INTRODUCTION

Suicide accounted for over 44,000 deaths in the United States in 2015, a 25% increase in rate since 2000. Approximately half of people dying by suicide – and 70% of those making non-fatal suicide attempts - have received some mental health outpatient care in the prior year, including both mental health specialty visits and primary care visits with mental health diagnoses. Mindful of these potential opportunities for prevention, the Joint Commission recently issued a Sentinel Event Alert regarding detection of suicide risk in all health care settings. Existing tools for identifying suicide risk include self-report questionnaires and predictions computed from electronic health records. Both have shortcomings in sensitivity and efficiency of risk concentration.

A previous supplement to the MHRN cooperative agreement supported development of a population-based suicide risk calculator, predicting risk of suicide attempt or suicide death following an outpatient visit using both responses to PHQ9 item 9 and discrete data extracted from health system electronic health records. Using a database of approximately 20 million visits by 3 million patients aged 13 and over, we developed and validated machine learning logistic regression models predicting risk of suicide attempt and suicide death within 90 days of an outpatient visit – either a visit to a specialty mental provider or a primary care visit in which a mental health diagnosis was recorded. Potential predictors included demographic and clinical data extracted from health system records for the 5 years prior to each visit: prior suicidal behavior, mental health and substance use diagnoses, general medical diagnoses, prescriptions for psychiatric medications, inpatient or emergency department mental health care, and responses to routinely administered PHQ9 depression questionnaires. Models were developed in a 65% random sample of visits and validated in the remaining 35%. Variable selection models considered 150 discrete predictors and 164 potential interactions. In the validation sample, the 5% of mental health specialty visits with highest risk scores accounted for 43% of subsequent suicide attempts and 47% of suicide deaths. Areas under the receiver operating characteristic curves (AUCs) for prediction of suicide attempt and suicide death were 0.85 and 0.86. In the validation sample, the 5% of primary care visits with highest risk scores accounted for 48% of subsequent suicide attempts and 43% of suicide deaths. AUCs for prediction of suicide attempt and suicide death were 0.85 and 0.83.

While these models represent a substantial advance over existing risk prediction or risk stratification tools, we identify several significant limitations. Fixed limits of our computational methods (penalized LASSO logistic regression in the R computing environment) forced us to limit both our sample size and the number of potential predictors and interaction terms. Those methods also limit ability to appropriate account for clustering of observations within patients and account for the sparse and skewed distributions of predictor data. Finally, we now recognize the need to extend these methods to predict risk following acute-care (inpatient and emergency department) encounters.

We now propose a next stage of work to address these limitations. Specific aims of this next stage include:

1. Expand and enhance the risk prediction dataset to: include larger numbers of observations with data regarding self-reported suicidal ideation (PHQ9 Item 9), include additional encounters and events following the transition from ICD9 to ICD10 diagnoses, and allow more detailed consideration of the timing of predictor events (diagnoses, encounters, prescription fills)
2. Expand sampling to include emergency department and inpatient encounters
3. Evaluate alternative modeling approaches, including classification- or tree-based approaches such as Classification and Regression Trees (CART), Mixed Effects Regression Trees (MERT), and Random Forest.
4. Rapidly disseminate all methods, tools and results to a wide range of stakeholders including health systems, researchers, and EHR vendors.

The new work proposed here will make several significant and unique contributions to suicide prevention research, including:

- An unprecedented sample size of patients, visits, and suicidal behavior outcomes
- A population-based sample including mental health and primary care visits from seven health systems
- Integration of electronic health records data (diagnoses, prescriptions) with systematically collected data regarding suicidal ideation
- Inclusion of more diverse and detailed predictors than any previous research in this area

SIGNIFICANCE

Suicide accounted for over 44,000 deaths in the United States in 2015, reflecting a 25% increase in suicide mortality rate since 2000[1]. Non-fatal suicide attempts account for almost 500,000 emergency department visits annually[2]. Of people dying by suicide, approximately half have received some mental health outpatient care in the prior year, including both mental health specialty visits and primary care visits with mental health diagnoses[3]. Over 70% of people making nonfatal suicide attempts have received some outpatient mental health care in the prior year [4]. Mindful of these potential opportunities for prevention, the Joint Commission recently issued a Sentinel Event Alert regarding detection of suicide risk in all health care settings [5].

