Data Science Approaches to Assess Suicide Risk in Individuals, Populations, & Communities: Current Practices, Opportunities, and Risks

The National Academies of Sciences, Engineering, and Medicine

John Pestian, Ph.D. MBA
Professor, Pediatrics, Psychiatry & Biomedical Informatics
Cincinnati Children's Hospital Medical Center, University of Cincinnati
Joint Faculty Oak Ridge National Laboratory
john.pestian@cchmc.org
http://pestianlab.cchmc.org

Cincinnati Children's
changing the outcome together

University of CINCINNATI

About Us

Facts and Figures 2015

Cincinnati Children's at a Glance

- Full-service, nonprofit pediatric academic medical center, established in 1883
- Our vision: to be the leader in improving child health
- 629 registered beds, including 95 inpatient behavioral-health beds and 36 residential behavioral-health beds
- Comprises the Department of Pediatrics of the University of Cincinnati College of Medicine
- #3 in pediatric research grants from the National Institutes of Health
- 850 volunteers contributed 90,263 hours at our facilities, an increase of 11% from the previous fiscal year

National Recognition

- Ranked #3 in the nation among all Honor Roll hospitals in *U.S. News & World Report's* 2015-16 Best Children's Hospitals list, including top 10 for all 10 pediatric specialties analyzed.
- Ranked #1 for pulmonology by *U.S. News* for the first time, and earned top-5 honors in cancer, nephrology, gastroenterology and GI surgery, orthopaedics and urology.
- Earned Magnet-designation (2013-2017) by the American Nurses Credentialing Center. Of the nearly 6,000 hospitals in the U.S., fewer than 7% have earned Magnet recognition.

Statistical Highlights (JULY 1, 2014 – JUNE 30, 2015)

Admissions (includes short stay)	33,692
Emergency Department Visits	102,557
Outpatient Visits	
Primary Care	84,711
Specialty	852,899
Test Referral Center	76,486
Urgent Care	63,481
Surgical Procedures	
Inpatient	6,529
Outpatient	25,810
Surgical Hours	44,031
Patient Encounters	1,246,447

Faculty and Staff

Faculty	
Pediatrics	724
Surgery	103
Radiology	50
Anesthesia	59
Clinical Fellows	262
Research Postdoctoral Fellows	162
Residents	211
Active Medical Staff	1,789
Total Employees	15,260

Operating Revenues and Expenses (dollars in thousands)

Net Hospital Patient Services Revenue	$1,527,936
Research Grants and Contracts	$198,052
Total Operating Revenues	$2,206,470
Total Operating Expenses	$1,996,846
Available to Reinvest in the Mission	$209,624

Cincinnati Children's

Children's changing the outcome together | University of CINCINNATI

The importance of early identification

- Half of all psychiatric disorders that begin by 18 years of age are usually preceded by subsyndromal pathologies that evolve into a major psychiatric disorder.
  - These disorders account for 45% of the global burden of disease for individuals under 25 years 4.
- The cost of doing nothing is staggering—the WHO reports that mental illnesses are the leading cause of disability-adjusted life years (DALYs) worldwide, and $2.5 trillion global with a projected increase to more than $6T by 2030.
- We have done nearly fifteen studies, today I'll review two – suicide corpus construction and NLP in the emergency department World Health Organization.
- "Global health estimates: leading causes of DALYs." *World Health Organization. Retrieved* 23 (2021).

Definition: One DALY represents the loss of the equivalent of one year of full health. DALYs for a disease or health condition are the sum of the years of life lost to due to premature mortality (YLLs) and the years lived with a disability (YLDs) due to prevalent cases of the disease or health condition in a population.

