Improving prediction of adolescent suicide attempts with electronic health records:

Leveraging external data sources of similar patients

Shane J Sacco, PhD
Postdoctoral Research Associate
Department of Statistics

April 14th, 2023

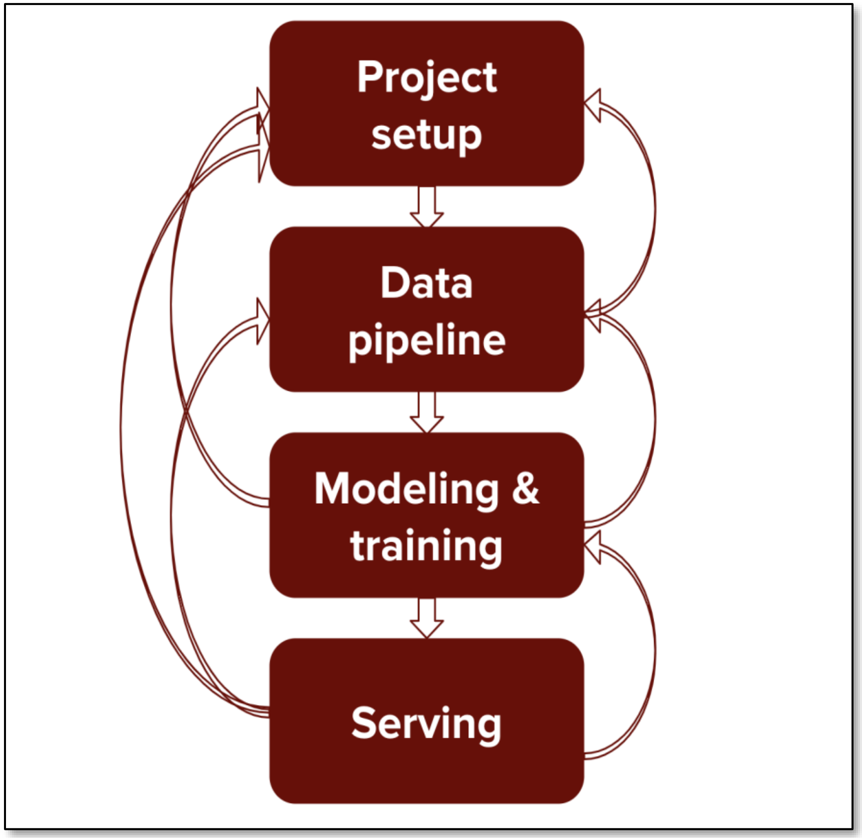
UCONN

Suicide is rising in adolescence

- According to the CDC, suicide rates of 10 to 24 year olds in the US increased 57% between 2007 and 2018[1,2]
- Almost 80% of individuals under 19 years old who died by suicide had contact with healthcare in year prior[3]
- With this, federal mandates have been established to identify and prevent suicide, especially in healthcare[4,5]

Suicide prediction

- One method to identify at-risk patient is creating statistical algorithms
- To date, several algorithms have been published using electronic health records (EHRs)[6-8]
- However algorithms only identify around half of true positive pediatric attempters when setting specificity to 90% (similar to declaring the 10% of patients as high risk)



Suicide prediction

Study	Cohort	Model	AUC	Sensitivity at 90% Specificity
Su et al (2020)[6]	10-18 patients in pediatric hospital	Logistic regression	0.84-0.86	0.53-0.65
Xu et al (2021)[7]	10-24 inpatients in CT	Logistic regression	0.82	0.43
Barak-Corren et al (2020)[8]	Patients in pediatric hospital	naïve Bayes classifier	0.72	0.37

Improving prediction

Since we can only detect ~50% of cases, we must continue to improve algorithms, but how?

- Machine learning models?
  - E.g, random forest
  - difficult to interpret for clinicians and does not increase information from EHR

- Use clinical notes?
  - E.g., Natural Language Processing
  - Very costly, time-consuming, and complex

- Collect more data like social determinants of health surveys?
  - I.e., patient-reported outcomes
  - Difficult in real-world context, time-consuming, costly, lack of staff compliance

- Leverage data from other patients or individuals?
  - Only if we have enough information to match patients

Improving prediction

Since we can only detect ~50% of cases, we must continue to improve algorithms, but how?

- Machine learning models?
  - E.g, random forest
  - difficult to interpret for clinicians and does not increase information from EHR
- Use clinical notes?
  - E.g., Natural Language Processing
  - Very costly, time-consuming, and complex
- Collect more data like social determinants of health surveys?
  - I.e., patient-reported outcomes
  - Difficult in real-world context, time-consuming, costly, lack of staff compliance
- Leverage data from other patients or individuals?
  - Only if we have enough information to match patients

Today's focus

Leveraging data with fusion

- Using a principle in transfer learning, we may be able to generate additional patient features

- Specifically, using ***data fusion***[9], we may be able to match patients on their known features (i.e., demographics and diagnosis codes) with patients or individuals in other datasets on the same and generate new features

- Other datasets (***external datasets***) may have information about risk not found in the EHR used for prediction (***target datasets***)

Leveraging data with fusion

What kind of risk information could external datasets have?

