ARTICLE   OPEN

Check for updates

Next-generation precision medicine for suicidality prevention

R. Bhagar[1,12], S. S. Gill [1,2,12], H. Le-Niculescu[1,10,12], C. Yin[3,12], K. Roseberry[1], J. Mullen[4], M. Schmitz[2], E. Paul[1,5], J. Cooke[5], C. Tracy[1,5], Z. Tracy[1,5], A. S. Gettelfinger[1], D. Battles[6], M. Yard[7], G. Sandusky[7], A. Shekhar[1,11], S. M. Kurian[8], P. Bogdan [3] and A. B. Niculescu [1,2,5,7,9,10✉]

© The Author(s) 2024

Suicidality remains a clear and present danger in society in general, and for mental health patients in particular. Lack of widespread use of objective and/or quantitative information has hampered treatment and prevention efforts. Suicidality is a spectrum of severity from vague thoughts that life is not worth living, to ideation, plans, attempts, and completion. Blood biomarkers that track suicidality risk provide a window into the biology of suicidality, as well as could help with assessment and treatment. Previous studies by us were positive. Here we describe new studies we conducted transdiagnostically in psychiatric patients, starting with the whole genome, to expand the identification, prioritization, validation and testing of blood gene expression biomarkers for suicidality, using a multiple independent cohorts design. We found new as well as previously known biomarkers that were predictive of high suicidality states, and of future psychiatric hospitalizations related to them, using cross-sectional and longitudinal approaches. The overall top increased in expression biomarker was SLC6A4, the serotonin transporter. The top decreased biomarker was TINF2, a gene whose mutations result in very short telomeres. The top biological pathways were related to apoptosis. The top upstream regulator was prednisolone. Taken together, our data supports the possibility that biologically, suicidality is an extreme stress-driven form of active aging/death. Consistent with that, the top subtypes of suicidality identified by us just based on clinical measures had high stress and high anxiety. Top therapeutic matches overall were lithium, clozapine and ketamine, with lithium stronger in females and clozapine stronger in males. Drug repurposing bioinformatic analyses identified the potential of renin-angiotensin system modulators and of cyclooxygenase inhibitors. Additionally, we show how patient reports for doctors would look based on blood biomarkers testing, personalized by gender. We also integrated with the blood biomarker testing social determinants and psychological measures (CFI-S, suicidal ideation), showing synergy. Lastly, we compared that to machine learning approaches, to optimize predictive ability and identify key features. We propose that our findings and comprehensive approach can have transformative clinical utility.

INTRODUCTION

"The cure of many diseases remains unknown to the physicians because they do not study the whole person."

    Socrates

Suicides are preventable tragedies that continue to occur, and in fact have increased in number in recent years [1]. In the US, suicide rates increased approximately 36% between 2000–2021. Suicide was responsible for 48,183 deaths in 2021, which is about one death every 11 minutes (CDC, 2023). The number of people who think about or attempt suicide is even higher. In 2021, an estimated 12.3 million American adults seriously thought about suicide, 3.5 million planned a suicide attempt, and 1.7 million attempted suicide (SAMSHA, 2022).

Our previous studies had pioneered the identification of blood gene expression biomarkers for suicidality [2–7], and other groups have validated this blood-based approach as well [8]. These gene expression studies are complementary to other genetic studies in the field [9–11], and in fact we integrate these different lines of work into our approach, as a convergent prioritization second step. We had also developed and tested a 22- item, simple to administer, quantitative risk evaluation and mitigation questionnaire called Convergent Functional Information for Suicidality (CFI-S), that focuses on social determinants and other known risk factors, and does not ask about current suicidal ideation [12, 13]. We wanted to expand upon those studies, as a way of deriving future scientific and practical insights, that would move these precision medicine approaches towards widespread utilization in clinical practice. First, we used larger cohorts of psychiatric patients for our biomarker studies (discovery, validation, and testing). Second, we used larger literature-derived databases for our convergent prioritization approaches, and the annotation of our biomarkers for drug modulatory effects. Third, we now have a

[1]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA. [2]MindX Sciences, Indianapolis, IN, USA. [3]University of Southern California, Los Angeles, CA, USA. [4]IT Core, Indiana University, Indianapolis, IN, USA. [5]Present address: VA Medical Center, Indianapolis, IN, USA. [6]Marion County Coroner's Office, Indianapolis, USA. [7]INBRAIN, Indianapolis, IN, USA. [8]Scripps Health, La Jolla, CA, USA. [9]Stark Neuroscience Research Institute, Indiana University School of Medicine, Indianapolis, IN, USA. [10]Present address: Department of Psychiatry, University of Arizona College of Medicine, Phoenix, AZ, USA. [11]Present address: Office of the Dean, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [12]These authors contributed equally: R. Bhagar, S. S. Gill, H. Le-Niculescu, C. Yin. ✉email: aniculescu@arizona.edu

longer duration follow-up on subjects, and also used longitudinal, not just cross-sectional, calculations for predictions. Fourth, we used newer methodologies, such as RNA sequencing, as well as machine-learning [14], to generate and analyze data. Fifth, we derived from the data a deeper biological, clinical, and therapeutical understanding of suicidality. Lastly, we show examples of practical applications, including bio-socio-psychological integration. We propose that such approaches could and should be used in clinical practice, to stem and reverse the tide of suicides.

MATERIALS AND METHODS
Ethics approval and consent to participate
All methods were performed in accordance with the relevant guidelines and regulations.

All subjects understood and signed informed consent forms detailing the research goals, procedure, caveats and safeguards, per Indiana University IRB approved protocol ("Mood State and Other Biomarkers: a Discovery-Based Approach", protocol no. 1011004024). No identifiable information or images are included in the publication.

Cohorts
We used three independent cohorts: discovery (a live psychiatric subjects cohort), validation (a postmortem coroner's office suicide completers cohort), and testing (an independent live psychiatric subjects test cohort for predicting suicidal ideation, and for predicting future hospitalizations for suicidality) (Fig. 1A).