Cincinnati Children's changing the outcome together | University of CINCINNATI

The complexity of suicide

- “Suicide is a complex public health problem of global importance.
- Suicidal behavior differs between sexes, age groups, geographic regions, and sociopolitical settings, and variably associates with different risk factors, suggesting etiological heterogeneity.
- Although there is no effective algorithm to predict suicide in clinical practice, improved recognition and understanding of clinical, psychological, sociological, and biological factors might help the detection of high-risk individuals and assist in treatment selection.”
- DOI: 10.1016/S0140-6736(15)00234-2

Some requirements for dealing with this complexity

3-D AI Representation of EHR
Estimated combinatorial size ---

High-performance computing for training models

Highly training multi-dimensional team

Early identification can reduce impulsive suicide

You think you are so wonderful. This is for you. I took the pills for you. She means so much to you.

Mom its all wrong. I took 14 of your pills. They Mother I am so tired. Please forgive me. I can't decide anything any more. Daddy* please hurry home. Mom I'm so tired. Please forgive me for everything.

Source: Pestian, Shneidman, Leenaars Suicide Note Database

Sentiment corpus construction, annotation, and shared task

- Use notes to create a corpus for linguistic models
- Collect 1,319 notes written by people before they died by suicide.
  - 1950 - 2020 by Drs. Shneidman, Leenaars, and John Pestian
- Corpus construction began in 2009
  - CCHMC's Institutional Review Board (#2009-0664).
  - Each note was scanned into the Suicide Note Module (SNM)
  - The scanned notes were then transcribed to a text-based version by a professional transcriptionist.
  - Each note was then reviewed for errors by three separate reviewers.
- Hold a competition to develop machine learning algorithms

- DOI: 10.4137/BII.S10213

Cincinnati Children's changing the outcome together | University of CINCINNATI

Annotations read and added emotions

- Annotators were asked to identify the following emotions:
  - abuse, anger, blame, fear, guilt, hopelessness, sorrow, forgiveness, happiness, peacefulness, hopefulness, love, pride, thankfulness, instructions, and information.
  - Emotion tokens were provided from Pubmed abstracts
    - Delphi analysis to rank
  - A special web-based tool was used to collect, monitor, and arbitrate annotator's activity. The tool collects annotation at the word and sentence level.

Cincinnati Children's changing the outcome together | University of Cincinnati

Most frequent sentiments found in suicide notes

Table 3

Frequency and example of assigned emotions.

Description	Frequency	Example
Instructions	609	Careful, cyanide gas in the bathroom
Hopelessness	601	I just didn't want to live anymore
Love	472	I love her
Information	430	I have no debts except for what my wife knows
Guilt	423	Forgive me please
Sorrow	342	Oh, how I suffer
Blame	235	I have been pushed around too much
Hopefulness	216	You will a happy and healthy life
Thankfulness	187	You, John have been so good to me and Jane
Anger	183	Well, Jane I hope this makes you happy!
Fear	154	I am terrified
Happiness/ peacefulness	119	I'm ready for the next step with joy and anticipation
Pride	89	We have another sweet little daughter
Forgiveness	61	I do not blame you for anything, my dear
Abuse	53	Life is so cruel when you are persecuted by in-laws and ex-wife

Pestian, John P., et al. "Sentiment analysis of suicide notes: A shared task." *Biomedical informatics insights* 5 (2012): BII-S9042.

Cincinnati Children's changing the outcome together | University of CINCINNATI

Shared task

- 35 International teams participated
  - Used the corpus to train and test their models for finding

Table 4. Team ranking using micro-average F_1, precision and recall.