- **More data.** May contain more data to create better “informed” suicide risk scores when modelling
- **Unique features.** May contain unique risk factors like social support not well-measured in the target data
- **Unique cases.** May contain attempters with diagnosis profiles not found in the target data to learn more about target attempters

The general framework

Now that you have a basic idea of the concept, let’s describe the general framework:

1) Identify an external dataset(s) and feature(s) of interest

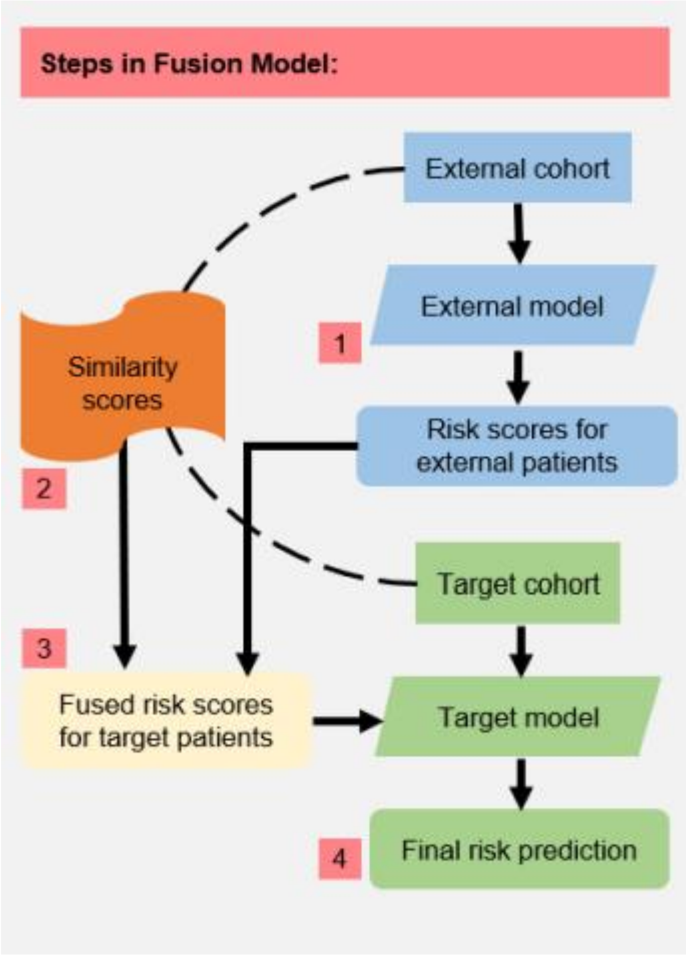
2) Match each patient in both datasets using a similarity metric like Pearson’s *r*

3) Generate new features in the target data by aggregating values from top similar patients in the external data

Time for an example

First study in context of suicide risk: Xu et al (2021)[7]

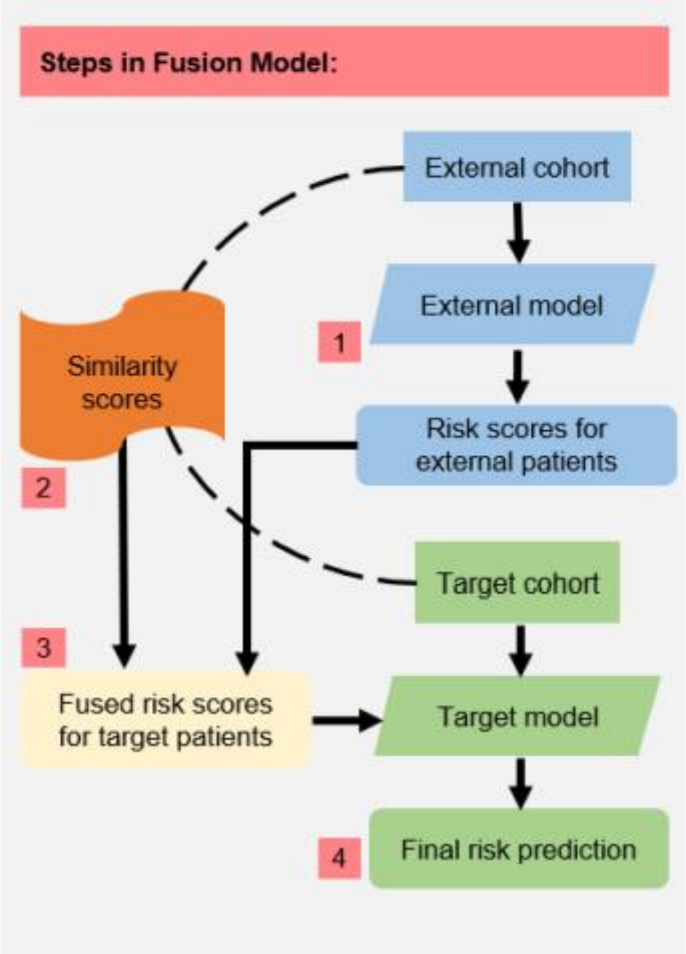
- They sought to improve suicide risk prediction in CT 10-24 pediatric inpatients (the Hospital Inpatient Discharge Database)
- 485 cases and 38806 controls from 2012-2017
- Transferred risk scores from a model using inpatient and outpatient medical claims data (All Claims Paid Database)



