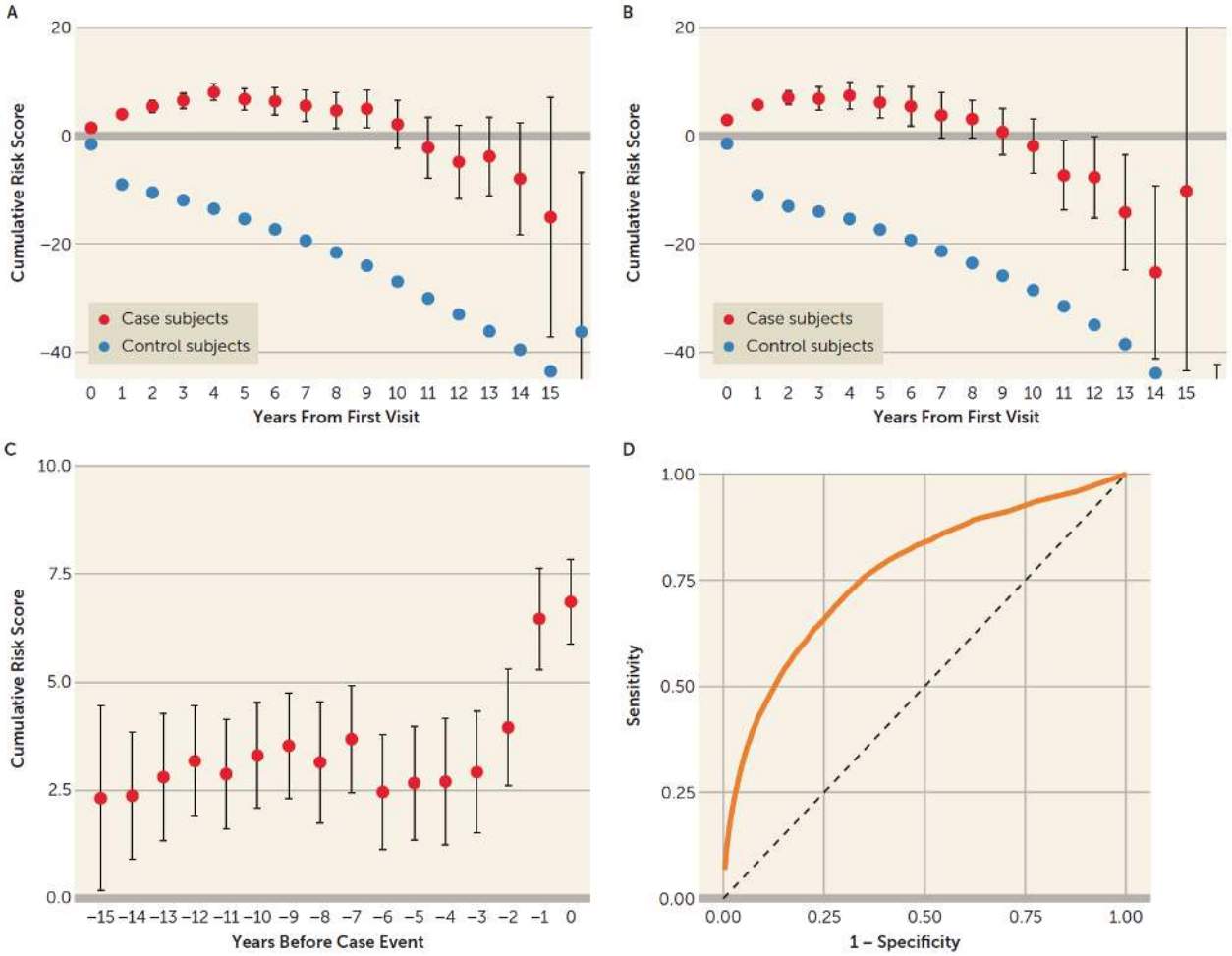
Yuval Barak-Corren, M.S., Victor M. Castro, M.S., Solomon Javitt, M.D., Alison G. Hoffnagle, M.S., Yael Dai, B.A., Roy H. Perlis, M.D., M.Sc., Matthew K. Nock, Ph.D., Jordan W. Smoller, M.D., Sc.D., Ben Y. Reis, Ph.D.

FIGURE 1. Summary View of Odds Ratios by Diagnostic Category[a]



These tools are then able to identify between a third and a half of all suicide attempts, on average three years before they occur.

FIGURE 2. Cumulative Risk Score Over Time[a]



JAMA Network Open

Original Investigation | Health Informatics

Validation of an Electronic Health Record–Based Suicide Risk Prediction Modeling Approach Across Multiple Health Care Systems

Yuval Barak-Corren, MD; Victor M. Castro, MS; Matthew K. Nock, PhD; Kenneth D. Mandl, MD, MPH; Emily M. Madsen, BS; Ashley Seiger, MSc; William G. Adams, MD; R. Joseph Applegate, BS; Elmer V. Bernstam, MD; Jeffrey G. Klann, PhD; Ellen P. McCarthy, PhD; Shawn N. Murphy, MD, PhD; Marc Natter, MD; Brian Ostasiewski, BS; Nandan Patibandla, MS; Gary E. Rosenthal, MD; George S. Silva, BS; Kun Wei, BS; Griffin M. Weber, MD, PhD; Sarah R. Weiler, PhD; Ben Y. Reis, PhD; Jordan W. Smoller, MD, ScD

Molecular Psychiatry (2017) 22, 544–551
© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved 1359-4184/17
www.nature.com/mp

ORIGINAL ARTICLE

Predicting suicides after outpatient mental health visits in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS)

RC Kessler[1], MB Stein[2,3], MV Petukhova[1], P Bliese[4], RM Bossarte[5], EJ Bromet[6], CS Fullerton[7], SE Gilman[8,9], C Ivany[10], L Lewandowski-Romps[11], A Millikan Bell[12], JA Naifeh[7], MK Nock[13], BY Reis[14], AJ Rosellini[1], NA Sampson[1], AM Zaslavsky[1] and RJ Ursano[7]

AJPH RESEARCH

Sexual Assault Victimization and Mental Health Treatment, Suicide Attempts, and Career Outcomes Among Women in the US Army

Anthony J. Rosellini, PhD, Amy E. Street, PhD, Robert J. Ursano, MD, Wai Tat Chiu, MS, Steven G. Heeringa, PhD, John Monahan, PhD, James A. Naifeh, PhD, Maria V. Petukhova, PhD, Ben Y. Reis, PhD, Nancy A. Sampson, BA, Paul D. Bliese, PhD, Murray B. Stein, MD, Alan M. Zaslavsky, PhD, and Ronald C. Kessler, PhD

Improving risk prediction for target subpopulations: Predicting suicidal behaviors among multiple sclerosis patients

Yuval Barak-Corren, Victor M. Castro, Solomon Javitt, Matthew K. Nock, Jordan W. Smoller, Ben Y. Reis

Published: February 16, 2023 • https://doi.org/10.1371/journal.pone.0277483

npj | digital medicine

Explore content · About the journal · Publish with us

nature > npj digital medicine > articles > article

Article | Open access | Published: 27 January 2022

Predictive structured–unstructured interactions in EHR models: A case study of suicide prediction

Ilkin Bayramli, Victor Castro, Yuval Barak-Corren, Emily M. Madsen, Matthew K. Nock, Jordan W. Smoller & Ben Y. Reis

JAMA Network

JAMA Psychiatry

Original Investigation

FREE

March 17, 2021

Accuracy Requirements for Cost-effective Suicide Risk Prediction Among Primary Care Patients in the US

Eric J. Ross, MD[1,2,3]; Kelly L. Zuromski, PhD[4]; Ben Y. Reis, PhD[5]; et al

Journal of the American Medical Informatics Association, 29(1), 2022, 62–71
doi: 10.1093/jamia/ocab225
Advance Access Publication Date: 2 November 2021
Research and Applications

Research and Applications

Temporally informed random forests for suicide risk prediction

Ilkin Bayramli[1,2], Victor Castro[3,4], Yuval Barak-Corren[1], Emily M. Madsen[5,6], Matthew K. Nock[4,7,8], Jordan W. Smoller[5,6,9], and Ben Y. Reis [1,9]

Psychiatry Research
Volume 323, May 2023, 115175

An efficient landmark model for prediction of suicide attempts in multiple clinical settings



Frist Episode Psychosis


MENTAL
HEALTH


ABUSE


CHILDREN CANNOT
STOP CHILD ABUSE.
ADULTS CAN...