Team	F_1	Precision	Recall
Open university	0.61390	0.58210	0.64937
MSRA	0.58990	0.55915	0.62421
Mayo	0.56404	0.57085	0.55739
Nrciit	0.55216	0.55725	0.54717
Oslo	0.54356	0.60580	0.49292
Limsi	0.53831	0.53810	0.53852
Swatmrc	0.53429	0.57890	0.49607
UMAN	0.53367	0.56614	0.50472
Cardiff	0.53339	0.54962	0.51808
LT3	0.53307	0.54374	0.52280
UTD	0.51589	0.55089	0.48506
OHSU	0.50985	0.53351	0.48821
Wolverine	0.50315	0.45334	0.56525
TPAVACOE	0.50234	0.49922	0.50550
CLiPS	0.50183	0.51889	0.48585
SIP	0.49727	0.67429	0.39387
SRI & UC Davis	0.48003	0.49831	0.46305
DIEGO-ASU	0.47506	0.41791	0.55031
Ebi	0.45636	0.60077	0.36792
Duluth	0.45269	0.45985	0.44575
Columbia	0.43017	0.42125	0.43947
Pxs697	0.40288	0.37192	0.43947
Lassa	0.38194	0.35089	0.41903
Saeed	0.37927	0.37059	0.38836
SNAPS	0.35294	0.58684	0.25236
Senti6	0.29669	0.30532	0.28852

Registered team locations

Table 5. Examples of sentence/label combinations that were misclassified by all systems.

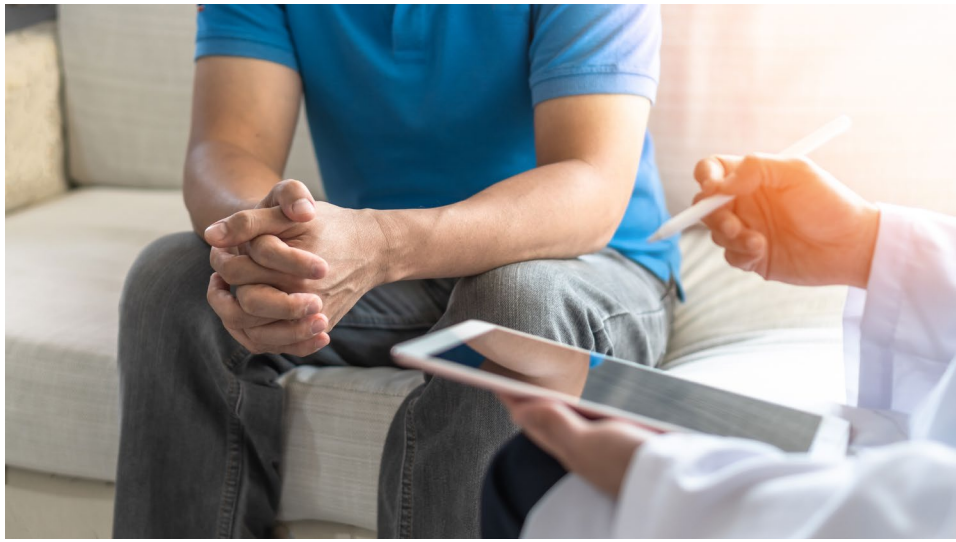
Error type	Text ID	Sentence	Annotator	System
False negative	200909031138 4664	"Goodbye my dear wife Jane."	love	none
False negative	200809091809 2119	"I ask God alone to judge my action."	guilt	none
False negative	200812181837 2227	"I hope something is done to John Johnson, for I do not wish to die in vain."	anger	none
False positive	200908201415 0445	"respectfully Mary P.S. I love you BABY."	none	love
False positive	200812181838 1506	"Dearest Jane I am about to commit suicide. x Please notify police that I am in the deserted garage at the top of Terrace in Cincinnati near the rose bowl."	none	instructions
False positive	200809091735 1923	"John: I can't take your cruel unkind treatment any longer."	none	hopelessness