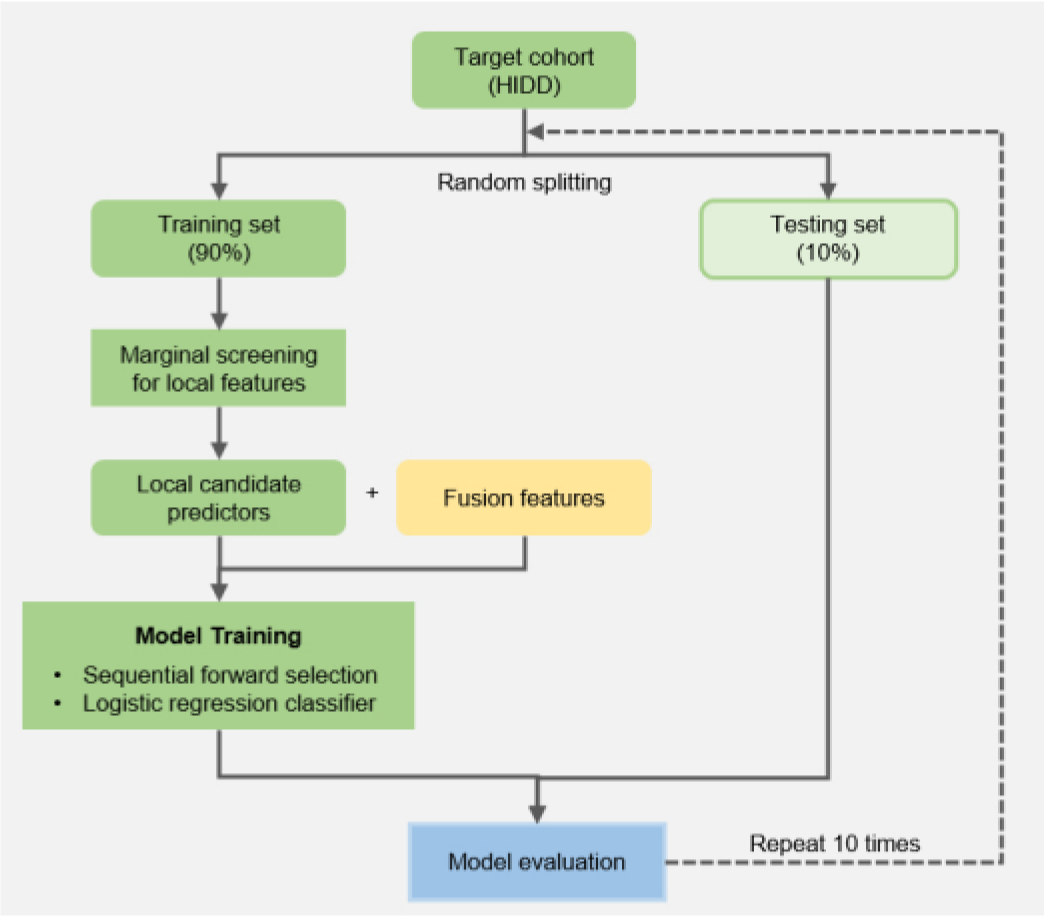
Time for an example

First study in context of suicide risk: Xu et al (2021)[7]

- Ran a logistic model in the external data and extracted patient risk scores
- Matched patients using Pearson's *r*
- Generated new target features by creating a weighted sum of top 1, 10, 20, 50, and 100 most similar patients
- Ran logistic regression in target data including local EHR and fused features



Time for an example



Time for an example

First study in context of suicide risk:
Xu et al (2021)[7]