We have previously reported that response to item 9 of the commonly-used PHQ9 depression questionnaire is a robust predictor of subsequent suicide attempt or suicide death. Those reporting thoughts of death or self-harm “nearly every day” were seven times as likely to attempt suicide and six times as likely to die by suicide over the following 90 days [6, 7].

This finding was immediately disseminated to our MHRN health systems and led to implementation of standard care processes using response to PHQ9 Item 9 as a risk stratification tool [8]. In mental health specialty clinics, patients reporting thoughts of death or self-harm “more than half the days” or “nearly every day” are expected to receive a structured second-stage assessment (typically using the Columbia Suicide Severity Rating Scale) and participate in creation of a structured safety plan.

While we were encouraged by this rapid implementation of our original findings, we identified significant shortcomings of risk stratification based solely on response to PHQ9 item 9. Sensitivity is only moderate. Over 35% of suicide attempts and deaths occurred among patients reporting suicidal ideation “not at all”. Efficiency in identifying high-risk patients is also only moderate. The 6% of patients reporting frequent suicidal ideation account for only 33% of subsequent suicide attempts and deaths. More sensitive and efficient risk prediction tools are needed. Most important, this risk stratification method is of no use for visits in which a PHQ9 score is not recorded. Even in MHRN health systems where use of the PHQ9 is well established, scores are recorded for only approximately 65% of visits with a mental health diagnosis.

Several recent reports have used data mining or machine learning methods to predict suicidal behavior from electronic health records data. For example, analyses have attempted to predict suicide death among Veterans Health Administration service users [9], suicide death following psychiatric hospitalization among Army soldiers [10], and suicide or accidental death following civilian general hospital discharge[11]. Two recent reports have attempted to predict suicidal behavior following outpatient visits. Kessler and colleagues[12] used health records data and military service records to predict suicide death following outpatient mental health visits by male US Army soldiers, finding that almost half of suicide deaths occurred among the 15% of visits with highest risk scores and approximately one quarter of deaths occurred among the 5% of visits rated as highest risk. Barak-Corren and colleagues[13] used health system records data to predict suicide attempt or death among patients receiving outpatient care in two large academic health systems. Approximately one-third of suicide attempts and deaths occurred in 5% of patients with highest risk scores.

PROGRESS REPORT

A previous supplement to the MHRN cooperative agreement supported development of a population-based suicide risk calculator, predicting risk of suicide attempt or suicide death following an outpatient visit using both responses to PHQ9 item 9 and discrete data extracted from health system electronic health records. The database assembled for this work included approximately 20 million outpatient visits in seven MHRN health systems between 1/1/2009 and 9/30/2015, including all specialty mental health visits and all visits to primary care providers associated with mental health diagnoses. Among the 3 million members/patients included in the database, health system records and state death certificate data identified approximately 29,000 non-fatal suicide attempts and 2100 suicide deaths. Potential predictors included approximately 150 indicators characterizing demographic characteristics, current and previous psychiatric and substance use diagnoses, current and previous medication use, chronic medical conditions, as well as current and previous responses to PHQ9 depression questionnaires. Because this sample covered a period when implementation of measurement-based care was variable across MHRN health systems, actual use of the PHQ9 varied both between health systems and within health systems over time. On average, PHQ9 scores were recorded for approximately 10% of visits and some data regarding prior PHQ9 scores were available for approximately 15%. Approximately 200 interaction terms (of thousands possible) were selected by investigators based on findings in our previous analyses (described above). Prediction models were developed in three steps:

- Variable selection – Using a 65% random sample of visits, logistic regression models with penalized lasso variable selection were estimated for prediction of suicide attempt and suicide death within 90 days of visit. Separate models were estimated for mental health specialty and primary care visits.
- Estimation – Because available software for LASSO variable selection cannot account for clustering (multiple visits per patient), predictors selected in the first step were entered into a second hierarchical logistic model (accounting for clustering) to estimate coefficients in the same 65% random sample.
- Validation – The final model (selected variables from the first step and coefficients estimated in the second step) was implemented in the remaining 35% random sample to evaluate model performance.

In models predicting suicide attempt and suicide death among mental health specialty patients, predictors selected included indicators of recent and lifetime suicidal behavior, specific psychiatric diagnoses (depression, eating disorder, bipolar disorder), and substance use diagnoses. Indicators regarding recent and past-year responses to PHQ9 item 9 were selected, even though these scores were available for only 10-15% of the sample. The table below shows model performance in the independent validation sample for prediction of suicide attempt and suicide death within 90 days of a mental health specialty visit or primary care visit with a mental health diagnosis.

SUICIDE ATTEMPTS FOLLOWING MENTAL HEALTH SPECIALTY VISIT				SUICIDE DEATHS FOLLOWING MENTAL HEALTH SPECIALTY VISIT			
Risk Score Percentile	Predicted Risk[1]	Actual Risk[2]	% of All Attempts[3]	Risk Score Percentile	Predicted Risk[1]	Actual Risk[2]	% of All Deaths[3]
>99.5th	13.0%	12.7%	10%	>99.5th	0.654%	0.694%	12%
99th to 99.5th	8.5%	8.1%	6%	99th to 99.5th	0.638%	0.595%	11%
95th to 99th	4.1%	4.2%	27%	95th to 99th	0.162%	0.167%	25%
90th to 95th	1.9%	1.8%	15%	90th to 95th	0.068%	0.088%	16%
75th to 90th	0.9%	0.9%	21%	75th to 90th	0.031%	0.029%	16%
50th to 75th	0.3%	0.3%	13%	50th to 75th	0.014%	0.015%	13%
<50th	0.1%	0.1%	8%	<50th	0.003%	0.003%	6%
SUICIDE ATTEMPTS FOLLOWING PRIMARY CARE VISIT				SUICIDE DEATHS FOLLOWING PRIMARY CARE VISIT			
Risk Score Percentile	Predicted Risk[1]	Actual Risk[2]	% of All Attempts[3]	Risk Score Percentile	Predicted Risk[1]	Actual Risk[2]	% of All Deaths[3]
>99.5th	8.6%	8.0%	15%	>99.5th	0.536%	0.435%	14%
99th to 99.5th	4.1%	4.2%	8%	99th to 99.5th	0.181%	0.197%	7%
95th to 99th	1.6%	1.6%	25%	95th to 99th	0.092%	0.083%	22%
90th to 95th	0.7%	0.7%	13%	90th to 95th	0.035%	0.038%	13%
75th to 90th	0.3%	0.3%	18%	75th to 90th	0.018%	0.019%	19%
50th to 75th	0.1%	0.1%	12%	50th to 75th	0.009%	0.009%	15%
<50th	0.04%	0.04%	9%	<50th	0.003%	0.003%	10%

Notes:
1 – Predicted risk in this stratum based on predictors and coefficients from development sample
2 – Observed risk in this stratum
3 – Percentage of all suicide attempts or deaths occurring in this stratum

Absolute risk levels for both suicide attempts and suicide deaths were lower in the primary care sample, but the specific predictors selected and the overall accuracy of prediction were similar to those seen in specialty mental health patients.