Similar to our previous studies [2–4, 6], the live psychiatric subjects are part of a larger longitudinal cohort of adults that we are continuously collecting, the Indy500+ cohort. Subjects are recruited from the patient population at the Indianapolis VA Medical Center and Indiana University School of Medicine. All subjects understood and signed informed consent forms detailing the research goals, procedure, caveats and safeguards, per Indiana University IRB approved protocol ("Mood State and Other Biomarkers: a Discovery-Based Approach", protocol no. 1011004024). Subjects completed up to nine testing visits, 3–6 months apart or whenever a new psychiatric hospitalization occurred. At each testing visit, they received a series of psychiatric rating scales, including the Convergent Functional Information for Suicidality (CFI-S) and the Hamilton Rating Scale for Depression-17, which includes a suicidal ideation (HAMDSI) rating item (Fig. S1), and their blood was drawn. We collected whole blood (5 ml) in two RNA-stabilizing PAXgene tubes, labeled with an anonymized ID number, and stored at -80 degrees C in a locked freezer until the time of future processing. Whole-blood RNA was extracted for microarray and RNA sequencing gene expression studies from the PAXgene tubes, as detailed below.

For this study, our within-subject discovery cohort, from which the biomarker data were derived, consisted of 90 subjects (69 males, 21 females) with psychiatric disorders and multiple testing visits, who each had at least one diametric change in SI scores from no SI to high SI (HAMD-SI ≥ 2), or vice versa, from one testing visit to another. There were 4 subjects with 6 visits each, 8 subjects with 5 visits each, 7 subjects with 4 visits each, 39 subjects with 3 visits each, and 32 subjects with 2 visits each resulting in a total of 273 blood samples for subsequent gene expression studies (Fig. 1, Table S1).

Our postmortem cohort (n = 101), in which the top biomarker findings were validated for behavior, consisted of 83 male and 18 female suicide completers obtained through the Marion County coroner's office (Table S1). We required a last observed alive postmortem interval of 24 h or less, and the majority of cases selected completed suicide by means other than overdose, which could potentially affect gene expression. 59 subjects completed suicide by gunshot head or chest, 27 by hanging, 4 by jumping, 2 by slit wrist, 2 by train, 1 by asphyxiation, 1 by electrocution, and only 2 by overdose, 2 by carbon monoxide poisoning. Next of kin signed informed consent at the coroner's office for donation of blood for research.

Our independent test cohort for predicting suicidal ideation (Table S1) consisted of 330 male and 76 female subjects with psychiatric disorders, with one or multiple testing visits in our lab, with either no SI, intermediate SI, or high SI, resulting in a total of 820 blood samples in which whole-genome blood gene expression data were obtained (Fig. 1, Table S1).

Our test cohort for predicting future hospitalizations within one year of testing (Fig. 1, Table S1) is a subset (279 males, 47 females) of the

independent test cohort for which we had longitudinal follow-up with electronic medical records for at least 365 days resulting in a total of 685 samples. The subjects' subsequent number of psychiatric hospitalizations, with or without suicidality (planning, ideation or attempt), was tabulated from electronic medical records. Subjects were evaluated for the presence of future hospitalizations for suicidality within one year. A hospitalization was deemed to be without suicidality if suicidality was not listed as a reason for admission, and no SI was described in the admission and discharge medical notes. Conversely, a hospitalization was deemed to be due to suicidality if suicidal acts or intent were listed as a reason for admission, and/or SI was described in the admission and discharge medical notes.

Our test cohort for predicting all future hospitalizations (Fig. 1, Table S1) is a subset (310 males, 55 females) of the independent test cohort for which we also had longitudinal follow-up with electronic medical records resulting in a total of 745 samples. The subjects' subsequent psychiatric hospitalizations were tabulated in the same way as for hospitalizations in the first year, but for all time following their testing visit. Subjects also were evaluated for the presence of future hospitalizations for suicidality, and for the frequency of such hospitalizations.

Medications. The subjects in the discovery cohort were all diagnosed with various psychiatric disorders and had various medical co-morbidities. Their medications were listed in their electronic medical records and documented by us at the time of each testing visit. Medications can have a strong influence on gene expression. However, there was no consistent pattern of any particular class of medication. Our subjects were on a wide variety of different medications, psychiatric and non-psychiatric. Furthermore, the independent validation and testing cohort's gene expression data was Z-scored by gender before being combined, to normalize for any such effects. Some subjects may be non-compliant with their treatment and may have changes in medications or drugs of abuse not reflected in their medical records. Our goal is to find biomarkers that track suicidality, regardless of if the reason for it is due to internal biology or driven by exogenous substances or medication. In fact, one would expect some of these biomarkers to be targets of medications, as we show in this paper. Furthermore, the prioritization step that occurs after discovery is based on a field-wide convergence with literature that includes genetic data, which are unrelated to medication effects. Overall, the discovery, validation, and replication by testing in independent cohorts of the biomarkers, with our design, occurs despite the subjects having different genders, diagnoses, being on various medications, and other variables.

Blood gene expression experiments
RNA extraction. Whole blood (2.5 ml) was collected into each PaxGene tube by routine venipuncture. PaxGene tubes contain proprietary reagents for the stabilization of RNA. Total RNA was extracted and processed as previously described [2–4, 6].

Microarrays. Microarray work was carried out using previously described methodology on a subset of subjects (n = 794) [2–4, 6].

Of note, all genomic data was normalized (RMA for technical variability, then z-scoring for biological variability—by gender), before being combined and analyzed.

RNA sequencing: Next-generation RNA sequencing studies were carried out on the rest of the subject samples collected more recently (n = 248). We then endeavored to integrate the two platforms.

Biomarker analyses
Step 1: Discovery. Analyses were done in separately for Affymetrix data and for RNAseq data. Results were then integrated for a final discovery score.