	HIDD (target cohort)		APCD (external cohort)	
	Case	Control	Case	Control
No. of patients	485	38 806	2053	153 433
Sex, *N* (%)				
Female	308 (63.51)	22 937 (59.11)	1281 (62.40)	76 533 (49.88)
Male	177 (36.49)	15 869 (40.89)	772 (37.60)	76 900 (50.12)
Age group, *N* (%)				
10–14	72 (14.85)	6266 (16.15)	368 (17.92)	46 374 (30.22)
15–19	253 (52.16)	12 798 (32.98)	931 (45.35)	53 184 (34.66)
20–24	160 (32.99)	19 742 (50.87)	754 (36.73)	53 875 (35.11)

Time for an example

First study in context of suicide risk:
Xu et al (2021)[7]

Models	**AUC (95% CI)**	**Sensitivity (*SD*)**	
		95% specificity	**90% specificity**
Conventional	0.82 (0.81, 0.84)	0.24 (0.07)	0.42 (0.07)
Fusion	0.86 (0.84, 0.89)	0.65 (0.02)	0.70 (0.03)

Time for an example

First study in context of suicide risk:
Xu et al (2021)[7]

- They found a substantial improvement in prediction
  - Likely provided outpatient information about target patients
  - Risk information about many other patients

- Coefficients of model predictors refined by fused risk scores

- Model of fused risk scores only provided “at-chance” performance

Many more avenues of fusion

- As mentioned earlier, we may leverage different types of external features and data
- In the next slides, I will present case studies including:
  - Some examples of these different data types
  - Various methods for similarity matching

Fusing
Social determinants of health

Social determinants of health

- In case study 1, I present my recent publication, using the same target dataset in Xu et al[7], HIDD

- However, instead of risk scores, we fuse ***social determinants of health (SDOH)*** from a large survey-based study

- SDOH represent a wide number of biopsychosocial factors that describe the circumstances in which we are born and live[10]

Social determinants of health

- SDOH includes concepts such as social support, housing, finances, education, religion, etc.

- Literature suggests SDOH are unique risk factors and protective factors related to suicide risk[11-14]

- In EHR, SDOH are not well-documented and if present, may only approximate many of these concepts[15]

Social determinants of health

- We extracted 23 SDOH from The National Longitudinal Study of Adolescent to Adult Health (Add Health)[16]
  - N cases = 230; N controls = 6271

- For example:
  - Degree to which mother cares and father cares
  - Frequency of hanging out with friends
  - Living in a one-family house
  - Religious orientation
  - Body image
  - Having physical altercations while intoxicated
  - Perception of being killed by age 21

Social determinants of health

- We followed the same framework as Xu et al (2021)[7] but expanded it in a number of ways:

- Fusion variables were generated using Pearson's *r* and Manhattan *d*

- A second set of fusion variables were generated by calculating weighted versions of similarity scores

  – Weights for each diagnosis code in matching derived from a logistic model of suicide risk in Add Health

Social determinants of health

Models	AUC (95% CI)	Sensitivity (95% CI)	
		95% specificity	90% specificity
Conventional	0.82 (0.81, 0.83)	0.28 (0.25, 0.31)	0.44 (0.39, 0.49)
Fusion	0.83 (0.82, 0.84)*	0.31 (0.27, 0.35)*	0.48 (0.43, 0.52)*

*Note. *improvements significant at 95% confidence level*

Social determinants of health

- Including fused SDOH variables improved prediction modestly
  - Totally unique patients
  - Used mock diagnosis codes to match patients as Add Health did not contain these codes directly
  - SDOH may work in-tandem/interact with each other, complex relationships

- Various SDOH with different matching methods appeared in all models and highly predictive in subsequent analyses
  - Provided unique prediction of attempts
    - Mother caring
    - Having no religious orientation

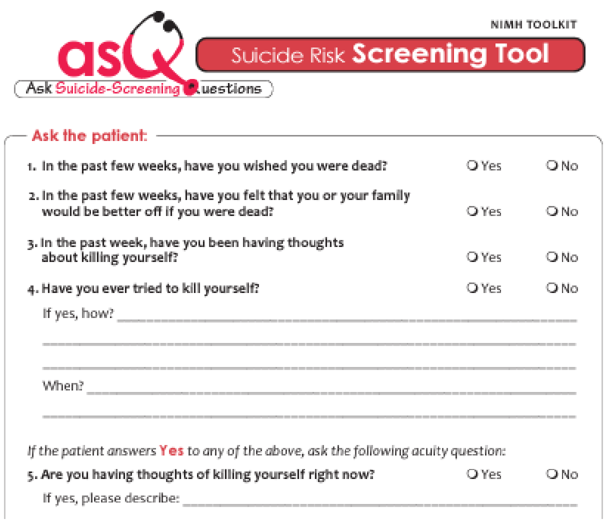
Fusing
Suicide risk screening results

Suicide risk screening

- In case study 2, I present preliminary work predicting suicide attempts in 10-18 patients that were hospitalized and/or seen in the emergency department of a **large urban medical center in CT**
- Here, I acknowledge the importance of suicide risk screening and the potential to fuse screening results from a pediatric hospital which universally screened patients seen in their ED
- As a component of federal mandates[4,5], suicide risk screening is becoming an essential task in healthcare facilities
- Screeners show validity in predicting ideation & suicide attempts[17,18]