BMJ

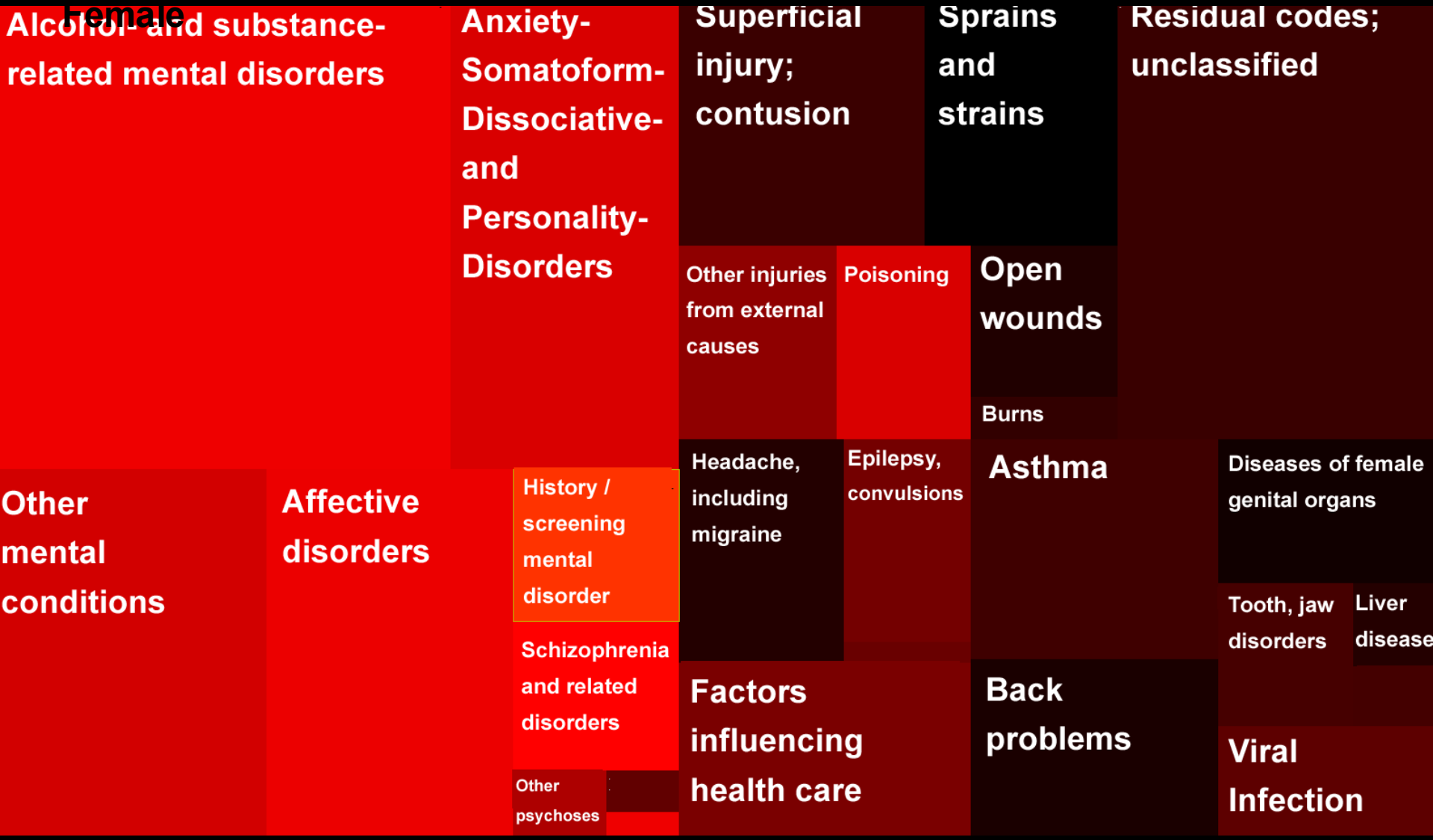
RESEARCH

Longitudinal histories as predictors of future diagnoses of domestic abuse: modelling study

Ben Y Reis, assistant professor,[1,2] Isaac S Kohane, professor,[1,2] Kenneth D Mandl, associate professor[1,2]

Table 2 | Assault and intentional injury related ICD-9 codes added to codes in table 1 to form broader case definition

E960	Fight, brawl, rape
E960.0	Unarmed fight or brawl
E960.1	Rape
E961	Assault—corrosive/caustic agent
E962.0	Assault—poisoning with medical agent
E962.1	Other solid and liquid substances
E962.2	Assault—poisoning with gas/vapour
E962.9	Unspecified poisoning
E963	Assault—hanging/strangulation
E964	Assault by submersion
E965.0	Assault—handgun
E965.1	Assault—shotgun
E965.3	Assault—military firearms
E965.4	Assault—firearm (not classified elsewhere)
E965.6	Gasoline bomb
E965.8	Assault—explosive (not classified elsewhere)
E965.9	Unspecified explosive
E966	Assault by cutting and piercing instrument
E968	Assault by other and unspecified means
E968.0	Assault—fire
E968.1	Assault—push from high place
E968.2	Assault—striking with object
E968.3	Assault—hot liquid
E968.4	Criminal neglect: abandonment of child, infant, or other helpless person with intent to injure or kill
E968.5	Assault—transport vehicle
E968.6	Assault—air gun
E968.7	Human bite—assault
E968.8	Assault (not classified elsewhere)
E968.9	Assault (not otherwise specified)
E969	Late effect assault


Female
Alcohol- and substance-related mental disorders
Anxiety-Somatoform-Dissociative-and Personality-Disorders
Superficial injury; contusion
Sprains and strains
Residual codes; unclassified
Other injuries from external causes
Poisoning
Open wounds
Burns
Other mental conditions
Affective disorders
History / screening mental disorder
Headache, including migraine
Epilepsy, convulsions
Asthma
Diseases of female genital organs
Tooth, jaw disorders
Liver disease
Schizophrenia and related disorders
Factors influencing health care
Back problems
Viral Infection
Other psychoses