For the Affymetrix dataset, a DE analysis was ran on the probeset level generating a raw score for each marker, and points were given when the probeset expression accurately corresponded to any changes in suicidality (from no to high, or high to no). The DE analysis detects how gradual changes in the gene expression track suicidality.

For the RNAseq dataset, a DE analysis was ran on the transcript level. Similarly, this analysis created a raw score for each transcript which corresponded with its ability to track suicidality. We substituted for the RNAseq data the 0 values with the next lowest non-zero values, so that the fold change calculations from visit to visit did not break in case of two

consecutive zeros. Ensembl 110 was then used to match transcripts to probeset(s). The average score from all transcripts corresponding to a probeset was generated. We only carried forward probesets that mapped to transcripts ($n = 46,600$), thus discarding probesets that did not map to transcripts as not being suitable for future practical applications, and transcripts that did not have probsets ($n = 53,002$) due to the hybrid integrated nature of our current approach. We anticipate that in the future when our studies consist of RNAseq data only, to not have to discard any transcripts and thus have a richer dataset.

Within each probeset, the score for each probeset from the microarray discovery cohort was summed with the average score of the transcripts corresponding to each probeset from the RNA seq discovery cohort, creating a final score for each probeset. A negative score indicated a decrease in expression biomarker, and a positive score an increase in expression biomarker.

A value percentile (separate for increased and decreased markers) was then assigned to each probeset based on its final score. The percentile scores were given points by thresholds (≥ 80% - 6pts; ≥50% - 4pts, ≥33.3% - 2 pt, <33.3% - 0pts) (Fig. 1). Only probesets with a greater than 2 pts. move on to the next step, Prioritization.

Step 2: Prioritization
Probeset to gene mapping for CFG prioritization step. After the discovery step, we mapped probesets to corresponding genes through Ensembl 110. The majority of probesets map to a single gene. However, some probesets map to multiple genes. Each probeset- gene pairing was carried out separately throughout the subsequent analyses. Probesets that did not have a gene match from Ensembl were further queried through the NetAffy database. Furthermore, if multiple or no genes matches were identified with NetAffy, those probesets were run through UCSC Genome Browser (https://genome.ucsc.edu/). Finally, all genes thus identified as corresponding to probesets were put through the GeneCards database to ensure the gene symbols were contemporary and up to date.

Databases: We have established in our laboratory (Laboratory of Neurophenomics, www.neurophenomics.info) manually curated databases of the human gene expression/protein expression studies (postmortem brain, peripheral tissue/fluids: CSF, blood and cell cultures), human genetic studies (association, copy number variations and linkage), and animal model gene expression and genetic studies, published to date on psychiatric disorders. Only findings deemed significant in the primary publication, by the study authors, using their particular experimental design and thresholds, are included in our databases. Our databases include only primary literature data and do not include review papers or other secondary data integration analyses to avoid redundancy and circularity. We also favored unbiased discovery studies over candidate genes hypothesis-driven studies. These large and constantly updated databases have been used in our CFG cross validation and prioritization platform (Fig. 1).

Data from 551 papers on suicidality were present in the databases at the time of the CFG analyses (human genetic studies-231, human brain studies-204, human peripheral tissue/fluids-116). We have developed in our lab a computerized CFG Wizard to automate and score in bulk large lists of genes by integrating evidence from these large databases, checked against manual scoring. Analyses were performed as previously described. Points were assigned to each gene (Human Brain Expression Evidence – 6pts, Human Peripheral Expression Evidence– 4pts, Human Genetic Evidence -2pts). Points were added to the Discovery score (0-6). There were $n = 2436$ markers with a combined score (discovery + prioritization) of 6 or above, which were progressed to the next step, Validation (Fig. 1).

Step 3: Validation
Affy and RNAseq data integration. Validation analyses were ran on a hybrid dataset of microarray and RNAseq data compiled into one. For the RNAseq samples, we used a sum of transcript TPM approach for all the transcripts corresponding to each probeset. This *virtual probeset* metric represents the equivalent of the expression of the probeset, allowing for adequate integration of RNAseq and microarray data into one dataset.

Housekeeping gene selection and data normalization
To normalize and account for technical variance, as well as for potential post-mortem effects in the postmortem Validation cohort, a housekeeping gene was selected to use for normalization. We compiled a list of the most used candidate housekeeping genes in the literature and all their corresponding probesets ($n = 35$). Not all classic housekeeping genes are "housekeeping", i.e. invariant or biologically not involved in the disorder, depending on the phenotype. An example of that is GAPDH [15]. So an empirical approach for each disease/phenotype/tissue is needed. The probeset with the lowest score after Discovery and integration of Affymetrix and RNAseq data (indicating the most invariance) was used as our housekeeping probeset (ACTB-224594_x_at). For the RNAseq samples, the TPM counts from all 14 transcripts of this probeset were summed up to generate a virtual probeset with a single value.

The microarray expression data probesets and the RNAseq expression data virtual probesets were divided by the housekeeping probeset/ virtual probeset expression levels.

Scaling factor
The final step to integrate Affymetrix and RNAseq data involves a scaling factor generation between the two platforms, to be able to bring data quantitatively to the same level/scale. Separate scaling factors were created for males and females. This process was done by taking the average expression for each probeset in the microarray data, as well as for each corresponding virtual probeset in the RNAseq data. The microarray average was then divided by the RNAseq average to generate a scaling factor. The RNAseq expression levels of each individual virtual probeset in each individual subject sample were then multiplied by their corresponding gender scaling factor.

$$\text{Scaling Factor for each probeset(by gender)} = \frac{\text{Microarray Average(Intensity)}}{\text{RNAseq Average(TPM)}}$$

$$\text{Scaling Factor} \times \text{Virtual Probeset Expression(TPM)in a subject} = \text{Scaled Expression}$$

This equation depicts how the Scaling Factor converts RNAseq expression TPM counts into the same, intensity-based levels, as microarray data.