Suicide risk screening

- A large pediatric hospital in CT used a pipeline of two screeners:
  - The Ask Suicide-Screening Questions (ASQ) survey[18]
    - 5 items result in negative, non-acute positive, acute positive
  - The Columbia Suicide-Severity Rating Scale (C-SSRS) screen[19]
    - 6 items results in no, low, moderate, high risk



NIMH TOOLKIT

asQ Suicide Risk Screening Tool

Ask Suicide-Screening Questions

Ask the patient:

1. In the past few weeks, have you wished you were dead? ❍ Yes ❍ No
2. In the past few weeks, have you felt that you or your family would be better off if you were dead? ❍ Yes ❍ No
3. In the past week, have you been having thoughts about killing yourself? ❍ Yes ❍ No
4. Have you ever tried to kill yourself? ❍ Yes ❍ No

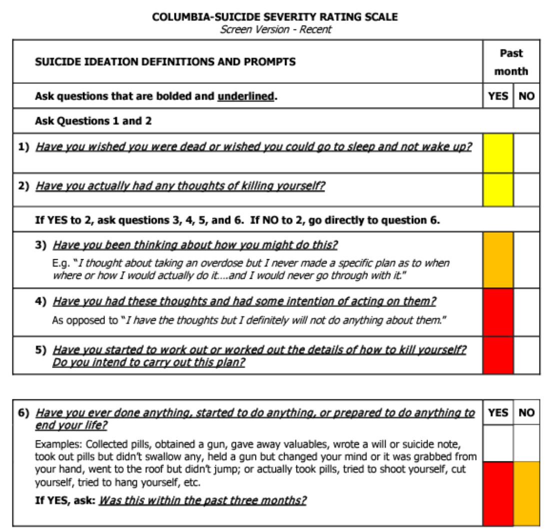
If yes, how? ______

When? ______

If the patient answers ***Yes*** *to any of the above, ask the following acuity question:*

5. Are you having thoughts of killing yourself right now? ❍ Yes ❍ No

If yes, please describe: ______



COLUMBIA-SUICIDE SEVERITY RATING SCALE
Screen Version - Recent

SUICIDE IDEATION DEFINITIONS AND PROMPTS	Past month	
Ask questions that are bolded and underlined.	YES	NO
Ask Questions 1 and 2		
1) *Have you wished you were dead or wished you could go to sleep and not wake up?*		
2) *Have you actually had any thoughts of killing yourself?*		
If YES to 2, ask questions 3, 4, 5, and 6. If NO to 2, go directly to question 6.		
3) *Have you been thinking about how you might do this?* E.g. "*I thought about taking an overdose but I never made a specific plan as to when where or how I would actually do it....and I would never go through with it.*"		
4) *Have you had these thoughts and had some intention of acting on them?* As opposed to "*I have the thoughts but I definitely will not do anything about them.*"		
5) *Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?*		

6) *Have you ever done anything, started to do anything, or prepared to do anything to end your life?* Examples: Collected pills, obtained a gun, gave away valuables, wrote a will or suicide note, took out pills but didn't swallow any, held a gun but changed your mind or it was grabbed from your hand, went to the roof but didn't jump; or actually took pills, tried to shoot yourself, cut yourself, tried to hang yourself, etc. **If YES, ask:** *Was this within the past three months?*	YES	NO

Suicide risk screening

- Target data included 338 cases and 7533 controls between 2012-2017

- External N = 17366 patients screened in the first two years (2019-2021)
  - 2799 Positive screens

- We use similarity matching as in case study 1 but also expand methods to include:
  - Jaccard's distance
  - cosine similarity

Suicide risk screening

Variable	Target		External	
	Cases	Controls	Positive Screen	Negative Screen
No of patients, N (%)	338 (4.29)	7533 (95.71)	2799 (16.12)	14567 (83.88)
Gender, N (%)				
Male	104 (30.77)	3564 (47.31)	845 (30.19)	7483 (51.37)
Female	234 (69.23)	3969 (52.69)	1954 (69.81)	7084 (48.63)
Age group, N (%)				
10-13	95 (28.11)	2877 (38.19)	991 (35.41)	6801 (46.69)
14-18	243 (71.89)	4656 (61.81)	1808 (64.59)	7766 (53.31)
Race/ethnicity, N (%)				
White	160 (47.34)	2672 (35.47)	1423 (50.84)	6201 (42.57)
Black or African American	70 (20.71)	2118 (28.12)	421 (15.04)	2635 (18.09)
Hispanic or Latino	81 (23.96)	2297 (30.49)	646 (23.08)	4251 (29.18)
Other	27 (7.99)	446 (5.92)	309 (11.04)	1480 (10.16)