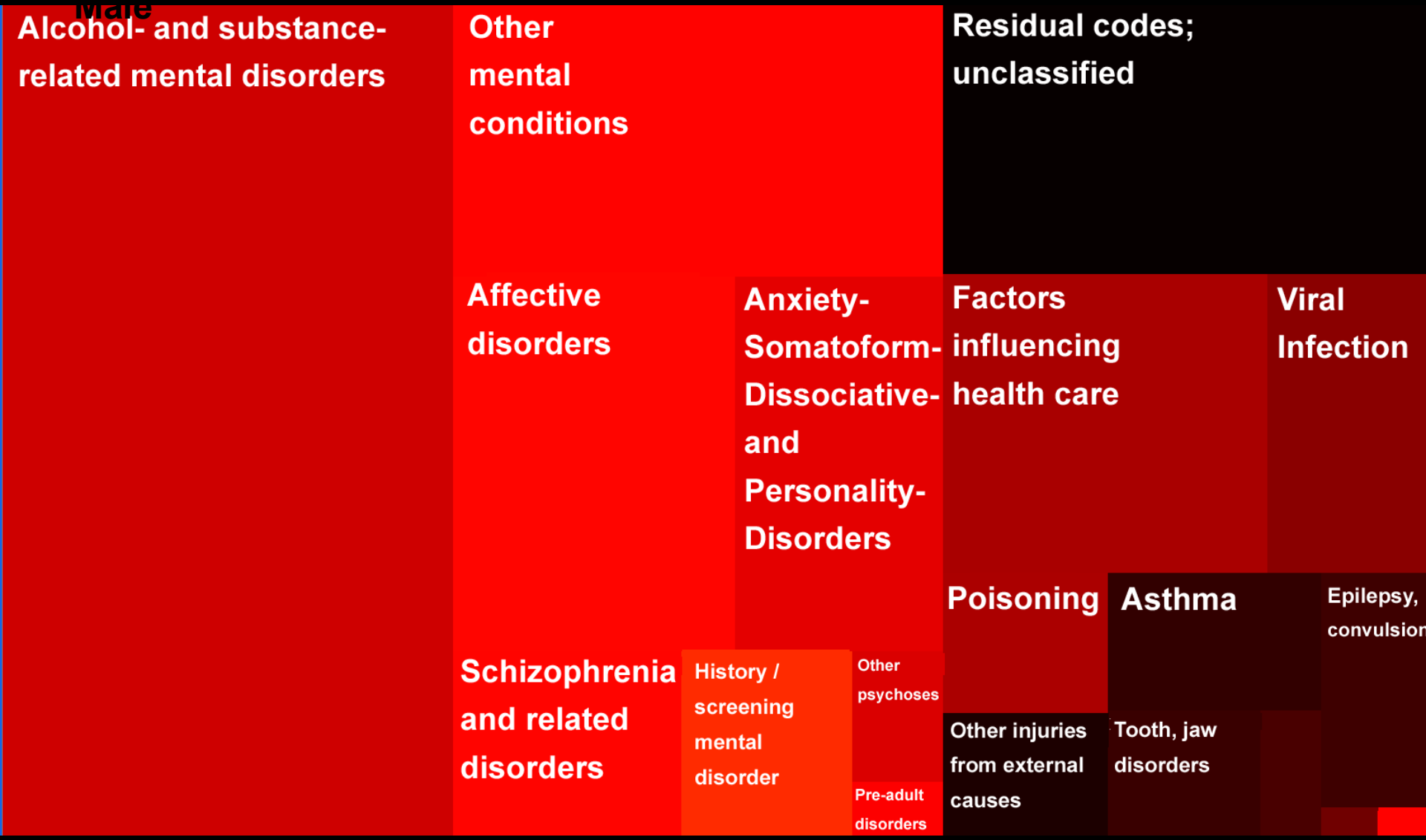
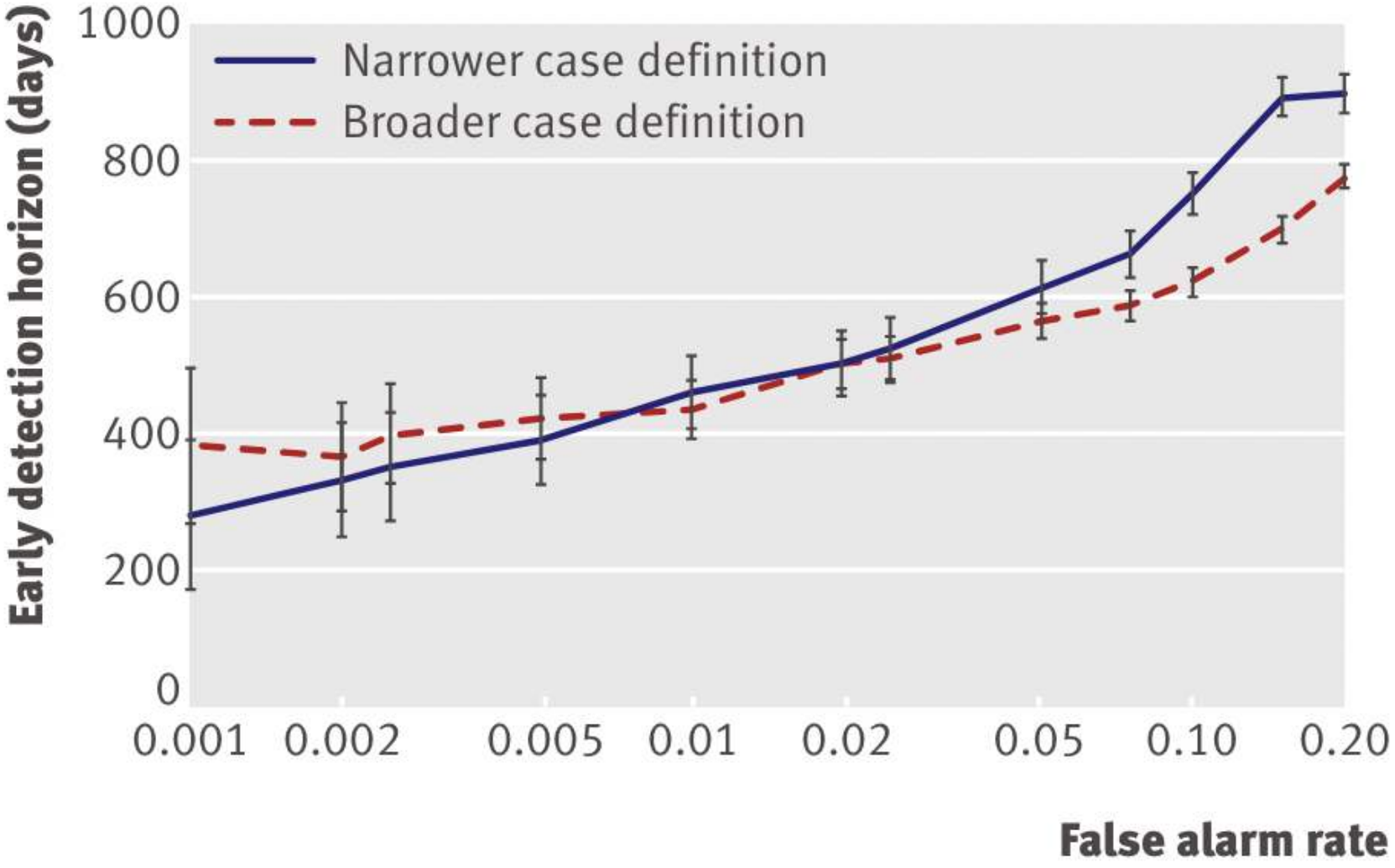
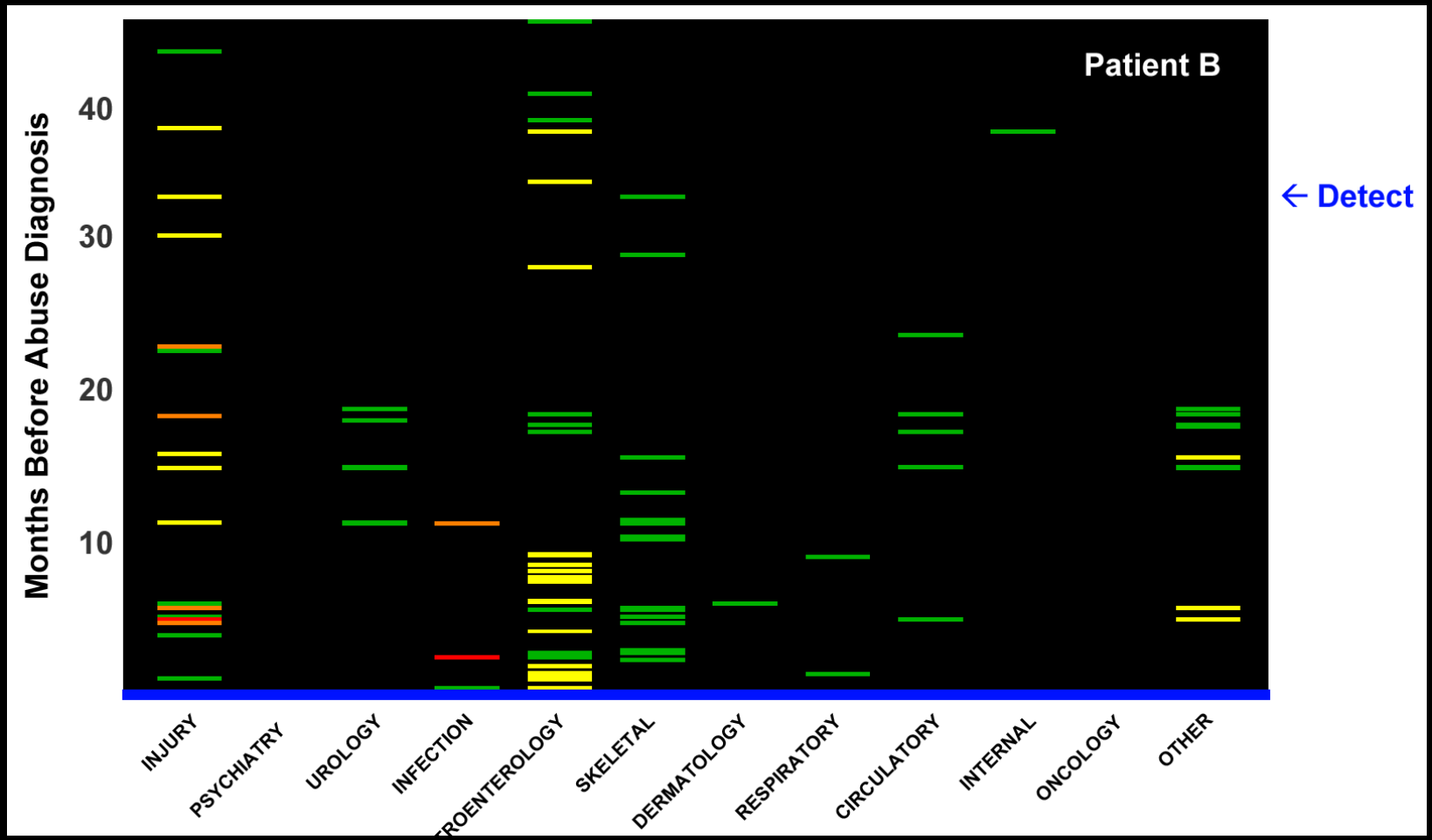

Male
Alcohol- and substance-related mental disorders
Other mental conditions
Residual codes; unclassified
Affective disorders
Anxiety-Somatoform-Dissociative- and Personality-Disorders
Factors influencing health care
Viral Infection
Poisoning
Asthma
Epilepsy, convulsion
Schizophrenia and related disorders
History / screening mental disorder
Other psychoses
Other injuries from external causes
Tooth, jaw disorders
Pre-adult disorders