Comparison of these model results with risk stratification based on PHQ9 item 9 score only (described above) demonstrates substantial improvement in both sensitivity and efficiency of risk stratification. Among mental health specialty patients, visits with risk scores below the 75th percentile accounted for only 21% of all suicide attempts and 19% of suicide deaths, compared to over 35% of attempts and deaths occurring among visits with responses of "not at all" to PHQ9 item 9. And visits with scores above the 95th percentile accounted for 43% of all suicide attempts and 47% of all suicide deaths, compared to approximately 30% of attempts and deaths occurring among the 6% of visits responding “more than half the days” or “nearly every day” to PHQ9 item 9. In the independent validation sample, areas under the ROC curves for prediction of suicide attempt and suicide death in the specialty mental health sample were 0.85 and 0.86 respectively. Corresponding AUCs in the primary care sample were 0.85 and 0.83. By this AUC measure, our prediction models substantially outperformed all previously described models predicting suicide death following outpatient visit[12, 13], psychiatric hospitalization[10], or general medical hospitalization[11], where AUC values ranged from 0.67 to 0.77.

This work represents a significant advance over risk prediction tools previously available, and a substantial improvement over the risk stratification procedures currently used in MHRN health systems (based on PHQ9 scores alone). Findings will be immediately applied in MHRN health systems as part of health system implementation of Zero Suicide strategies. Implementation of model-based risk scores within the Epic electronic health record is planned for KP Washington (previously Group Health Cooperative) in late 2017 and in other KP regions in 2018. We have shared methods and preliminary results with colleagues at Epic, and are exploring rapid nationwide dissemination of risk scoring tools to all Epic-using health systems (over 50% of US physician practices). Given that analytic datasets were created from data resources following the HCSRN/PCORnet/Sentinel common data model, these methods could be applied to electronic health records or insurance claims data from other large health systems. To facilitate dissemination, all specifications and computer code for development of analytic datasets are now available via the MHRN public code repository (https://github.com/MHResearchNetwork).

While these results and generalizable tools represent a significant advance over existing risk stratification tools, we identify several significant limitations:

- Limited capacity for predictors – Because traditional regression-based methods operate on the entire design matrix, and because standard software limits the size that matrix/vector to approximately 2.1 billion bits, variable selection models could only include a total of approximately 350 predictors (main effects plus interactions) rather than the originally intended total of approximately 500.
- Limited sample size – The same limitation on matrix size/vector length required that we reduce the size of the development sample from the originally intended 80% to 65%, compromising precision for estimation of suicide death outcomes (approximately 10% as common as suicide attempt outcomes).
- Limited availability of questionnaire data – In our sample of visits from 2009 to 2015, PHQ9 questionnaire data were recorded for approximately 10% of encounters (increasing from 4% in 2009 to 12% in 2015). Accuracy of prediction was greater in health systems with more frequent use of the PHQ9. In addition, responses to PHQ9 contributed more to risk prediction among primary care patients, for whom more detailed data regarding mental health diagnoses were less often available.
- Limited to outpatient visits – Our work completed and in progress has focused on risk following outpatient mental health and primary care visits, reflecting the priorities of delivery system leaders in our MHRN health systems. But we recognize substantial interest (nationally and in our MHRN health systems) in tools to predict suicide risk following inpatient and emergency department encounters.
- Limited ability to account for clustering within individuals – Available software for variable selection in traditional regression models (such as LASSO) cannot account for clustering of observations within individuals, and this necessitates the three-step process described above. Appropriately accounting for clustering is essential to distinguish between stable risk predictors (i.e. between-person variation) and time-varying predictors (i.e. within-person variation).
- Reliance on parametric methods – For many predictors of interest, original count data (e.g. number of emergency department mental health visits) were reduced to dichotomous predictors (e.g. presence/absence of any emergency department mental health visit) in order to accommodate parametric assumptions of a logistic regression model.

We propose below a next stage of work to address these limitations.