Suicide risk screening

Models	AUC (95% CI)	Sensitivity (95% CI)	
		95% specificity	90% specificity
Conventional	0.83 (0.79, 0.87)	0.49 (0.42, 0.57)	0.63 (0.55, 0.72)
Fusion	0.84 (0.80, 0.87)	0.54 (0.46, 0.62)*	0.64 (0.56, 0.72)

*Note. *improvements significant at 95% confidence level*

Suicide risk screening

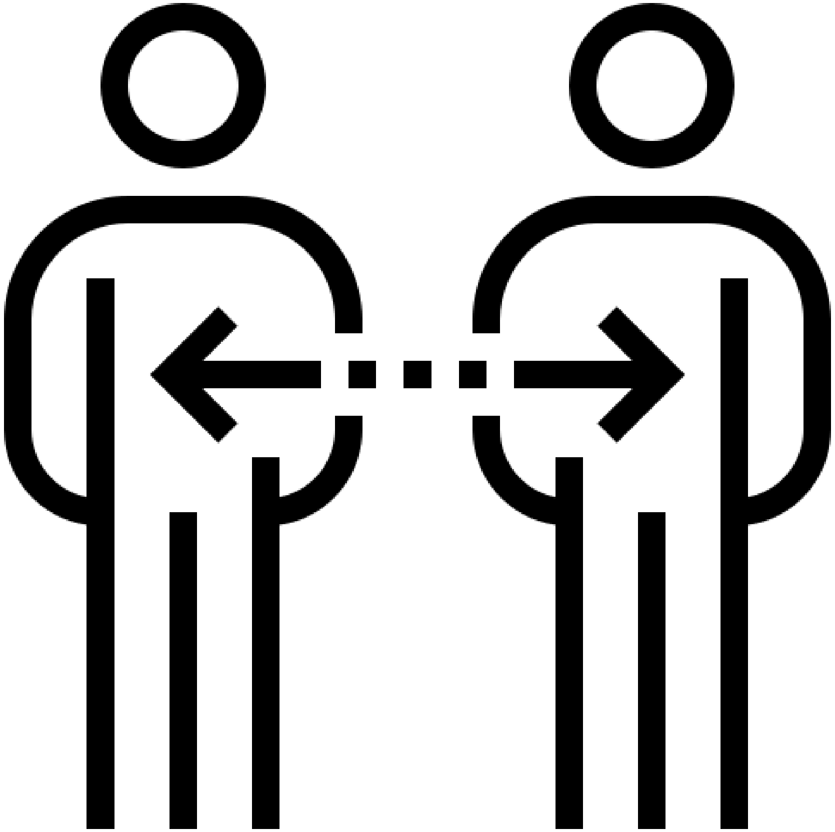
- Fusion improved prediction in a similar magnitude as case study 1, but only when setting specificity to 95%

- Inclusion of fused suicide risk screening variables provided both unique prediction and adjusted coefficient weights of local diagnosis codes

- Particular suicide risk scores were important to prediction:
  - Cosine: top 1000 similar patients
  - Weighted Jaccard’s distance: Top 10 similar patients
  - Weighted Manhattan distance: Top 10 similar patients
    - Methods generate variables with differing importance

Fusing
Similarity with cases

Case similarity



- In the final case study, I again present preliminary results predicting attempts in 10-24 patients that were hospitalized or seen in the ED, but within a ***small suburban*** medical center

- Here, another avenue of similarity is explored, which is to *skip* feature generation and include similarity scores directly

Case similarity

- With this, following the same framework, various similarity scores were fused to the target cohort (N cases = 173; N controls = 4322)
  - Similarity with cases in a large multicenter dataset of hospitalized and emergency patients in CT (N cases = 2828; N controls = 92752)
  - Expand the framework by not only calculating average scores but also median scores
  - Aggregating top *k* similar patients is not required

Case similarity

Variable		Target		External
		Cases	Controls	Cases
No of patients, N (%)		174 (3.87)	4322 (96.13)	2828 (2.96)
Gender, N (%)				
	Male	58 (33.33)	1910 (44.19)	1023 (36.17)
	Female	116 (66.67)	2412 (55.81)	1802 (63.72)
Age group, N (%)				
	10-14	34 (19.54)	1063 (24.60)	548 (19.38)
	15-19	91 (52.30)	1455 (33.66)	1253 (44.31)
	20-24	49 (28.16)	1804 (41.74)	1027 (36.32)
Race/ethnicity, N (%)				
	White	129 (74.14)	2505 (57.96)	1491 (52.72)
	Black or African American	6 (3.45)	359 (8.31)	363 (12.84)
	Hispanic or Latino	22 (12.64)	1171 (27.09)	752 (26.59)
	Other	17 (9.77)	287 (6.64)	222 (7.85)