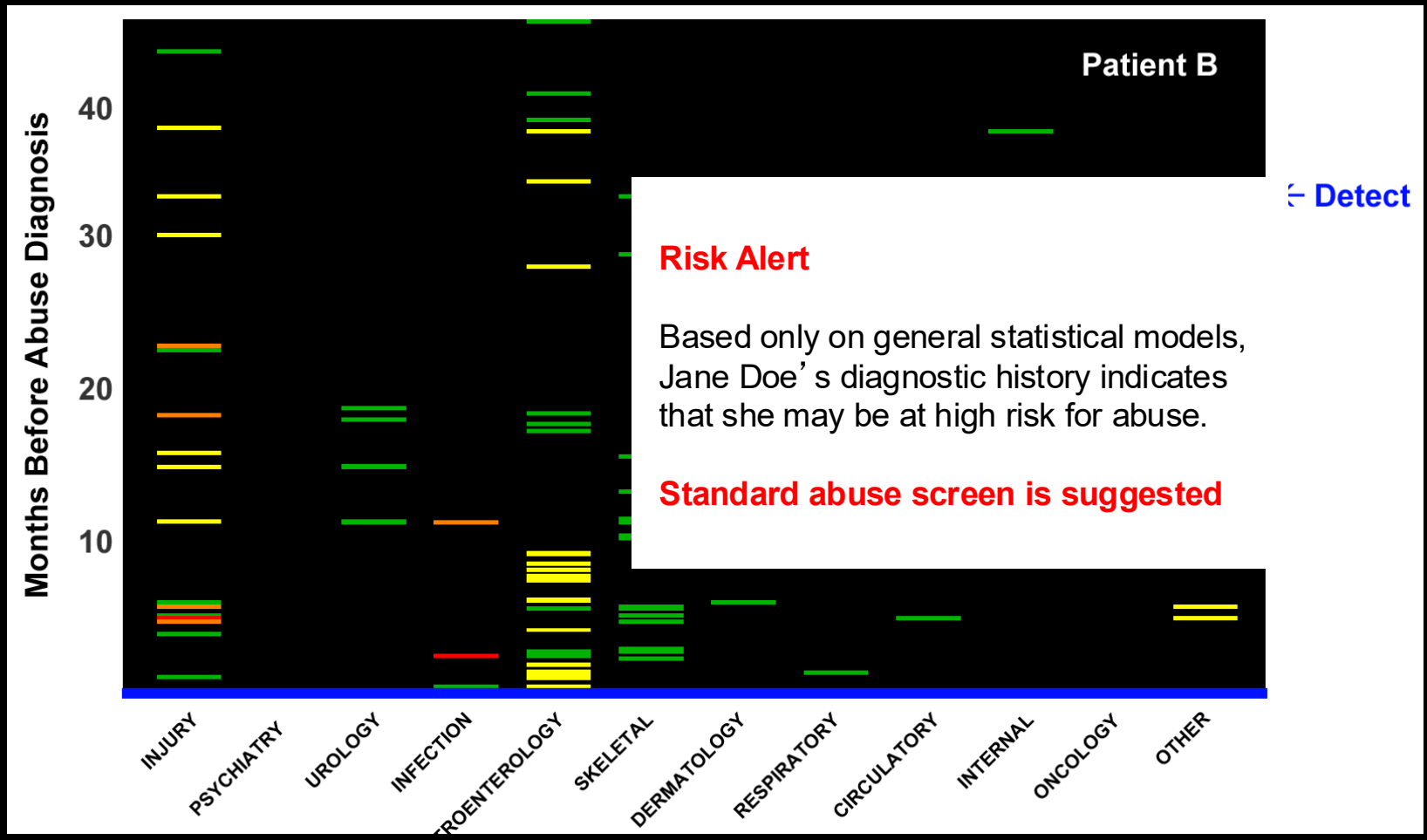
Table 4 | Partial risk scores* for women and men for select clinical categories

Category†	Women (95% CI)	Men (95% CI)
Alcohol, substance related mental disorders	1.455 (1.440 to 1.471)	1.253 (1.235 to 1.271)
Injuries from external causes	0.885 (0.843 to 0.925)	0.175 (0.098 to 0.249)
Poisoning	1.326 (1.279 to 1.373)	1.039 (0.960 to 1.115)
Affective disorders	1.435 (1.410 to 1.459)	1.726 (1.688 to 1.764)
Other mental conditions	1.283 (1.260 to 1.305)	1.640 (1.606 to 1.673)
Other psychoses	1.065 (0.980 to 1.148)	1.326 (1.209 to 1.434)

*The higher the partial risk score, the more predictive the category of diagnoses is of abuse.
†First three categories listed are more predictive of abuse in women than in men. Second three categories listed are more predictive of abuse in men than in women.


Early detection horizon (days)
1000
800
600
400
200
0
Narrower case definition
Broader case definition
0.001
0.002
0.005
0.01
0.02
0.05
0.10
0.20
False alarm rate


Patient B
Months Before Abuse Diagnosis
40
30
20
10
← Detect
INJURY
PSYCHIATRY
UROLOGY
INFECTION
SKELETAL
DERMATOLOGY
RESPIRATORY
CIRCULATORY
INTERNAL
ONCOLOGY
OTHER


Patient B
Months Before Abuse Diagnosis
40
30
20
10
Detect
Risk Alert
Based only on general statistical models, Jane Doe's diagnostic history indicates that she may be at high risk for abuse.
Standard abuse screen is suggested
INJURY
PSYCHIATRY
UROLOGY
INFECTION
SKELETAL
DERMATOLOGY
RESPIRATORY
CIRCULATORY
INTERNAL
ONCOLOGY
OTHER

Prediction in Practice

JAMIA
A SCHOLARLY JOURNAL OF INFORMATICS IN HEALTH AND BIOMEDICINE

Integration of physical abuse clinical decision support into the electronic health record at a Tertiary Care Children's Hospital

Srinivasan Suresh, Richard A Saladino, Janet Fromkin, Emily Heineman, Tom McGinn, Rudolph Richichi, Rachel P Berger

CDS system for pediatric and general EDs that notified providers of concern for CA based on a combination of active and passive screens that triggered EHR alerts.

Good accuracy at identifying cases diagnosed as possible, probable, or definite abuse by the child protection team and showed increased identification of high-risk injuries and reports made to CPS by ED providers.

JAMIA
A SCHOLARLY JOURNAL OF INFORMATICS IN HEALTH AND BIOMEDICINE

Developing machine learning–based models to help identify child abuse and neglect: key ethical challenges and recommended solutions

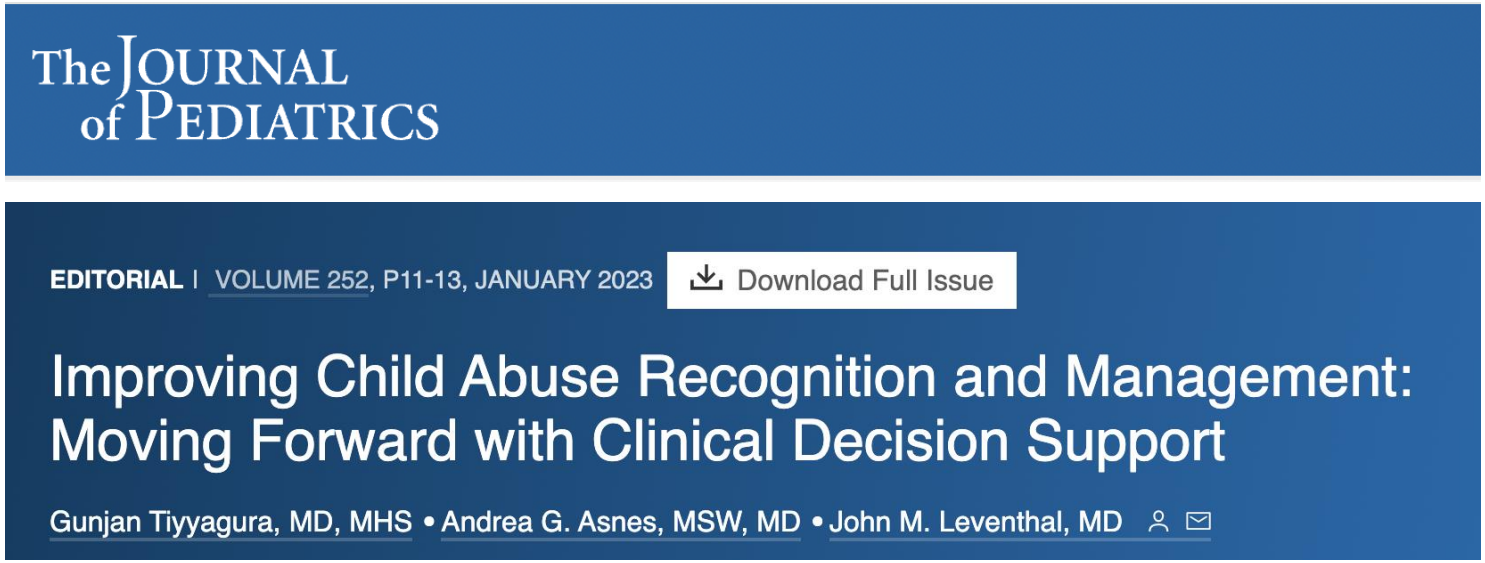
Get access >

Aviv Y Landau ✉, Susi Ferrarello, Ashley Blanchard, Kenrick Cato, Nia Atkins, Stephanie Salazar, Desmond U Patton, Maxim Topaz

Journal of the American Medical Informatics Association, Volume 29, Issue 3, March 2022,

Key Challenges:

- Biases in the data
- Lack of “gold standard “in assessment and diagnosis of child abuse and neglect;
- Challenges in testing predictive models in practice;
- Challenges in presentation of machine learning–based prediction to clinicians and patients.