APPROACH

AIM 1: EXPAND AND ENHANCE EXISTING RISK PREDICTION DATASET

We propose to assemble new visit-level datasets from each participating site in the fall of 2017. These new data will expand and enhance the existing dataset in the following ways:

- Extend visit sampling through 12/31/2016 – Extrapolating from the current database, we anticipate that extending the visit sampling period by 15 months will increase the total visit sample from approximately 20 million to approximately 24 million.
- Extend outcome observation through 12/31/2016 – – Extrapolating from the current database, we anticipate that extending the outcome period by 15 months will increase the number of unique suicide attempts observed from 29,000 to approximately 35,000 and increase the number of unique suicide deaths observed from 2,100 to 2,600. This expanded sample will allow increased precision of risk prediction in patient subgroups of specific interest (adolescents, people with substance use disorders, people with psychotic disorders).
- Increase the proportion of visits for which PHQ9 Item 9 scores were recorded – Given rapid increases in use of the PHQ9 questionnaire at the KP Northwest and KP Southern California sites, we anticipate that extending visit sampling through 2016 will increase the proportion of visits with recorded PHQ9 scores from approximately 10% in the current sample to approximately 18%. This will enable sensitivity analyses limited to visits in 2014 or later when use of the PHQ9 was more prevalent at all sites. Examining the additional contribution of PHQ9 scores to risk prediction will address both clinical questions (How important is self-reported suicidal ideation in prediction of suicide risk?) and policy questions (Will systematic implementation of measurement-based care facilitate suicide prediction efforts?).
- Include new ICD10-CM codes for ascertainment of non-fatal suicide attempts – As we have recently published[14], extending the outcome period to included ICD10-CM diagnoses is expected to improve overall ascertainment of definite or possible self-harm and more accurately distinguish between definite and possible self-harm. While we expect improved prediction of suicide attempts under the new diagnostic classification, effects of this change should be empirically evaluated.
- Expand predictors to include recent and long-term “non-adherence” behavior, including: missed mental health and general medical appointments, failure to refill long-term mental health and general medical medications.
- Include additional detail regarding classification and timing of prior suicide attempts – We propose to expand the number of predictors regarding prior suicide attempts from 12 dichotomous indicators (4 categories in 3 time periods) to 36 count variables (6 categories in 6 time periods) reflecting the number of times a definite or possible self-harm diagnosis was recorded in each category.
- Include additional detail regarding timing of diagnoses – We propose to expand the number of predictors regarding prior mental health and substance use diagnoses from 42 dichotomous indicators (14 categories in 3 time periods) to 84 count variables (14 categories in 6 time periods) reflecting the number of times a diagnosis in each category was recorded.
- Include additional detail regarding timing of prior medication exposures - We propose to expand the number of predictors regarding medication exposures from 12 dichotomous indicators (4 categories in 3 time periods) to 36 count variables (6 categories in 6 time periods) reflecting the number of days supply of medication dispensed in each category.
- Include additional detail regarding Emergency Department (ED) and inpatient care – We propose to expand the number of predictors regarding ED and inpatient care from 6 dichotomous indicators (any ED or inpatient utilization for mental health diagnosis in 3 time periods) to 36 count indicators (6 categories in 6 time periods) reflecting the number of ED or inpatient encounters in 3 diagnosis groups (mental health, self-harm, other).

AIM 2: EXPAND SAMPLING TO INCLUDE INPATIENT AND EMERGENCY DEPARTMENT ENCOUNTERS

We propose to extend our current methods to develop parallel models predicting risk of suicide attempt and suicide death following emergency department and inpatient encounters. These models would consider the same range of potential predictors described above. If sample sizes permit, models for emergency department and inpatient encounters will be stratified according to diagnosis (mental health vs. general medical diagnosis).

AIM 3: EVALUATE ALTERNATIVE MODELING METHODS

We propose to implement alternative non-parametric prediction methods to accommodate the larger dataset and increased number and precision of predictor variables. We will evaluate predictive accuracy and characterize computational burden for the following approaches.