Validation analysis
Three groups were used for Validation Analyses: the No SI (HAMD-SI = 0) and High SI (HMAD-SI ≥ 2) groups from the Discovery cohort, along with the suicide completers from the Coroner's Office.

Expression levels from the hybrid dataset of samples run on Affymetrix (probesets) and RNAseq (virtual probesets) were normalized using the housekeeping gene and scaling factors, and then z-scored by gender across the 3 groups. We carried out an ANOVA in the biomarkers that were stepwise changed in expression from No SI to High SI to Suicide Completion. Biomarkers that survived Bonferroni correction for number of biomarkers tested received 6 points, those nominally significant 4 points, and those that were just stepwise 2 points. The rest were 0 points. These points contributed to each marker's CFE3 score (discovery + prioritization + validation) (Fig. 1).

Top candidate biomarkers (after the first 3 steps)
Adding the scores from the first three steps into an overall convergent functional evidence (CFE) score (Fig. 1), we ended up with a list of 2340 top candidate biomarkers for suicidality that had a CFE score greater than 8 (1/3 of the possible maximum score of 24 after the first 3 steps). These top candidate biomarkers were carried forward into additional testing for clinical utility (Step 4).

Testing for clinical utility in independent cohorts
We tested in independent cohorts of psychiatric patients the ability of each of the top candidate biomarkers to assess state severity (measured by HAMD-SI ≥ 2), and predict trait risk (future hospitalizations for suicidality in the first year of follow-up, and in all future years of follow-up). We conducted our analyses across all patients, as well as personalized by gender.

The test cohorts for predicting suicidality severity (state), and the test cohorts for predicting future hospitalizations with suicidality (trait), were assembled out of hybrid datasets of Affymetrix and RNAseq samples, that were integrated as described above for the Validation step (housekeeping gene normalization, scaling factor and Z-scoring by gender). The cohorts were completely independent from the discovery and validation cohorts, there was no subject overlap with them. Individual biomarkers used for predictions were normalized as described above to avoid potential
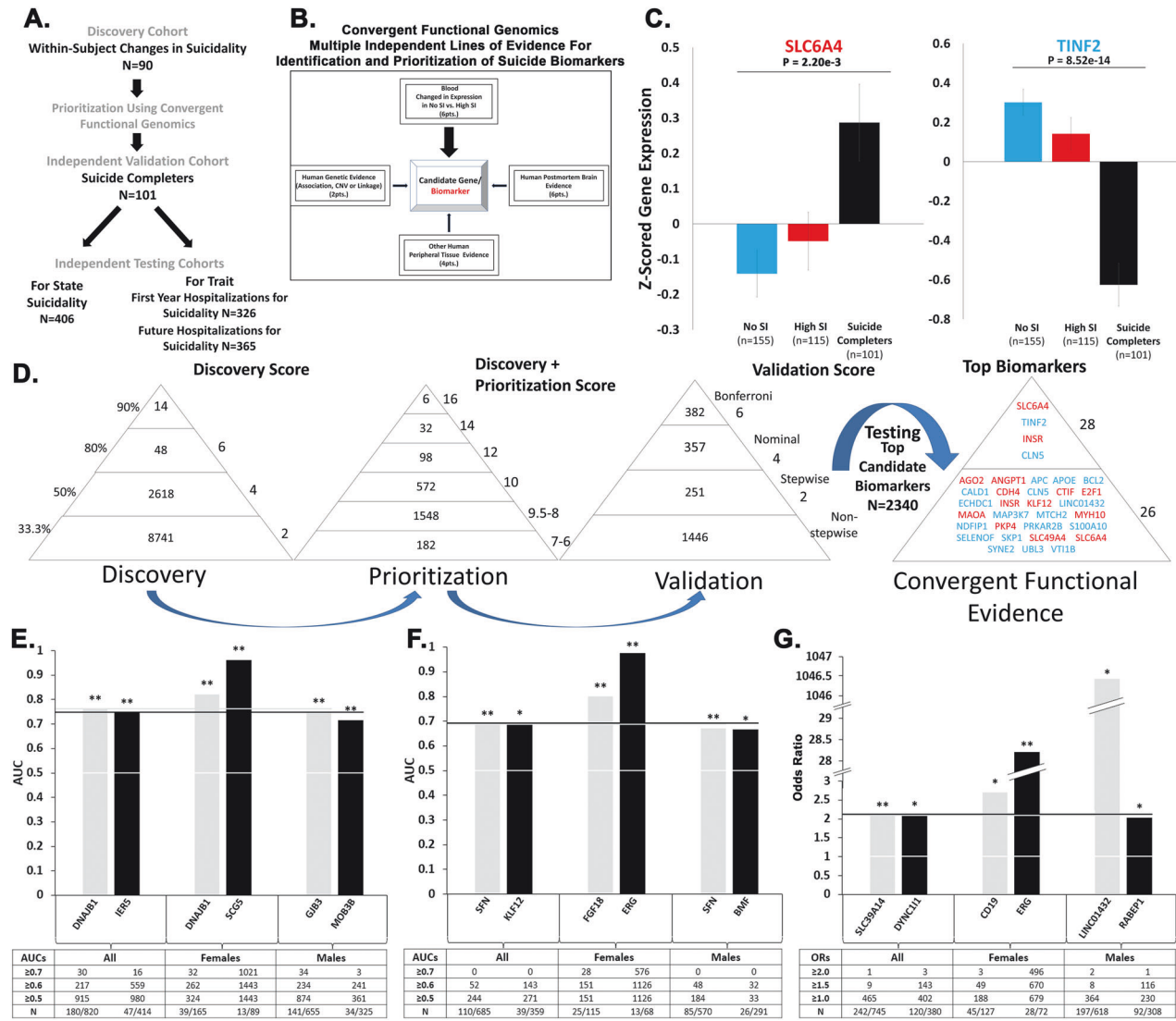
Fig. 1 Steps 1-3: discovery, prioritization, validation and testing of biomarkers for suicidality. A Cohorts used in study, depicting flow of discovery, prioritization, and validation of biomarkers from each step. **B** Prioritization using Convergent Functional Genomics (CFG). **C** Validation -biomarkers are assessed for stepwise change from discovery subjects with no symptoms, high symptoms to the validation subjects where samples were collected from suicide completers, using ANOVA. The histograms depict a top increased (I) and a top decreased biomarker (**D**). Number of probesets and scoring at each of the Steps. Step 1 -Discovery probesets are identified based on their score for tracking symptoms and ranked 33.3% (2 pt), 50% (4 pt) and 80% (6 pt). Step 2- Prioritization with CFG for prior evidence of involvement in Suicidality. Maximum of 6 pt. Genes scoring at least 6 pt out of a maximum possible of 18 pt after Discovery and Prioritization are carried forward to the validation step. Step 3- Validation in an independent cohort of suicide completers. We selected the top CFE score ≥8 ($n = 2340$) for further testing and characterization. **E** Predictions for State—High Suicidality. Top cross-sectional and longitudinal markers are shown in all subjects, males, and females. Table below displays number of significant markers within each prediction group by AUC. **F** Predictions for Trait —Hospitalizations in the First Year. Top cross-sectional and longitudinal markers are shown in all subjects, males, and females. Table below displays number of significant markers within each prediction group by AUC. **G** Predictions for Trait—All Future Hospitalizations. Top cross-sectional and longitudinal markers are shown in all subjects, males, and females. Table below displays number of significant markers within each prediction group by odds ratio.