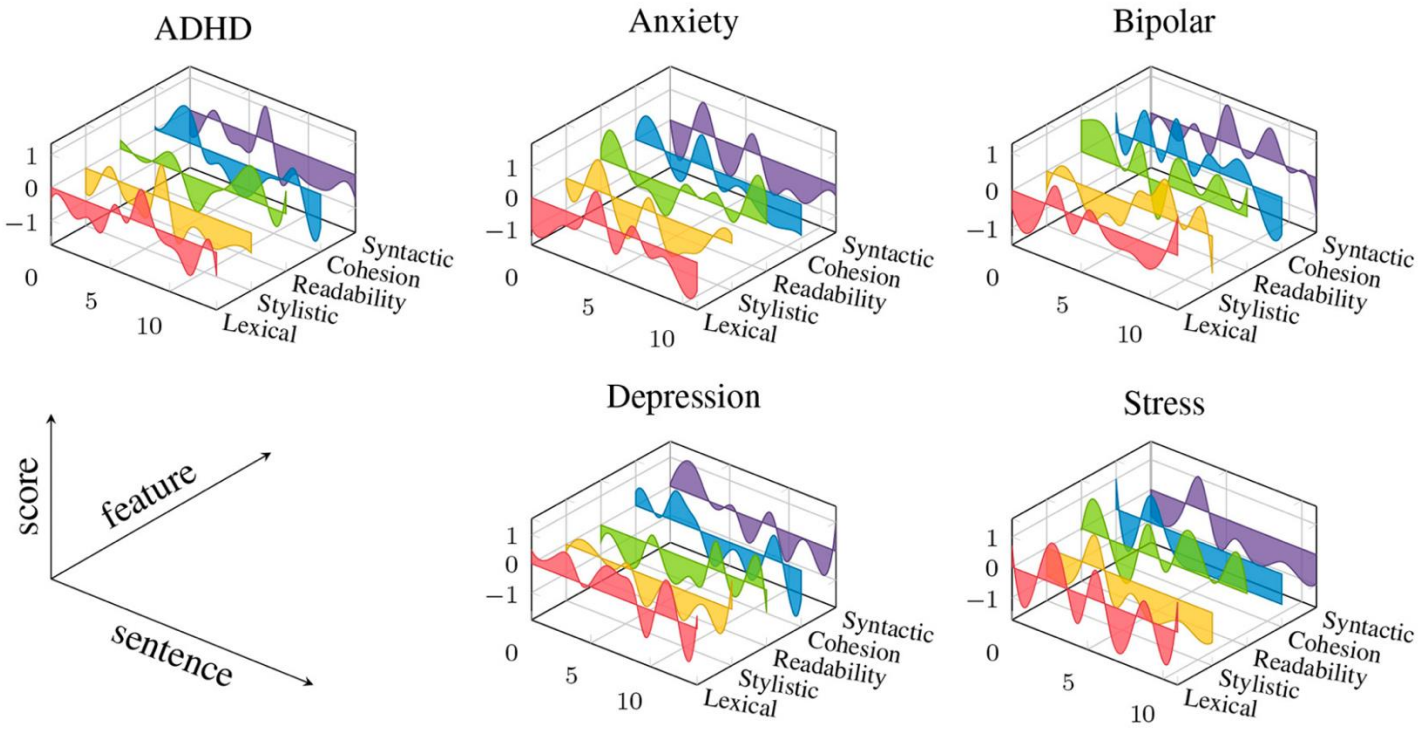


- Broad implementation of a CA-CDS system in both general and pediatric ED settings for all children <4 y.o.
- Use of active alerts to providers with links to evidence-based order sets embedded in the EHR
- Standardized hand-offs between medical professionals and incorporated into the CA-CDS system;
- Integration of CA-CDS that is feasible, sustainable, and easily disseminated among institutions.

The Importance of Explainability

Toward explainable AI (XAI) for mental health detection based on language behavior

Elma Kerz[1]*, Sourabh Zanwar[1], Yu Qiao[1] and Daniel Wiechmann[2]*



How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review

Rikard Rosenbacke[1], MSc; Åsa Melhus[2], MD, PhD; Martin McKee[3], MD, DSc; David Stuckler[4], PhD

Out 10 studies, 5 reported that XAI increased clinicians' trust, particularly when the explanations were clear, concise, and relevant to clinical practice. 2 studies highlighted that XAI could either enhance or diminish trust, depending on the complexity and coherence of the provided explanations.

Integrating Novel Data Sources

Article

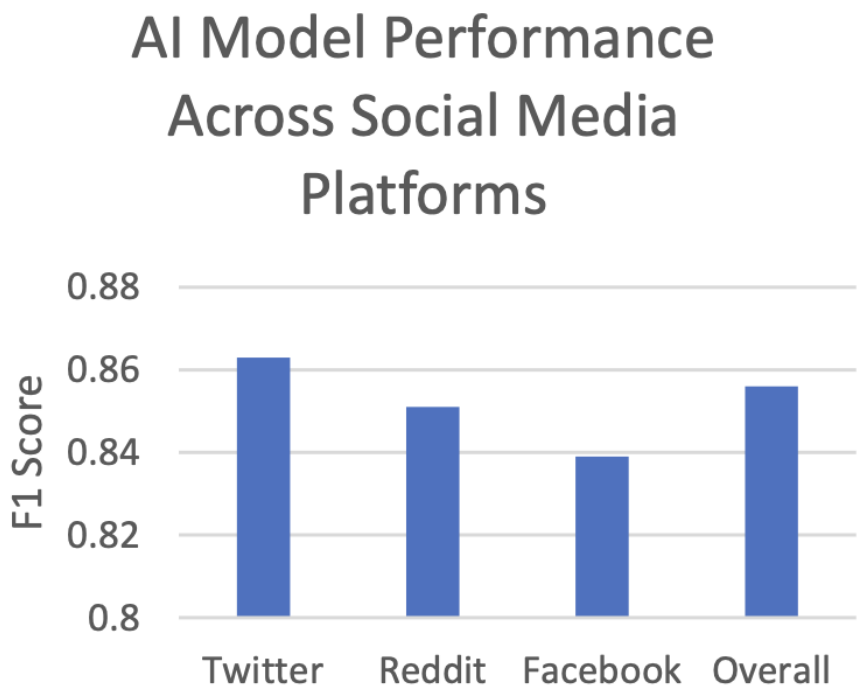
Early Detection of Mental Health Crises through Artifical-Intelligence-Powered Social Media Analysis: A Prospective Observational Study

Masab A. Mansoor [1,*] and Kashif H. Ansari [2]

996,452 social media posts in English, Spanish, Mandarin, and Arabic from Twitter, Reddit, and FB;

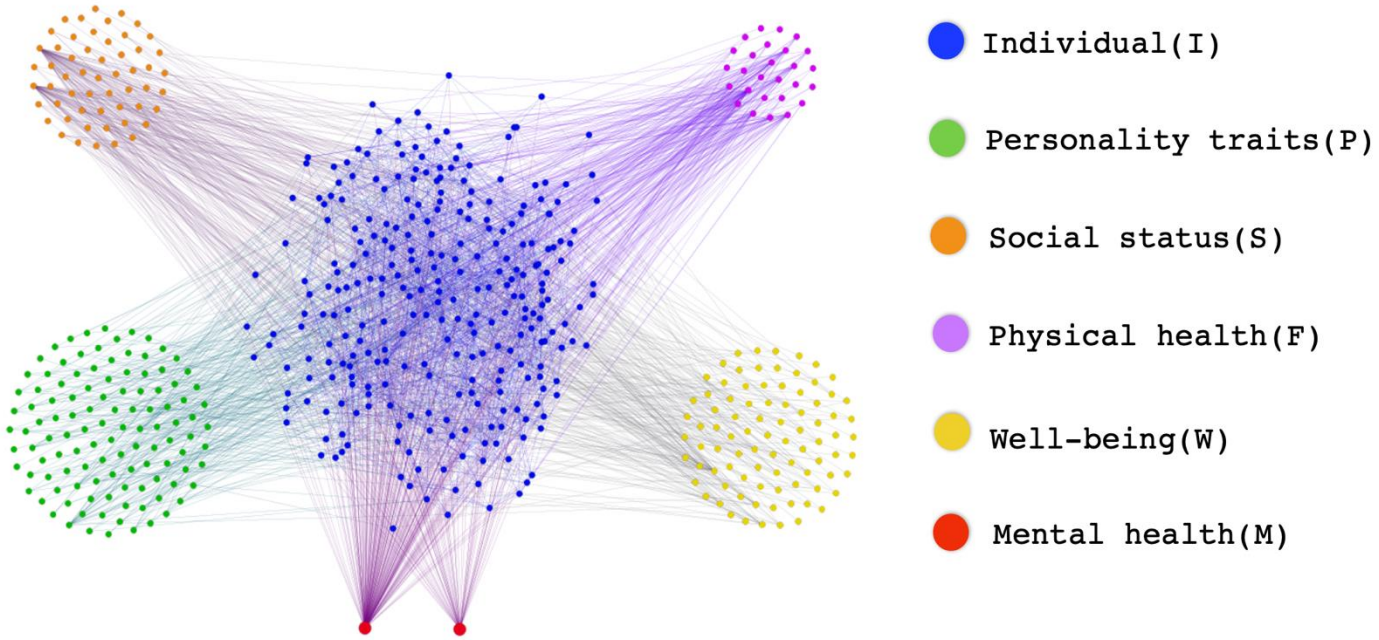
Identified depressive episodes (91.2%), manic episodes (88.7%), suicidal ideation (93.5%), and anxiety crises (87.3%).

Average lead time of 7.2 days before human expert identification.