- Estimate risk of suicide attempt in the expanded dataset using classification and regression trees (CART) [15], a non-parametric risk stratification method that can accommodate larger sample sizes and a greater number of predictors than traditional, parametric regression methods. CART model estimation proceeds via a recursive search to identify the binary partition of a covariate space that accounts for the most variability in the outcome. The partitioning algorithm is repeated over subsets of the covariate space until the marginal gain from additional partitions falls below a pre-specified threshold. In so doing, the procedure uncovers covariate values more predictive of the outcome. As an example, stratifying the population based on two or more suicide attempts in the previous 3 months may better predict a patient's risk than our current variable definition (any attempt vs. none). Moreover, within a subset of the patients, e.g., those with a particular diagnosis, the partition of previous suicide attempts that better fits the outcome data may differ. This approach is particularly appealing in the context of suicide risk prediction as risk factors may very across patient subgroups. Following established methods, we will employ bootstrap aggregation, or "bagging", to average over many trees estimated on bootstrapped datasets and, as a result, reduce prediction bias [16]. Advantages of this approach include:
    - CART does not require a priori selection of variables or interactions. Thus, accuracy of CART predictions does not depend on correct variable selection, and the additional computational step required to perform model selection is avoided.
    - By evaluating single, binary partitions of the covariate space in turn, CART is also more computationally tractable with a large number of records and covariates than parametric regression procedures that entail inverting the entire covariate matrix to estimate parameters.
    - Results from CART are also easy to interpret; classification trees illustrate salient predictors to end users, so that face validity of the evidence can be evaluated by stakeholders and potential end users.
- One potential disadvantage of CART and similar approaches is that non-parametric or tree-based classification does not yield a simple polynomial expression for calculation of individual-level risk in a new sample of patients. Consequently, findings from CART models would be more difficult to implement in standard electronic health record systems (such as Epic). It will therefore be important to evaluate the advantages of non-parametric or tree-based approaches over the logistic regression models already developed.
- Initial work will evaluate predictive performance of the CART model on our current dataset with approximately 20 million records and approximately 150 predictors in order to compare accuracy to our current parametric regression approach. This will allow us to develop and evaluate analytic methods while new expanded datasets are being extracted and assembled.
- We will extend standard CART approaches to account for correlated nature of data, that is, for multiple observations within each patient, instead of assuming that observations are independent conditional on observed covariates from the EHR. We will consider random effect estimation (RE-EM) [17, 18] trees and mixed effects regression trees (MERT) [19], as time permits. Accounting for correlated suicide attempt outcomes within patients who access mental health services on multiple occasions will enable us to more accurately assess the temporal or within-person association between predictors and the outcome as well as the uncertainty of those association estimates. Furthermore, estimates of patient-level random effects could, in turn, be used to make more precise suicide risk predictions for patients with multiple visits.
- We will explore additional non-parametric methods including random forest regression [20], multivariate adaptive regression splines for classification (PolyMARS) [21], and multiple additive regression trees (MART) [22]. As time allows, we will compare performance and computational burden of alternative methods.
- All analyses will be performed with the open source statistical computing software R. The following packages will be used for the various analytic approaches we are considering: rpart (CART), REEMtree (RE-EM trees), RandomForest (random forests), polspline (PolyMARS), and gbm (MART),

<u>AIM 4: RAPIDLY DISSEMINATE ALL TOOLS AND PRODUCTS</u>

Continuing the practice established in our previous work, all technical tools developed in this work will be immediately distributed to any interested users via our public repository of technical materials. Specifically:

- This complete research plan will be posted on notification of funding
- All specifications and code for extraction of data from health system records will be posted as soon as extraction and data quality control are complete

- All specifications and code for final analytic datasets will be posted as soon as datasets are created.
- All analytic code will be posted as soon as final analyses are complete.
- All analytic results (including complete risk prediction models) will be posted as soon as final results have completed peer review.