doi.org/10.4137/BII.S9042

Cincinnati Children's changing the outcome together | University of CINCINNATI

Next study: Linguistic and Acoustic Clinical Trial

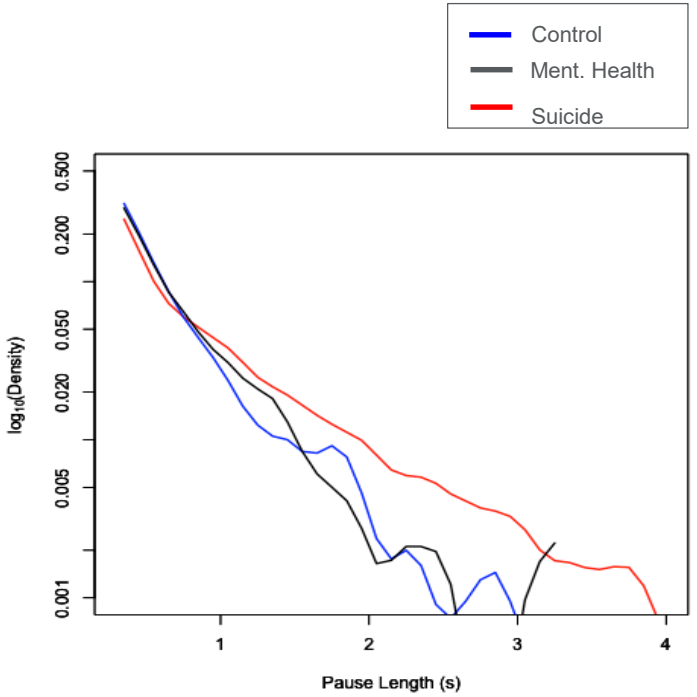
- Enroll patients from ED
- collect linguistic, acoustic, video data from EHR
  - Interview with semi-structured ubiquitous questions to extract language for machine learning
    - Administered by clinicians during ED visit
      - *Do you have hope?*
      - *Do you have any fear?*
      - *Do you have any secrets?*
      - *Are you angry?*
      - *Where does it hurt emotionally?*
- Doi: 10.1111/sltb.12180



Cincinnati Children's changing the outcome together | University of Cincinnati

Audio & Facial Features

- Pauses Longer
- Smaller vowel spacing
- Smile less
- Gaze down



- **DOI:** 10.1109/ICASSP.2015.7178880
- **DOI:** 10.1109/FG.2017.96

Cincinnati Children's changing the outcome together | University of CINCINNATI

Results: machine learning linguistics

Table 1: The AROC for the machine learning algorithm. The non-suicidal group comprises of either mentally ill and control subjects. Class performances are shown for adolescents, adults, and the combined adolescent and adult cohorts.

	Suicidal vs. Controls			Suicidal vs. Mentally Ill			Suicidal vs. Mentally Ill and Controls		
	Adolescents ROC (SD)	Adults ROC (SD)	Adolescents+Adults ROC (SD)	Adolescents ROC (SD)	Adults ROC (SD)	Adolescents+Adults ROC (SD)	Adolescents ROC (SD)	Adults ROC (SD)	Adolescents+Adults ROC (SD)
Linguistics	0.87 (0.04)	0.91 (0.02)	0.93 (0.02)	0.82 (0.05)	0.77 (0.04)	0.79 (0.03)	0.82 (0.04)	0.84 (0.03)	0.87 (0.02)
Acoustics	0.74 (0.05)	0.82 (0.03)	0.79 (0.03)	0.69 (0.06)	0.74 (0.04)	0.76 (0.03)	0.74 (0.05)	0.80 (0.03)	0.76 (0.03)
Linguistics + Acoustics	0.83 (0.05)	0.93 (0.02)	0.92 (0.02)	0.80 (0.05)	0.77 (0.04)	0.82 (0.03)	0.81 (0.04)	0.84 (0.03)	0.87 (0.02)

- Next Steps: Spreading Activation Mobile (SAM)
- Real Time Demo

Cincinnati Children's changing the outcome together | University of CINCINNATI

Future approaches to data collection

Steering
Purpose: Oversight, Scientific agenda, direction, project selection

Ethical, Regulatory, and Management
Purpose: Ethical, legal, management, and social principles of using AI/machine learning (ML) in biomedicine. Project management

Data Engineering
Purpose: Adopt best practices for AI data management. Limit risks and maximize opportunities by following WHO principles for AI data collection and governance.