Case similarity

Models	AUC (95% CI)	Sensitivity (95% CI)	
		95% specificity	90% specificity
Conventional	0.81 (0.77, 0.85)	0.41 (0.36, 0.46)	0.59 (0.53, 0.66)
Fusion	0.84 (0.78, 0.89)	0.51 (0.43, 0.60)*	0.61 (0.52, 0.72)

Case similarity

- Alike case study 2, fusion only improved predictions when setting specificity to 95%, however, AUC appeared to trend toward improvement

- INCLUDING DIAGNOSIS CODES, the strongest predictor in models (by magnitude) was case similarity using standard Pearson's *r*

- Inclusion of case similarity lessened importance of key local diagnosis codes:
  - Prior attempts
  - Suicidal ideation
  - Depression

Concluding statements

- All of these studies are generally **proof of concept**, however, improvements in prediction appear hopeful
- Identifying as many attempters as possible before an event occurs is critical, especially in pediatric populations
- Applied and methodological work is needed to better understand the mechanics underlying differences in improvement, similarity metrics and top patients included, and generated fusion variables
- All of these studies have limitations and full-scale grants dedicated to fusion may lead to larger improvements, better understanding

Concluding statements

- Importantly, obtaining information such as SDOH or suicide risk screening results directly from patients is ideal
- However data collection, especially in healthcare settings, is time-consuming, costly, and subject to noncompliance[20]
- “Estimating” these important features via data fusion is a free, time efficient method
- If such fusion algorithms are deployed in real-world settings, we may be able to help prevent additional suicide attempts in pediatric populations, as well as, in general populations

Thank you!

Thank you all for attending my talk!

I would also like to thank my mentors Kun Chen and Robert Aseltine

As well as Wanwan Xu and Fei Wang for their work on the topic

References

1. Centers for Disease Control and Prevention. Web-based Injury Statistics Query and Reporting System (WISQARS) [online]. National Center for Injury Prevention and Control, CDC (producer); 2020. Available from: www.cdc.gov/ncipc/wisqars/index.html.
2. Curtin S. State Suicide Rates Among Adolescents and Young Adults Aged 10–24: United States, 2000–2018. Hyattsville, MD: National Center for Health Statistics, 2020.
3. Ahmedani BK, Simon GE, Stewart C, Beck A, Waitzfelder BE, Rossom R, et al. Health care contacts in the year before suicide death. J Gen Intern Med. 2014; 29(6): 870-7. doi: 10.1007/s11606-014-2767-3.
4. Gordon JA, Avenevoli S, Pearson JL. Suicide Prevention Research Priorities in Health Care. JAMA Psychiatry. 2020; 77(9): 885-6. doi: 10.1001/jamapsychiatry.2020.1042.
5. Labouliere CD, Vasan P, Kramer A, Brown G, Green K, Rahman M, Kammer J, Finnerty M, Stanley B. "Zero Suicide" - A model for reducing suicide in United States behavioral healthcare. Suicidologi. 2018;23(1):22-30. PMID: 29970972; PMCID: PMC6022755.

References

6. Su C, Aseltine R, Doshi R, Chen K, Rogers SC, Wang F. Machine learning for suicide risk prediction in children and adolescents with electronic health records. Transl Psychiatry. 2020; 10(1): 413. doi: 10.1038/s41398-020-01100-0.
7. Xu W, Su C, Li Y, Rogers S, Wang F, Chen K, et al. Improving suicide risk prediction via targeted data fusion: proof of concept using medical claims data. J Am Med Inform Assoc. 2022; 29(3): 500-11. doi: 10.1093/jamia/ocab209.
8. Barak-Corren Y, Castro VM, Nock MK, Mandl KD, Madsen EM, Seiger A, et al. Validation of an Electronic Health Record-Based Suicide Risk Prediction Modeling Approach Across Multiple Health Care Systems. JAMA Netw Open. 2020; 3(3): e201262. doi: 10.1001/jamanetworkopen.2020.1262.
9. Pan S, Yang Q. A survey on transfer learning. IEEE Trans Know Data Eng. 2010; 22: 1345-59. doi: 10.1109/TKDE.2009.191.
10. Hahn RA. What is a social determinant of health? Back to basics. J Public Health Res. 2021; 10(4): 2324. doi: 10.4081/jphr.2021.2324.