Heterogeneous network approach to predict individuals' mental health

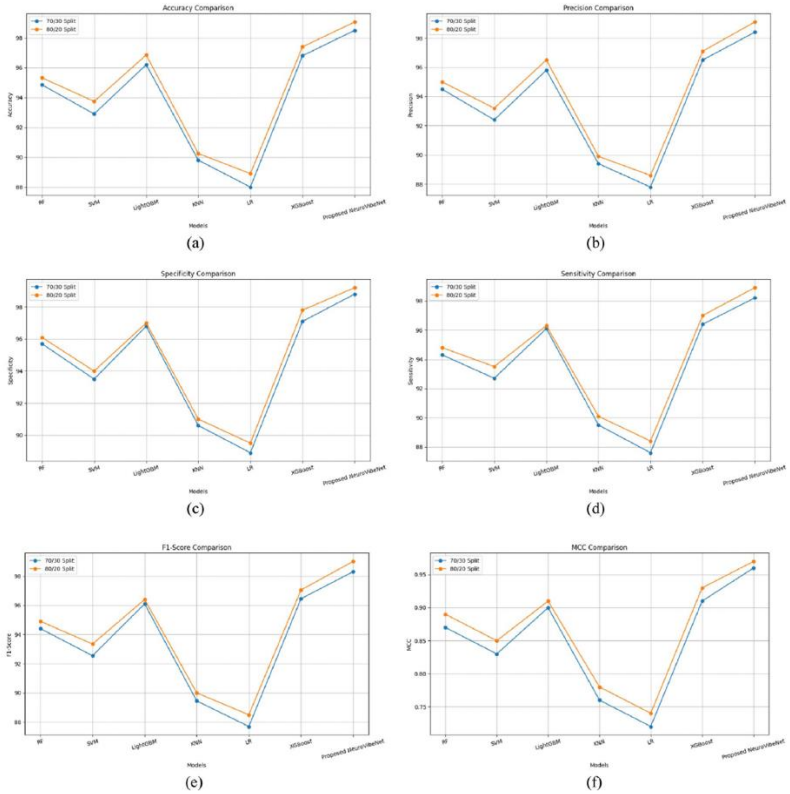
Shikang Liu[1], Fatemeh Vahedian[1], David Hachen[2], Omar Lizardo[3], Christian Poellabauer[1], Aaron Striegel[1], and Tijana Milenković[1,4,5,*]



Early detection of mental health disorders using machine learning models using behavioral and voice data analysis

Sunil Kumar Sharma[1,2]✉, Ahmed Ibrahim Alutaibi[3], Ahmad Raza Khan[4], Ghanshyam G. Tejani[5,6]✉, Fuzail Ahmad[7,8] & Seyed Jalaleddin Mousavirad[9]✉

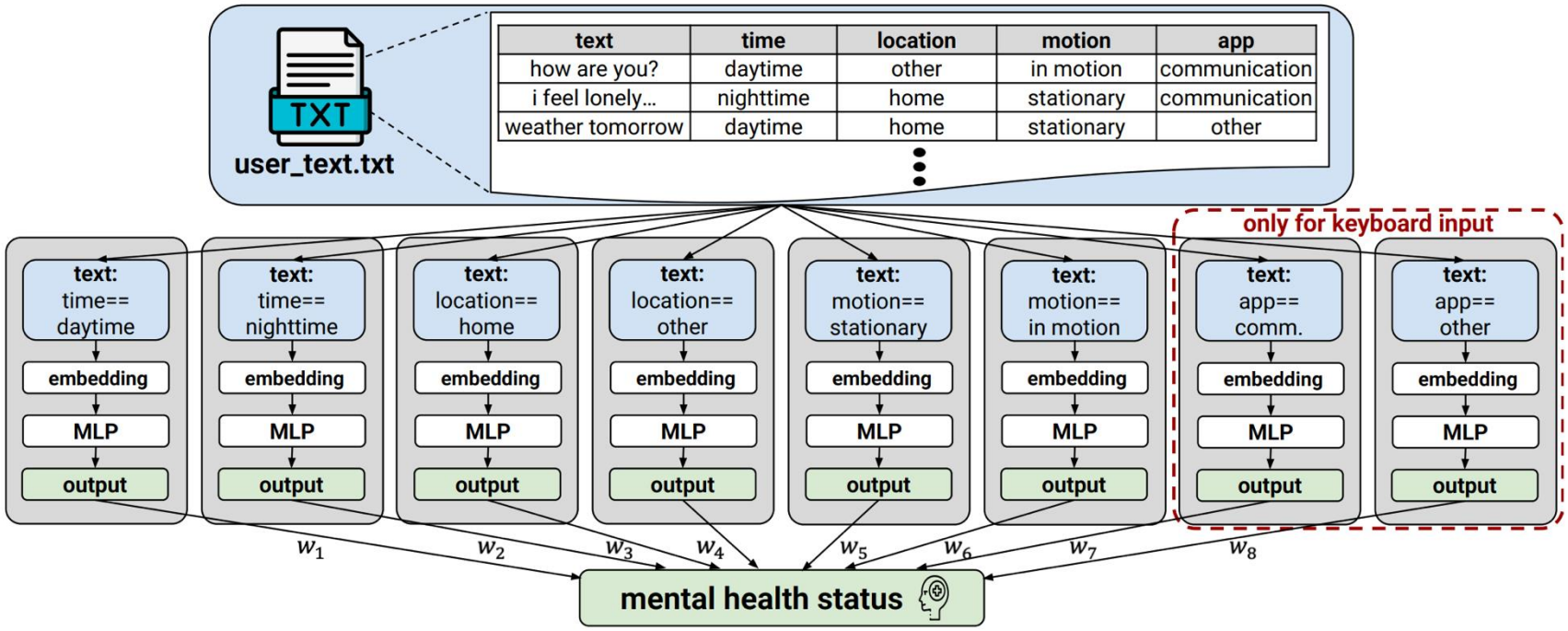
scientific reports



Class	Precision-Recall AUC
Normal	0.91
Depression	0.89
Anxiety	0.87
Bipolar Disorder	0.88
Schizophrenia	0.86
Average	0.88

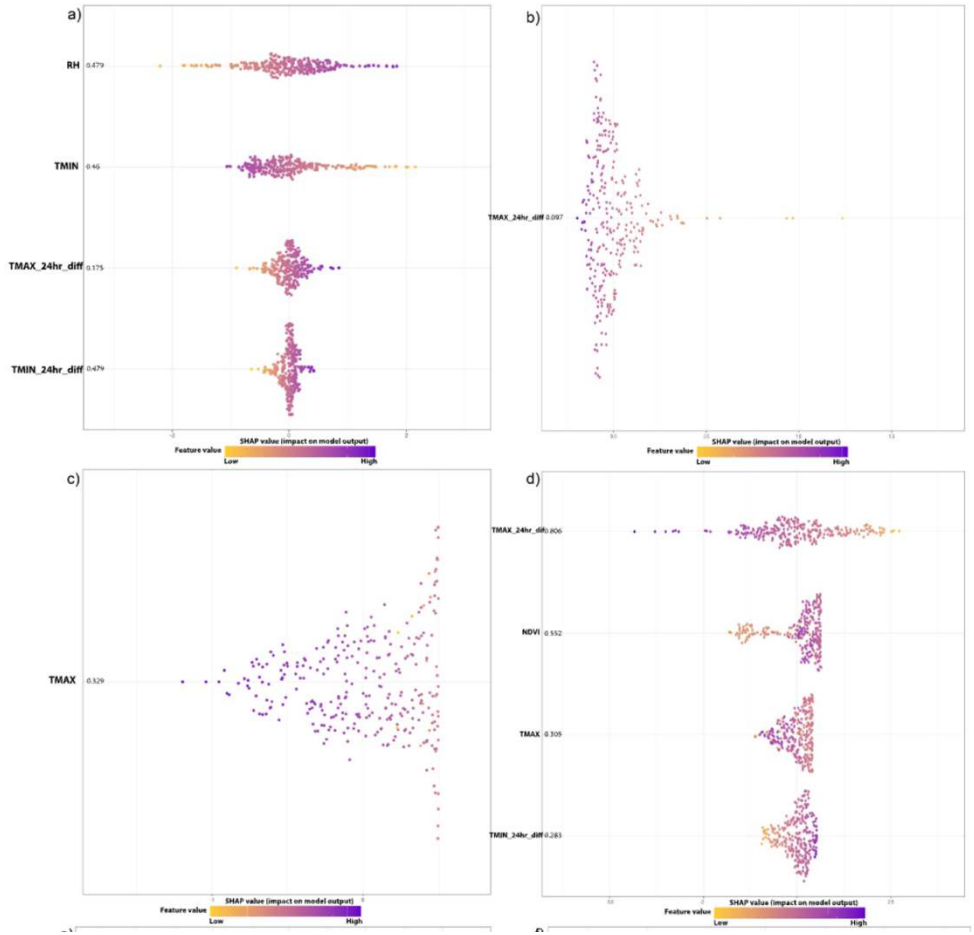
FedTherapist: Mental Health Monitoring with User-Generated Linguistic Expressions on Smartphones via Federated Learning

Jaemin Shin[1], Hyungjun Yoon[1], Seungjoo Lee[1], Sungjoon Park[2], Yunxin Liu[3], Jinho D. Choi[4], Sung-Ju Lee[1]