Algorithmic Development
Purpose: Build and provide software, hardware, and firmware tools for the biomedical and behavioral research community

Translation to Care
Purpose: Implement AI tools that improve our ability to identify, understand, and help individuals experiencing and suffering from mental illness

Workforce Education
Purpose: Provide training and education for teaming, ethics, standards, tools development, and data acquisition

Data for suicide holistic measures

Clinical
Environmental
Social Determinant
Census
Pubmed
Social media

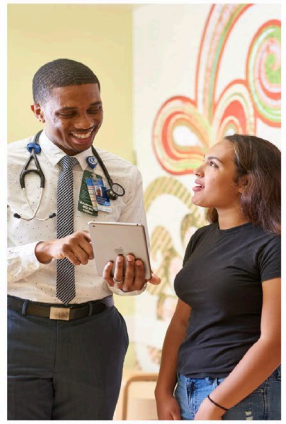
.

.

Cincinnati Children's
changing the outcome together

University of CINCINNATI

Clinical and High-Speed Computing Collaboration



Being in the right place at the right time, with the right tools, and delivering the right care

Scientific Goals

- Create a multimodal data repository for computing trajectories.
- Create methods for extracting all scientifically validated indicators of mental illness and use machine learning to discover new ones.
- Create and validate trajectories of depression, anxiety, and suicide.
- Determine how to operationalize the results to clinical settings.

Administrative Goals

- Establish the framework for data collection and sharing.
- Establish the necessary regulatory and legal framework.
- Establish data-sharing partnerships with other regional stakeholders.
- Establish new strategic partnerships collaboration.

Cincinnati Children's
changing the outcome together

Developing Mental Health Trajectories Using High-Performance Computing

A Collaboration Between Cincinnati Children's Hospital Medical Center and the United States Department of Energy's Oak Ridge National Laboratory

See back-side for citation

50% of adult mental illness begins by **age 14**

75% by **age 24**

On average, treatment is started **11 years** after symptoms start

ONLY 51% of youth (6-17 yrs) were treated for their mental illness in any given year

70% of youth in the juvenile justice system have a diagnosed mental illness

1 in 8 emergency room visits involve adult mental illness and substance abuse

8.4 million people in the U.S. provide care to the mentally ill

25% of the homeless have a serious mental illness

Depression and anxiety disorders cost the global economy **$1 trillion** each year in lost productivity

What inspired this initiative?

Innovations in mental health treatment that use advanced technology lag the pace of innovations in medical care and basic science. We want to use advances in artificial intelligence to help identify children with mental illness earlier and provide families and caregivers with precision information that can be useful for treatment.

What are we hoping to accomplish?

Similar to pediatric growth charts that measure and plot trajectories of growth in height, weight, body mass, and head circumference, we will predict patient trajectories of anxiety, depression, and suicide prevention that are informed by genetic traits, current physical and mental health state, social determinants of health, and environmental exposure to known indicators of risk.

Through our unique national collaboration, we will use our clinical and computational expertise to develop near real-time visualizations of patient-specific mental health trajectories that enable earlier identification of patients with clinically-elevated symptoms and connect them to services, leading to better outcomes and quality of life.

By combining information that is regularly collected during healthcare encounters, we can use advanced artificial intelligence and machine learning methods to identify children at highest risk of developing a mental health concern over time. The first three conditions of focus are depression, anxiety, and suicidal ideation.