artefacts due to different ranges of expression by gene expression platform and gender, and to be able to combine different biomarkers into panels. Predictions were performed using R-studio. For cross-sectional analyses, we used biomarker expression levels. For longitudinal analyses, we combined four measures: biomarker expression level, slope (defined as ratio of levels at current testing visit vs. previous visit, divided by time between visits), maximum levels (at any of the current or past visits), and maximum slope (between any adjacent current or past visits), as described in previous studies [16–18]. For decreased biomarkers, we used the minimum rather than the maximum for level calculations.

Predicting state-suicidality severity. Receiver-operating characteristic (ROC) analyses between marker levels and suicidality state were performed by assigning subjects visits with a HAMD-SI score ≥2 in the high suicidality category vs. the rest of the subjects in this independent test cohort (406 subjects, 820 visits). We used the pROC package of R (Table 1 and Fig. 1). Additionally, a one-tailed *t*-test was performed between high suicidality group vs. the rest, and Pearson R (one-tail) was calculated between suicidality scores and biomarker levels.

Predicting trait- future psychiatric hospitalization due to suicidality as a symptom/reason for admission. We conducted analyses for predicting future psychiatric hospitalizations with suicidality as a symptom/reason for admission in the first year following each testing visit, in subjects that had at least one year of follow-up in the VA system, in which we have access to complete electronic medical records (326 subjects, 685 visits). ROC

Table 1. Top biomarkers after 4 Steps of convergent functional evidence (CFE).

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/ p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
SLC6A4 Solute Carrier Family 6 Member 4	241811_x_at	(I) 4 76.08%	12	2.21E−03/ 4	**All** L: (47/414) 0.59/2.44E−02 **Gender** Females L: (13/89) 0.68/2.22E−02	**Gender** Females L: (13/68) 0.79/6.75E−04	**Gender** Females L: (28/72) 2.71/2.16E−03	Addictions Alcohol Anxiety Depression Pain Stress **Distant Output**	Agomelatine Citalopram Duloxetine Fluoxetine Imipramine Lamotrigine Omega-3 fatty acids Probiotic PS128 Sertraline Vortioxetine	28
TINF2 TERF1 Interacting Nuclear Factor 2	220052_s_at	(D) 4 56.87%	10	8.52E−14/ 6	**Gender** Females C: (39/165) 0.59/3.92E−02 Females L: (13/89) 0.72/4.95E−03	**Gender** Females L: (13/68) 0.67/3.12E−02	**All** C: (242/745) 1.2/4.73E−03 **Gender** Males C: (197/618) 1.18/1.55E−02	Alcohol Dementia Depression Stress SZ SZA		28
INSR Insulin Receptor	226216_at	(I) 4 52.67%	8	1.15E−04/ 4	**All** L: (47/414) 0.6/1.38E−02 **Gender** Males L: (34/325) 0.61/1.57E−02	**All** L: (39/359) 0.58/4.68E−02	**All** L: (120/380) 1.38/2.22E−03 **Gender** Females L: (28/72) 1.86/1.83E−02 Males L: (92/308) 1.36/1.07E−02	Addictions Aging Anxiety BP Dementia Depression Stress SZ	Antidepressants Lithium Valproate	28
CLN5 CLN5 Intracellular Trafficking Protein	214252_s_at	(D) 4 52.23%	6	6.23E−09/ 6	**All** C: (180/820) 0.57/1.52E−03 L: (47/414) 0.66/2.13E−04 **Gender**	**All** C: (39/359) 0.59/3.04E−02 **Gender** Females L:	**All** C: (242/745) 1.22/7.35E−03 L: (120/380) 1.85/4.03E−03 **Gender**	Alcohol Anxiety Dementia Depression Stress	Lithium	28