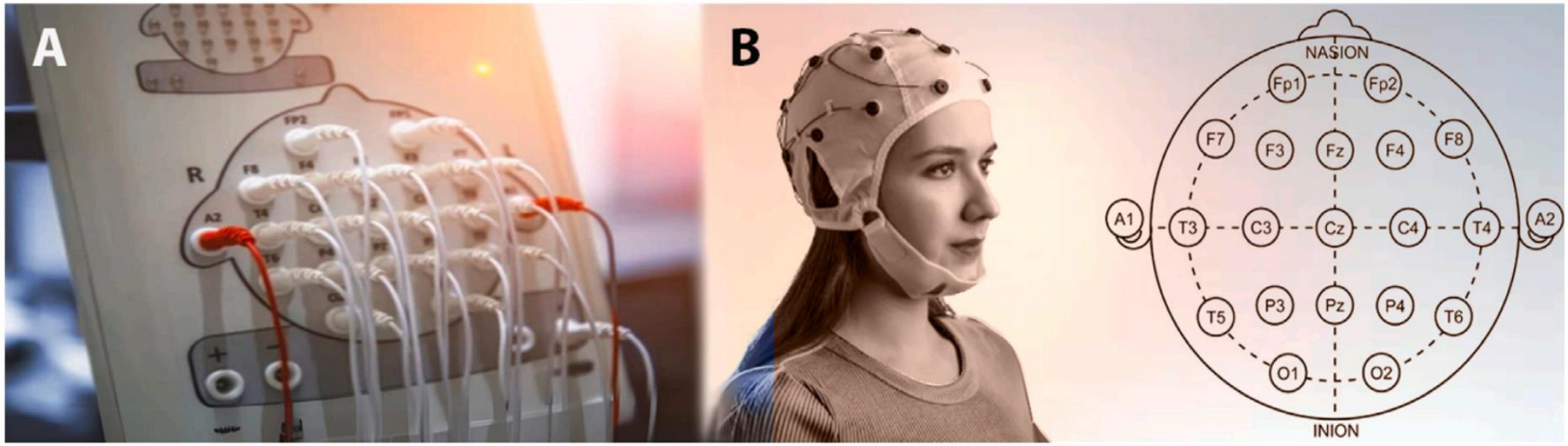
Socio-Environmental Determinants of Mental and Behavioral Disorders in Youth: A Machine Learning Approach

Luke Wertis[1], Margaret M. Sugg[1], Jennifer D. Runkle[2], and Douglas Rao[2]



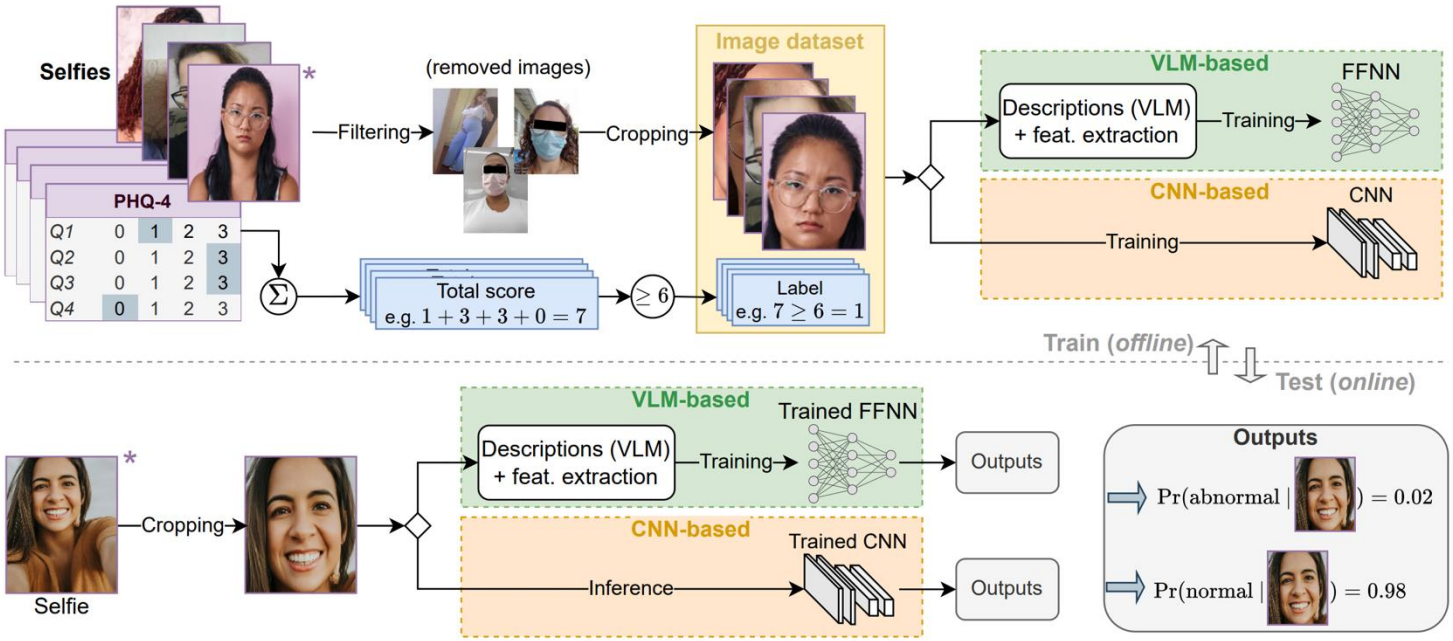
Response prediction of antidepressants: Using graph theory tools for brain network connectivity analysis

Akbar Davoodi [a,*], Martin Holeňa [a], Martin Brunovský [b,c], Aditi Kathpalia [a,d], Jaroslav Hlinka [a], Martin Bareš [b,c], Milan Paluš [a]



AI-Driven Early Mental Health Screening: Analyzing Selfies of Pregnant Women

Gustavo A. Basílio[1], Thiago B. Pereira[1], Alessandro L. Koerich[2], Hermano Tavares[3], Ludmila Dias[1], Maria das Graças da S. Teixeira[4], Rafael T. Sousa[5], Wilian H. Hisatugu[4], Amanda S. Mota[3], Anilton S. Garcia[6], Marco Aurélio K. Galletta[7], and Thiago M. Paixão[1]



* *IA-generated image for illustration purpose.*

Leveraging LLMs to Predict Affective States via Smartphone Sensor Features

Tianyi Zhang, Songyan Teng, Hong Jia, Simon D'Alfonso
{t.zhang59,stteng}@student.unimelb.edu.au,{hong.jia,dalfonso}@unimelb.edu.au
University of Melbourne
Australia

Item	Feature	Sensor
1	the count of episodes using Email applications	Applications
2	the count of episodes using all applications	
3	the count of episodes using social media applications	
4	the count of episodes using dating applications	
5	the count of episodes using social applications	
6	the count of episodes using entertainment applications	
7	the count of episodes using Facebook Moments	
8	the count of usage episodes for the application with top usage	
9	the count of episodes using YouTube	
10	the count of episodes using Twitter	
11	the duration of using Email	
12	the duration of using all applications	
13	the duration using social media applications	
14	the duration of using dating applications	
15	the duration of using social applications	
16	the duration of using entertainment applications	
17	the duration of using Facebook Moments	
18	the duration of using the application with top usage	
19	the duration of using YouTube	
20	the duration of using Twitter	
21	the count of discharging episodes	Battery
22	the count of charging episodes	
23	the number of missed calls	Calls
24	the number of distinct contacts associated with missed calls	
25	the number of missed calls from the most frequent contact	
26	the number of incoming calls	
27	the number of distinct contacts associated with incoming calls	
28	the mean of incoming call duration	
29	the sum of incoming call duration	
30	the mode of incoming call duration	
31	the count of incoming calls from the most frequent contact	
32	the number of outgoing calls	
33	the number of distinct contacts associated with outgoing calls	
34	the mean of outgoing call duration	
35	the sum of outgoing call duration	
36	the mode of outgoing call duration	
37	the count of outgoing calls from the most frequent contact	
38	the count of typing events	Keyboard
39	the count of keyboard typing or swiping event where the length changes exactly one more character	
40	the count of keyboard typing or swiping event where the length changes less than one fewer character	
41	the count of keyboard typing or swiping event where the length changes exactly one fewer character	
42	the number of characters in average in a session	
43	the number of typing sessions	
44	the average time between keystrokes	
45	time spent at the second most visited location	Locations
46	maximum time spent at any location cluster	
47	the ratio of time spent moving between locations to time spent stationary at a location	
48	total travelled distance	
49	standard deviation of the time spent at location clusters	
50	time spent at the most visited location	
51	average time spent at location clusters	
52	normalized entropy of location visits	
53	variance of speed during movement between locations	
54	time spent at home	
55	time spent at the third most visited location	
56	the number of transitions between distinct locations	
57	minimum time spent at any location cluster	
58	radius of Gyration (RoG) indicating the area covered	
59	average speed during movement between locations	
60	percent of time considered outliers in location data	
61	the variance of locations visited	
62	the logarithm of the variance of locations visited	
63	the number of significant places visited	
64	the entropy of location visits	
65	the number of received messages from the most frequent contact	Messages
66	the number of received messages	
67	the number of distinct contacts associated with received messages	
68	the number of sent messages to the most frequent contact	
69	the number of sent messages	
70	the number of distinct contacts associated with sent messages	
71	the number of unlock episodes	Screen status
72	total duration of unlock episodes	
73	the length of longest unlock episode	
74	average time of unlock episodes	
75	the length of shortest unlock episode	
76	standard deviation of unlock episodes	
77	time between the first unlock episode and midnight	

Item (True label)	Predicted Score	Reasoning
Active (5)	3	The person had a moderate number of unlocking episodes (average 8 minutes) and travelled a moderate distance (total travelled 4166 meters on 2023-08-08, 9261 meters on 2023-08-10, 17230 meters on 2023-08-12) during the week.
Determined (4)	3	There is no clear evidence to suggest the person was determined during the week.
Attentive (5)	2	The person did not use any social media applications or dating applications, and there were no emails or messages sent or received during this period. This suggests they were not attentive to these aspects of their phone.
Inspired (4)	1	There is no clear evidence to suggest the person was inspired during the week.
Alert (4)	3	The person had a moderate number of unlocking episodes (average 6 minutes across the week) and travelled a moderate distance (total travelled ranging from 4166 meters on 2023-08-08 to 17230 meters on 2023-08-12) which could indicate some alertness.
Upset (3)	1	There is no clear evidence to suggest the person was upset during the week.
Hostile (2)	1	There is no clear evidence to suggest the person was hostile during the week.
Ashamed (1)	1	There is no clear evidence to suggest the person was ashamed during the week.
Nervous (5)	2	There is no clear evidence to suggest the person was nervous during the week. They did not use social media or dating applications, and there were no emails or messages sent or received.
Afraid (4)	1	There is no clear evidence to suggest the person was afraid during the week.