For more information, contact John Pestian, PhD, MBA (john.pestian@cchmc.org) or Tracy Glauser, MD (tracy.glauser@cchmc.org)

TEAM

Principal Investigators
Tracy A. Glauser, MD
John P. Pestian, PhD, MBA

Oak Ridge National Laboratory
Greeshma Agasthya, PhD
Blair Christian, PhD
Katherine Evans, PhD
Daniel A. Jacobson, PhD
Anuj Kapadia, PhD
Joe Lake, MS

Co-Investigators
Bob Ammerman, PhD
Bruce Aronow, PhD
Drew Barzman, MD
Sarah Beal, PhD
Andrew Beck, MD, MPH
Cole Brokamp, PhD
Lee Ann Conard, MD
Kevin Dufendach, MD
Craig Erickson, MD
Stephanie Loos, MS
Christopher D. King, PhD
Tesfaye Mersha, PhD
Jareen Meinzen-Derr, PhD
Alexis Mitulpunkt, PhD
Yizhao Ni, PhD
Ernie Pedapati, MD
Scott Powers, PhD
Laura Ramsey, PhD
Dan Santel, PhD
Mayur Sarangdhar, PhD
Robert Shapiro, MD
Andy Spooner, MD
Michael Sorter, MD
Lori Stark, PhD
Jeffrey Strawn, MD
Rhonda Szczesniak, PhD
Sander Vinks, Pharm.D, PhD
Cynthia Zion

Biomedical Informatics
Ron Bryson, MA
Jay Gholap, MS
Stephanie Loos, MS
Adriana Navarro Sainz, MDes
Jeremy Nix
Natalie Reno, MS
Prakash Velayutham, MS
Janet Zahner

Infographic of Statistics Abstracted from *Mental Health By the Numbers* by the National Alliance on Mental Illness (NAMI), Online: https://www.nami.org/mhstats; Copyright 2021 NAMI based on studies conducted by organizations like Substance Abuse and Mental Health Services Administration (SAMHSA), Centers for Disease Control and Prevention (CDC), and the U.S. Department of Justice.

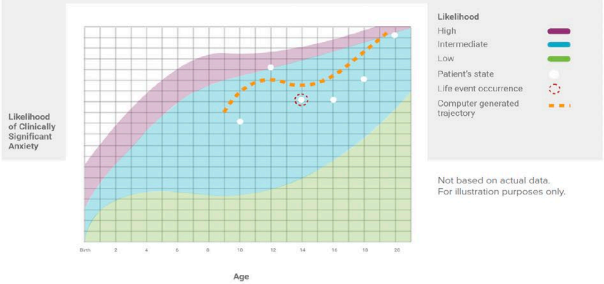
BRV424008

A Multidisciplinary Collaboration

To develop a computation model for early detection of anxiety, depression, and suicide, Cincinnati Children's has partnered with Oak Ridge National Laboratories—the U.S. Department of Energy's largest science and engineering lab and home to the nation's most powerful supercomputer. From its roots in the Manhattan Project, Oak Ridge National Laboratory's mission continues to evolve at the frontiers of science.

Most recently, Oak Ridge National Laboratory has dedicated substantial resources to health sciences biomedical research and has partnered with governmental and non-governmental agencies like Cincinnati Children's, who, as collaborators, can impact the nation's health by using advanced computational tools and methods. Cincinnati Children's and Oak Ridge National Lab have also established a Diversity Committee of clinicians and scientists who will ensure that the scientific approach, selection of data sources, and interpretation of findings reflect the diversity of experiences for children in our community and leads to improvements in care delivery and health for all youth, regardless of age, race, ethnicity, gender, sexual orientation, or economic background.

Patient's Anxiety Trajectory



What do we need to accomplish our objectives?

Amazing opportunities exist to gain insights into children's mental health trajectories using existing clinical data from the medical record. The lower costs of accumulating and accessing biological and clinical data, the availability of diverse datasets, and advances in artificial intelligence-based algorithms have ushered in a new era of research specifically focused on identifying children and families at high risk for mental health concerns, with opportunities for early intervention, prevention, and treatment.

While we have experience in trajectory computing, this collaboration allows us to scale-up and think about how to help children and families coping with mental health concerns regionally, nationally, and internationally.

Cincinnati Children's
changing the outcome together

ATI

Thank you.

Cincinnati Children's
changing the outcome together

University of CINCINNATI