As described above, we have already established a collaboration with the largest national EHR vendor to facilitate rapid dissemination of risk score calculations to a wide range of health care settings, including those lacking sophisticated analytic capability.

<u>POTENTIAL LIMITATIONS</u>

- Portability of tree-based risk predictions – While tree-based classification methods (such as random forest) are less computationally intensive and can accommodate a larger number of potential predictors, the resulting predicted probabilities represent a weighted sum of an ensemble of tree structures rather than a simple polynomial expression (as in logistic regression). Consequently, those predictions are not as easily transported into standard electronic health records systems (such as the Epic “Healthy Planet” predictive analytics module). Consequently, we plan explicit comparison of the accuracy of risk predictions derived from tree-based methods with those derived from logistic regression methods – to quantify this potential trade-off between portability and prediction accuracy.
- Limitation to patients with identified mental health need - Our analyses consider only the 25% of health system members seeking specialty mental health or receiving a mental health diagnosis in primary care. Previous research indicates that one third of suicide attempts and half of suicide deaths occur among those with no recorded mental health diagnosis or treatment [3, 4]. Risk prediction models using health records could also prove useful among those without identified mental health need, but predictors would certainly be different from the mental health diagnoses and treatments selected in this sample. A separate MHRN project, led by Brian Ahmedani, is identifying predictors of suicidal behavior across the entire health system membership, with a specific focus on general medical diagnoses and patterns of medical utilization
- Health records data do not consider important predictors - Health system records do not contain readily accessible information regarding many important risk factors for suicidal behavior. Consequently, we cannot examine the effects of negative events such as job loss, bereavement, or relationship disruption. Suicidal behavior likely reflects the intersection of longer-term clinical risk factors with the more immediate influence of negative life events. We are engaged with MHRN health system leaders to develop more standardized methods for assessing and recording important social risk factors and life events.

INNOVATION

The new work proposed here will make several significant and unique contributions to suicide prevention research, including:

- An unprecedented sample size of patients, visits, and suicidal behavior outcomes
- A population-based sample including mental health and primary care visits from seven health systems
- Integration of electronic health records data (diagnoses, prescriptions) with systematically collected data regarding suicidal ideation
- Inclusion of more diverse and detailed predictors than any previous research in this area

In addition, building on data infrastructure and methods already developed will allow remarkable efficiency in both speed and cost. As described below, we propose to assemble a data resource including approximately 25 million visits in seven healthcare systems, accomplish primary analyses in 12 months, and complete all analyses in approximately 18 months – with direct costs predicted to be approximately $450,000.