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
					Females L: (13/89) 0.74/3.17E−03 Males C: (141/655) 0.58/2.24E−03 Males L: (34/325) 0.62/8.83E−03	(13/68) 0.69/1.94E−02	Males C: (197/618) 1.23/1.18E−02 Males L: (92/308) 1.8/1.11E−02			
PKP4 Plakophilin 4	214874_at	**(I)** 4 65.65%	6	1.37E−03/ 4	**All** L: (47/414) 0.59/1.88E−02 **Gender** Females L: (13/89) 0.68/1.83E−02	**All** L: (39/359) 0.61/1.46E−02 **Gender** **Females** L: (13/68) 0.8/3.67E−04	**All** L: (120/380) 1.29/1.61E−02 **Gender** Females L: (28/72) 3.49/7.16E−08	Addictions Alcohol Dementia Depression Pain Psychosis Stress SZ		26
SLC49A4 Solute Carrier Family 49 Member 4	229213_at	**(I)** 4 64.89%	4	6.23E−06/ 6	**All** L: (47/414) 0.62/2.77E−03 **Gender** Females L: (13/89) 0.65/3.78E−02 Males L: (34/325) 0.61/1.73E−02	**All** L: (39/359) 0.65/1.04E−03 **Gender** Females L: (13/68) 0.75/2.44E−03 Males L: (26/291) 0.61/3.69E−02	**All** L: (120/380) 1.33/1.68E−03 **Gender** Females L: (28/72) 3.54/9.74E−05 Males L: (92/308) 1.25/2.89E−02	Aging Alcohol Dementia Depression	Indomethacin	26
SKP1 S-Phase Kinase Associated Protein 1	200719_at	**(D)** 4 63.19%	10	0	**All** C: (180/820) 0.58/9.11E−04 L: (47/414) 0.62/3.13E−03 **Gender**	**All** C: (110/685) 0.57/8.40E−03 L: (39/359) 0.61/1.24E−02 **Gender**	**All** C: (242/745) 1.15/2.51E−02 **Gender** Females L:	Aging Alcohol Anxiety Dementia Depression Stress SZ	Clozapine Lithium	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
					Females L: (13/89) 0.77/8.21E−04 Males C: (141/655) 0.58/3.12E−03	Females L: (13/68) 0.78/9.81E−04 Males C: (85/570) 0.57/1.93E−02	(28/72) 2.47/3.89E−02	Immediate Input		
ECHDC1 Ethylmalonyl-CoA Decarboxylase 1	223087_at	**(D)** 4 60.49%	6	7.84E−06/6	**All** C: (180/820) 0.58/9.61E−04 L: (47/414) 0.61/6.52E−03 **Gender** Females L: (13/89) 0.73/3.64E−03 Males C: (141/655) 0.57/3.96E−03	**Gender** Females L: (13/68) 0.66/3.71E−02	**All** C: (242/745) 1.15/3.33E−02 **Gender** Females C: (45/127) 1.35/4.98E−02	Addictions Alcohol BP Dementia Depression Pain Stress SZ	Chlorpromazine Ketamine	26
BCL2 BCL2 Apoptosis Regulator	203685_at	**(D)** 4 60.34%	10	1.74E−08/6	**All** C: (180/820) 0.56/7.51E−03 L: (47/414) 0.6/1.04E−02 **Gender** Females C: (39/165) 0.59/4.80E−02 Females L: (13/89) 0.77/8.21E−04 Males	**Gender** Females L: (13/68) 0.72/6.72E−03		Addictions Aging Alcohol BP Dementia Depression Neurological Pain Stress SZ	Amisulpride Bariatric Surgery Bupropion Cannabidiol (CBD) Chinese Sumac Clozapine CoQ 10 Curcumin Doxepin Duloxetine Fisetin Folic Acid and Omega-3 Haloperidol Lamotrigine Lithium	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/ p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
					C: (141/655) 0.55/2.83E−02				Olanzapine Omega-3 fatty acids Retinol Risperidone Sesamol Thymol Valproate Venlafaxine Wheat grass Xiao Yao San	
SELENOF Selenoprotein F	200902_at	**(D)** 4 56.33%	6	1.37E−05/ 6	**All** C: (180/820) 0.55/1.39E−02 L: (47/414) 0.59/1.84E−02 **Gender** Females L: (13/89) 0.76/1.36E−03 Males C: (141/655) 0.55/4.25E−02	**Gender** Females L: (13/68) 0.73/5.89E−03	**All** C: (242/745) 1.14/4.02E−02 L: (120/380) 1.52/2.76E−02 **Gender** Males L: (92/308) 1.5/4.71E−02	Addictions Stress	Atomoxetine	26
SYNE2 Spectrin Repeat Containing Nuclear Envelope Protein 2	202761_s_at	**(D)** 4 56.31%	12	0	**All** C: (180/820) 0.54/4.39E−02 L: (47/414) 0.58/4.66E−02 **Gender** Females L: (13/89) 0.67/2.84E−02	**Gender** Females C: (25/115) 0.63/2.75E−02 Females L: (13/68) 0.66/4.10E−02	**All** C: (242/745) 1.15/2.12E−02	Aging Alcohol BP Dementia Depression Pain PTSD Sleep Stress SZ	Ketamine Prednisolone	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
NDFIP1 Nedd4 Family Interacting Protein 1	222422_s_at	**(D)** 4 56.02%	6	3.46E−03/ 4	**All** C: (180/820) 0.56/6.85E−03 L: (47/414) 0.6/1.24E−02 **Gender** Females L: (13/89) 0.74/2.47E−03 Males C: (141/655) 0.57/6.98E−03	**All** C: (110/685) 0.55/3.98E−02 L: (39/359) 0.62/8.03E−03 **Gender** Females L: (13/68) 0.66/4.10E−02 Males C: (85/570) 0.59/4.73E−03	**All** C: (242/745) 1.14/4.41E−02 L: (120/380) 1.5/2.63E−02 **Gender** Males C: (197/618) 1.21/1.47E−02 Males L: (92/308) 1.59/2.00E−02	Addictions Aging Alcohol ASD BP Dementia Depression Stress SZ	Bariatric Surgery Carbamazepine Ketamine Omega-3 fatty acids	26