References

11. Kim YJ, Quinn CR, Moon SS. Buffering Effects of Social Support and Parental Monitoring on Suicide. Health Soc Work. 2021; 46(1): 42-50. doi: 10.1093/hsw/hlaa037.
12. Lawrence RE, Oquendo MA, Stanley B. Religion and Suicide Risk: A Systematic Review. Arch Suicide Res. 2016; 20(1): 1-21. doi: 10.1080/13811118.2015.1004494.
13. O'Connor RC, Smyth R, Williams JM. Intrapersonal positive future thinking predicts repeat suicide attempts in hospital-treated suicide attempters. J Consult Clin Psychol. 2015; 83(1): 169-76. doi: 10.1037/a0037846.
14. DeVylder JE, Frey FJ, Cogburn CD, Wilcox HC, Sharpe TL, Oh HY, Nam B, et al. Elevated Prevalence of Suicide Attempts among Victims of Police Violence in the USA. J Urban Health. 2017; 94(5): 629-36. doi: 10.1007/s11524-017-0160-3.
15. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: A systematic review. J Am Med Inform Assoc. 2020; 27(11): 1764-73. doi: 10.1093/jamia/ocaa143.

References

16. Harris KM, Halpern CT, Whitsel EA, Hussey JM, Killeya-Jones LA, Tabor J, et al. Cohort Profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). Int J Epidemiol. 2019; 48(5): 1415-1415k. doi: 10.1093/ije/dyz115.
17. King CA, O'Mara RM, Hayward CN, Cunningham RM. Adolescent suicide risk screening in the emergency department. Acad Emerg Med. 2009; 16(11): 1234-41. doi: 10.1111/j.1553-2712.2009.00500.x.
18. Horowitz L, Ballard E, Teach SJ, Bosk A, Rosenstein DL, Joshi P, et al. Feasibility of screening patients with nonpsychiatric complaints for suicide risk in a pediatric emergency department: a good time to talk? Pediatr Emerg Care. 2010; 26(11): 787-92. doi: 10.1097/PEC.0b013e3181fa8568.
19. Posner, Kelly, Gregory K. Brown, Barbara Stanley, David A. Brent, Kseniya V. Yershova, Maria A. Oquendo, Glenn W. Currier et al. "The Columbia–Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults." American journal of psychiatry 168, no. 12 (2011): 1266-1277.
20. Basch E. Patient-Reported Outcomes - Harnessing Patients' Voices to Improve Clinical Care. N Engl J Med. 2017; 376(2): 105-8. doi: 10.1056/NEJMp1611252.

Similarity equations

- Below is Pearson's *r* and Manhattan *d*
- For i in n_t = total sample of target, For j in n_e = total sample of external, For q in p features

$$r_{ij} = \frac{\dfrac{\sum_q^p w\left(x_{iq}^t - \dfrac{\sum_q^p w x_{iq}^t}{\sum_q^p w_q}\right)\left(x_{jq}^e - \dfrac{\sum_q^p w x_{jq}^e}{\sum_q^p w_q}\right)}{\sum_q^p w_q}}{\left(\dfrac{\sum_q^p w\left(x_{iq}^t - \dfrac{\sum_q^p w x_{iq}^t}{\sum_q^p w_q}\right)^2}{\sum_q^p w_q}\right)\left(\dfrac{\sum_q^p w\left(x_{jq}^e - \dfrac{\sum_q^p w x_{jq}^e}{\sum_q^p w_q}\right)^2}{\sum_q^p w_q}\right)}$$

$$d_{ij} = 1 - \frac{\sum_q^p W_q \left|x_{iq}^t - x_{jq}^e\right|}{p}$$

Time for an example

First study in context of suicide risk:
Xu et al (2021)[7]

Models	PPV (*SD*)	
	95% specificity	**90% specificity**
Conventional	0.057 (0.018)	0.050 (0.009)
Fusion	0.134 (0.015)	0.082 (0.008)

Social determinants of health

Models	PPV (95% CI)	
	95% specificity	90% specificity
Conventional	0.068 (0.059, 0.077)	0.054 (0.046, 0.063)
Fusion	0.074 (0.064, 0.085)	0.058 (0.050, 0.067)

Suicide risk screening

Models	PPV (95% CI)	
	95% specificity	**90% specificity**
Conventional	0.30 (0.27, 0.33)	0.22 (0.20, 0.24)
Fusion	0.32 (0.29, 0.35)	0.22 (0.20, 0.24)

Case similarity

Models	PPV (95% CI)	
	95% specificity	90% specificity
Conventional	0.24 (0.22, 0.26)	0.19 (0.18, 0.21)
Fusion	0.28 (0.25, 0.32)	0.20 (0.17, 0.22)