REFERENCES

1. Centers for Disease Control and Prevention. WISQARS Fatal Injury Reports, National, Regional, and State, 1981-2015. 2017 [updated 2017; cited 2017 April 4]; Available from: https://webappa.cdc.gov/sasweb/ncipc/mortrate.html.
2. Centers for Disease Control and Prevention. WISQARS Nonfatal Injury Reports, 200-2014. 2017 [updated 2017; cited 2017 April 4]; Available from: https://webappa.cdc.gov/sasweb/ncipc/nfirates.html.
3. Ahmedani BK, Simon GE, Stewart C, Beck A, Waitzfelder BE, Rossom R, Lynch F, Owen-Smith A, Hunkeler EM, Whiteside U, Operskalski BH, Coffey MJ, Solberg LI. Health care contacts in the year before suicide death. J Gen Intern Med. 2014;29(6):870-7. PMCID: PMC4026491.
4. Ahmedani BK, Stewart C, Simon GE, Lynch F, Lu CY, Waitzfelder BE, Solberg LI, Owen-Smith AA, Beck A, Copeland LA, Hunkeler EM, Rossom RC, Williams K. Racial/Ethnic differences in health care visits made before suicide attempt across the United States. Med Care. 2015;53(5):430-5. PMCID: PMC4397662.
5. Patient Safety Advisory Group. Detecting and treating suicidal ideation in all settings. The Joint Commission Sentinel Event Alerts. 2016;56.
6. Simon GE, Rutter CM, Peterson D, Oliver M, Whiteside U, Operskalski B, Ludman EJ. Does Response on the PHQ-9 Depression Questionnaire Predict Subsequent Suicide Attempt or Suicide Death? Psychiatr Serv. 2013;64(12):1195-202.
7. Simon GE, Coleman KJ, Rossom RC, Beck A, Oliver M, Johnson E, Whiteside U, Operskalski B, Penfold RB, Shortreed SM, Rutter C. Risk of suicide attempt and suicide death following completion of the Patient Health Questionnaire depression module in community practice. J Clin Psychiatry. 2016;77(2):221-7.
8. Rossom RC, Simon GE, Beck A, Ahmedani BK, Steinfeld B, Trangle M, Solberg L. Facilitating Action for Suicide Prevention by Learning Health Care Systems. Psychiatr Serv. 2016;67(8):830-2. PMCID: PMC4969117.
9. McCarthy JF, Bossarte RM, Katz IR, Thompson C, Kemp J, Hannemann CM, Nielson C, Schoenbaum M. 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.
10. Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ, Brown M, 3rd, Cai T, Colpe LJ, Cox KL, Fullerton CS, Gilman SE, Gruber MJ, Heeringa SG, Lewandowski-Romps L, Li J, Millikan-Bell AM, Naifeh JA, Nock MK, Rosellini AJ, Sampson NA, Schoenbaum M, Stein MB, Wessely S, Zaslavsky AM, Ursano RJ, Army SC. 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. PMCID: PMC4286426.
11. McCoy TH, Jr., Castro VM, Roberson AM, Snapper LA, Perlis RH. Improving Prediction of Suicide and Accidental Death After Discharge From General Hospitals With Natural Language Processing. JAMA Psychiatry. 2016;73(10):1064-71.
12. Kessler RC, Stein MB, Petukhova MV, Bliese P, Bossarte RM, Bromet EJ, Fullerton CS, Gilman SE, Ivany C, Lewandowski-Romps L, Millikan Bell A, Naifeh JA, Nock MK, Reis BY, Rosellini AJ, Sampson NA, Zaslavsky AM, Ursano RJ, Army SC. Predicting suicides after outpatient mental health visits in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). Mol Psychiatry. 2016((epub July 19)).
13. Barak-Corren Y, Castro VM, Javitt S, Hoffnagle AG, Dai Y, Perlis RH, Nock MK, Smoller JW, Reis BY. Predicting Suicidal Behavior From Longitudinal Electronic Health Records. Am J Psychiatry. 2017;174(2):154-62.
14. Stewart C, Crawford PM, Simon GE. Changes in Coding of Suicide Attempts or Self-Harm With Transition From ICD-9 to ICD-10. Psychiatr Serv. 2017;68(3):215.
15. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Monterey, CA: Wadsworth and Brooks/Cole Adanced Books and Software; 1984.
16. Breiman L. Bagging Predictors. Machine Learning. 1996;24:123-40.
17. Sela RJ, Simonoff JS. RE-EM trees: A data mining approach for longitudinal and clustered data. Machine Learning. 2012;82(2):169-207.
18. Fu W, Simonoff JS. Unbiased regression trees for longitudinal and clustered data. Computational Statistics and Data Analysis. 2015;88:53-74.

19. Hajjem A, Bellavance F, Larocque D. Mixed effects regression trees for clusted data. Statistics and Probability Letters. 2011;81(4):451-9.
20. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32.
21. Friedman JH. Multivariate adaptive regression splines. Ann Stat. 1991;19(1):1-141.
22. Friedman JH, Meulman JJ. Multiple additive regression trees with application in epidemiology. Stat Med. 2003;22(9):1365-81.