VTI1B Vesicle Transport Through Interaction With T-SNAREs 1B	225926_at	**(D)** 4 53.7%	6	2.29E−07/ 6	**All** L: (47/414) 0.6/1.44E−02 **Gender** Females L: (13/89) 0.71/8.90E−03	**Gender** Females L: (13/68) 0.67/2.70E−02	**All** C: (242/745) 1.13/4.57E−02 L: (120/380) 1.46/3.33E−02 **Gender** Males C: (197/618) 1.16/3.33E−02 Males L: (92/308) 1.45/4.90E−02	Addictions Aging Alcohol Anxiety ASD BP Dementia Pain Stress SZ	Magnesium Prednisolone	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
E2F1 E2F Transcription Factor 1	204947_at	(I) 4 53.44%	6	5.80E−10/ 6	**All** L: (47/414) 0.69/8.82E−06 **Gender** Females L: (13/89) 0.7/1.11E−02 Males L: (34/325) 0.67/4.83E−04	**All** L: (39/359) 0.63/3.40E−03 **Gender** Females L: (13/68) 0.79/5.14E−04	**Gender** Females C: (45/127) 1.3/3.58E−02 Females L: (28/72) 2.31/3.30E−05	Alcohol Dementia Pain Stress	American Ginseng Citalopram Escitalopram	26
CTIF Cap Binding Complex Dependent Translation Initiation Factor	204302_s_at	(I) 4 52.67%	6	5.05E−03/ 4	**All** L: (47/414) 0.65/2.86E−04 **Gender** Females L: (13/89) 0.73/4.79E−03 Males L: (34/325) 0.62/9.80E−03	**All** L: (39/359) 0.6/2.05E−02 **Gender** Females L: (13/68) 0.74/4.10E−03	**All** L: (120/380) 1.37/3.88E−03	Addictions Aging Alcohol Autism Dementia Depression Pain		26
MTCH2 Mitochondrial Carrier 2	217772_s_at	(D) 4 52.59%	8	1.31E−03/ 4	**All** L: (47/414) 0.58/2.96E−02 **Gender** Females L: (13/89) 0.68/2.16E−02 Males C: (141/655) 0.55/4.87E−02	**Gender** Females L: (13/68) 0.66/3.58E−02	**All** L: (120/380) 1.55/9.64E−03 **Gender** Females L: (28/72) 2.5/2.79E−02 Males L: (92/308) 1.42/4.75E−02	Addictions Alcohol BP Dementia Depression Stress SZ	Donepezil	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
PRKAR2B Protein Kinase CAMP-Dependent Type II Regulatory Subunit Beta	203680_at	**(D)** 4 52.12%	10	1.17E−06/ 6	**Gender** Females L: (13/89) 0.7/1.14E−02		**All** L: (120/380) 1.63/1.32E−03 **Gender** Females L: (28/72) 4.76/4.76E−04 Males L: (92/308) 1.46/2.24E−02	ADHD Aging Alcohol Anxiety ASD Dementia Depression PTSD Stress	Clozapine Valproate	26
ANGPT1 Angiopoietin 1	1552939_at	**(I)** 4 51.15%	8	1.98E−01/ 2	**All** L: (47/414) 0.64/1.14E−03 **Gender** Females L: (13/89) 0.74/2.51E−03 Males L: (34/325) 0.6/3.11E−02	**All** L: (39/359) 0.59/3.68E−02 **Gender** Females L: (13/68) 0.7/1.37E−02	**All** L: (120/380) 1.31/8.33E−04 **Gender** Females C: (45/127) 1.48/2.53E−03 Females L: (28/72) 2.53/2.24E−06	Addictions Aging Alcohol Anxiety Autism BP Dementia Depression Stress SZ	Ketamine Lactobacillus plantarum 299 v Lithium Omega-3 fatty acids Rofecoxib	26
KLF12 KLF Transcription Factor 12	206966_s_at	**(I)** 4 51.15%	6	1.07E−09/ 6	**All** L: (47/414) 0.63/2.26E−03 **Gender** Females L: (13/89) 0.65/4.17E−02 Males L: (34/325) 0.62/1.24E−02	**All** L: (39/359) 0.59/2.64E−02 **Gender** Females L: (13/68) 0.73/5.39E−03	**Gender** Females (28/72) 2.01/2.29E−03	Aging Alcohol Anxiety BP Dementia Pain Stress SZ	Indomethacin Rofecoxib Valproate Vortioxetine	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/ p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
CDH4 Cadherin 4	220227_at	(I) 4 50.38%	8	1.27E−02/ 4	**All** L: (47/414) 0.65/3.19E−04 **Gender** Females L: (13/89) 0.72/5.48E−03 Males L: (34/325) 0.62/1.10E−02	**All** L: (39/359) 0.6/1.69E−02 **Gender** Females L: (13/68) 0.76/1.81E−03	**Gender** Females L: (28/72) 2.25/4.95E−06	Alcohol Dementia Depression Stress	Lithium	26
APOE Apolipoprotein E	212884_x_at	(D) 2 43.75%	12	0	**All** C: (180/820) 0.74/1.30E−22 L: (47/414) 0.68/2.92E−05 **Gender** Females C: (39/165) 0.82/1.01E−09 Females L: (13/89) 0.88/7.35E−06 Males C: (141/655) 0.71/6.42E−15 Males L: (34/325) 0.61/1.75E−02	**All** C: (110/685) 0.67/9.04E−09 L: (39/359) 0.67/2.29E−04 **Gender** Females C: (25/115) 0.79/6.30E−06 Females L: (13/68) 0.88/1.07E−05 Males C: (85/570) 0.64/1.66E−05	**All** C: (242/745) 1.38/1.26E−05 L: (120/380) 1.38/1.00E−02 **Gender** Females C: (45/127) 1.48/1.48E−02 Females L: (28/72) 1.9/1.60E−03 Males C: (197/618) 1.37/1.20E−04	Aging Alcohol Autism Dementia Depression Neurological Pain Psychosis Stress SZ	Agaricus Blazei American Ginseng escitalopram Magnesium nortriptyline	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/ p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