Table 2: Zero-Shot Chain-of-Thought reasoning for one participant.

Item (True label)	Predicted Score	Reasoning
Active (5)	4	The student's average time spent moving between locations is consistently around 10 km/h, and they have a significant amount of total traveled meters (often over 10,000 meters). These indicate a generally active lifestyle. While their future week involves slightly less movement than usual, it is still above average, so I predict a score of 4.
Determined (4)	4	The student frequently uses applications for extended periods. This suggests a consistent engagement with tasks, potentially indicating a sense of determination. The future week involves similar levels of application usage, making a score of 4 appropriate.
Attentive (5)	4	The student's average time spent at the most visited location is very high, often exceeding 1000 minutes. This could imply a focus on tasks or activities within this location. This pattern persists in the future week. Thus, a score of 4 is likely.
Inspired (4)	4	While there isn't direct evidence in the provided data, the student's frequent use of applications, particularly for extended periods, suggests a potential engagement with learning or creative pursuits. The future week maintains this trend, so a score of 4 seems fitting.
Alert (4)	3	The student has a relatively low number of unlock episodes per day and a low average duration of unlock episodes. This could indicate a less frequent need to access the phone, suggesting a possible lower level of alertness. The future week maintains this pattern, so I predict a score of 3.
Upset (3)	3	The student often experiences missed calls, suggesting potential disruptions and stress. In the future week, there is a small increase in missed calls, so I predict a score of 3.
Hostile (2)	3	There is no clear indication of hostile behaviors based on the provided data. The student has consistent low numbers of missed and outgoing calls, suggesting a general calmness. However, the future week sees a small uptick in missed calls, potentially signaling a slight increase in tension, making a score of 3 more likely.
Ashamed (1)	2	No clear indicators of shame can be found in the provided data. The student's low levels of social media and dating app usage suggest a possible preference for privacy or a lack of social engagement. I predict a score of 2 for the future week.
Nervous (5)	4	The student often shows a high number of typing events per session. This could indicate a possible nervousness or anxiousness, particularly if there are many events where the length of the text changes in fewer characters, suggesting possible backspacing or edits. The future week reflects this pattern, so a score of 4 seems appropriate.
Afraid (4)	3	The provided data lacks clear indicators of fear. The student's consistent low numbers of missed and outgoing calls suggest a general feeling of safety and security. However, the future week features a slight increase in missed calls, which might indicate a potential for feeling slightly more vulnerable or anxious, leading to a score of 3.

Table 3: Ten-Shot Chain-of-Thought reasoning for one participant

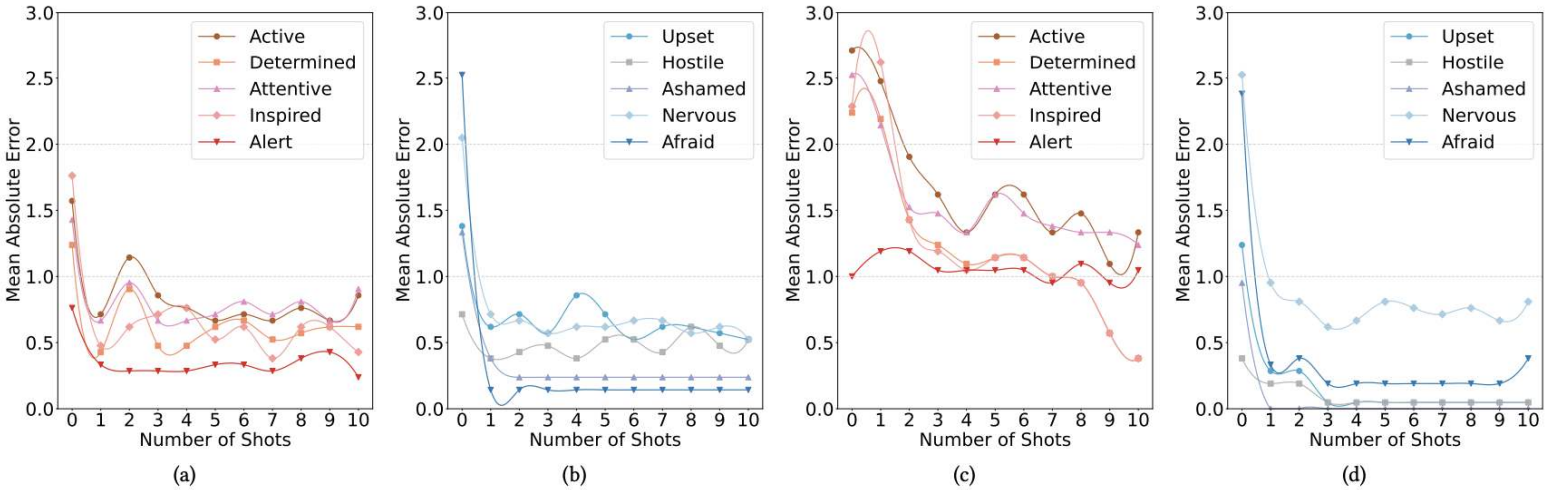


Figure 3: Learning curve for data from two participants

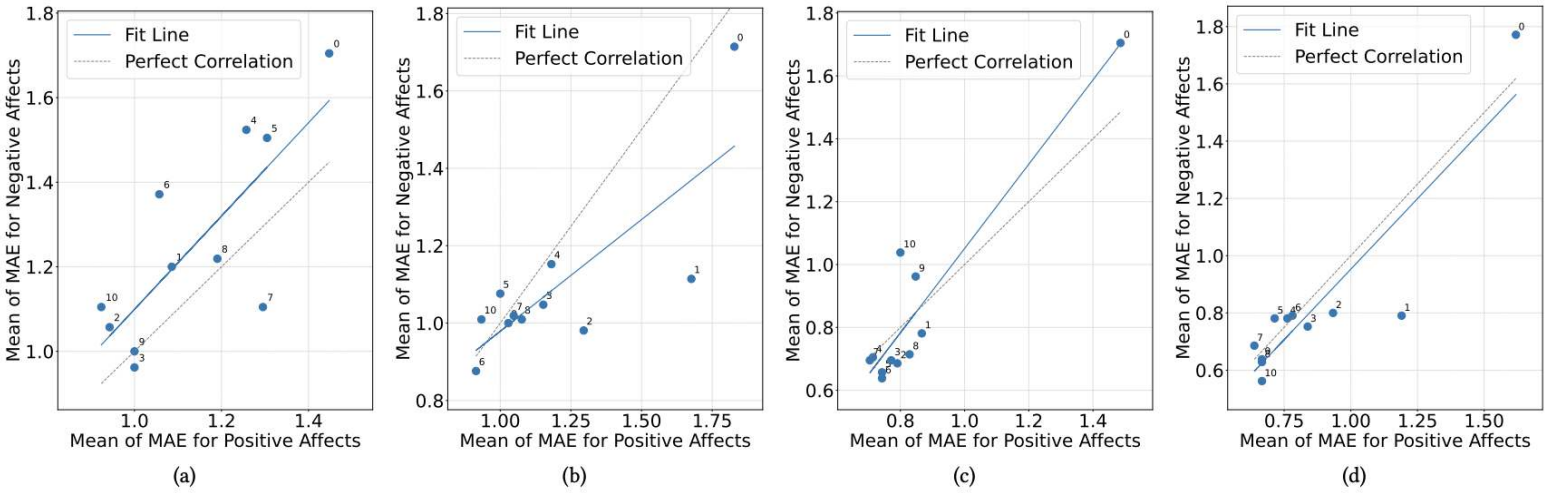


Figure 4: Four participant examples of the linear relationship between MAEs for predicting positive and negative affects.

Alert	High screen unlocks, high activity levels	Suggests high engagement with environment and awareness
Inspired	Increased messaging and calls	Reflects social engagement and a drive to act
Determined	Consistent typing and mobility	Indicates focus and purposeful behavior
Afraid	High ambient noise, erratic screen usage	Suggests a possibly unstable or uncomfortable environment
Nervous	Increased backspacing, frequent short typing bursts	Reflects hesitancy or lack of confidence
Upset	Many missed calls, reduced mobility	May indicate withdrawal or avoidance behavior
Hostile	Abrupt typing patterns, fewer social interactions	Suggests irritability or frustration



Ben_Reis@harvard.edu

dhip
Boston Children's Hospital
HARVARD MEDICAL SCHOOL