MYH10 Myosin Heavy Chain 10	212372_at	(I) 2 41.77%	6	5.28E−14/ 6	**All** L: (47/414) 0.65/4.33E−04 **Gender** Females L: (13/89) 0.68/1.93E−02 Males L: (34/325) 0.62/9.55E−03	**All** L: (39/359) 0.66/6.31E−04 **Gender** Females L: (13/68) 0.88/1.32E−05	**All** L: (120/380) 1.38/1.19E−03 **Gender** Females L: (28/72) 3.26/2.09E−07	Alcohol BP Depression Pain SZ **Distant Output**	Clozapine Mianserin	26
UBL3 Ubiquitin Like 3	201535_at	(D) 2 40.74%	8	1.02E−05/ 6	**All** C: (180/820) 0.54/3.65E−02 L: (47/414) 0.61/5.57E−03 **Gender** Females L: (13/89) 0.7/1.07E−02 Males C: (141/655) 0.56/1.96E−02	**Gender** Females L: (13/68) 0.65/4.99E−02	**All** L: (120/380) 1.6/1.56E−02 **Gender** Males L: (92/308) 1.5/4.80E−02	Addictions Aging Alcohol Dementia Depression SZ		26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/ p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
CALD1 Caldesmon 1	215198_s_at	**(D)** 2 38.89%	12	0	**All** C: (180/820) 0.7/2.75E−16 L: (47/414) 0.61/6.20E−03 **Gender** Females C: (39/165) 0.67/7.83E−04 Females L: (13/89) 0.76/1.59E−03 Males C: (141/655) 0.7/7.12E−14	**All** C: (110/685) 0.62/2.35E−05 **Gender** Females L: (13/68) 0.71/8.69E−03 Males C: (85/570) 0.63/1.00E−04	**All** C: (242/745) 1.32/4.03E−04 **Gender** Females L: (28/72) 2.94/2.74E−04 Males C: (197/618) 1.4/1.63E−04	Alcohol Dementia Depression Memory formation Pain Stress **Distant Output**	Carbamazepine Clozapine Estradiol Ketamine	26
APC APC Regulator Of WNT Signalling Pathway	203525_s_at	**(D)** 2 38.89%	10	1.96E−07/ 6	**All** L: (47/414) 0.59/1.79E−02 **Gender** Females L: (13/89) 0.69/1.41E−02		**All** L: (120/380) 1.55/4.31E−03 **Gender** Males C: (197/618) 1.17/3.06E−02 Males L: (92/308) 1.63/5.78E−03	Aging Alcohol Anxiety BP Dementia Depression Stress SZ	Amisulpride Brexanolone Clozapine Echinacea Imipramine Lithium Risperidone Valproate	26
MAP3K7 Mitogen-Activated Protein Kinase Kinase Kinase 7	206853_s_at	**(D)** 2 38.8%	10	8.51E−04/ 4	**All** L: (47/414) 0.58/3.12E−02 Females L: (13/89) 0.74/3.28E−03	**Gender** Females L: (13/68) 0.67/3.12E−02	**All** L: (120/380) 1.52/3.04E−02	Aging Alcohol Anxiety BP Stress SZ	Lithium Nutramil Vortioxetine	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/Score Up to 6 pts	Step 4 Significant Predictions of High Suicidality State ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
MAOA Monoamine Oxidase A	212741_at	**(I)** 2 38.6%	12	2.84E−09/ 6	**Gender** Females L: (13/89) 0.65/4.07E−02	**Gender** Females L: (13/68) 0.74/3.17E−03	**Gender** Females L: (28/72) 2.42/1.21E−03	Addictions Alcohol Anxiety Autism BP Dementia Depression Pain Panic Disorder Stress SZ **Immediate Output**	Amphetamine Black Olive Extract Citalopram Fluoxetine Green Olive Extract Indantadol Ketamine PHENELZINE Psychotherapy SELEGILINE tranylcypromine Zonisamide	26
LINC01432 Long Intergenic Non-Protein Coding RNA 1432	1563246_at	**(D)** 2 35.19%	6	9.27E−07/ 6	**All** C: (180/820) 0.74/5.49E−23 L: (47/414) 0.68/4.53E−05 **Gender** Females C: (39/165) 0.78/2.70E−08 Females L: (13/89) 0.84/5.25E−05 Males C: (141/655) 0.72/1.28E−16 Males L: (34/325) 0.61/1.41E−02	**All** C: (110/685) 0.69/2.98E−10 L: (39/359) 0.59/2.71E−02 **Gender** Females C: (25/115) 0.78/6.99E−06 Females L: (13/68) 0.81/2.78E−04 Males C: (85/570) 0.64/1.02E−05	**All** C: (242/745) 1.87/5.51E−04 **Gender** Females C: (45/127) 1.7/1.96E−03 Females L: (28/72) 1.73/1.68E−02 Males C: (197/618) 1046.47/6.49E −05		escitalopram	26

Table 1. continued

Symbol/Gene Name	Probeset ID Affy data	Step 1 Discovery (Direction of Change in High Suicidality) Score % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence For Involvement in Suicidality Score Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Significant Predictors of High Suicidality State ROC AUC/ p-value 4 pts All 2pts Gender	Step 4 Significant Prediction of First Year Hosp. for Suicidality ROC AUC/p-value 4 pts All 2pts Gender	Step 4 Significant Predictions of Future Hosp for Suicidality OR/p-value 4 pts All 2pts Gender	Other Psychiatric and Related Disorders Evidence in Same Direction to High Suicidality Role in circadian clock	Drugs that Modulate the Biomarker in Opposite Direction to High Suicidality	CFE Polyevidence Score for Involvement in Suicidality (Based on Steps 1-4)
S100A10 S100 Calcium Binding Protein A10	200872_at	**(D)** 2 34.88%	12	0	**All** C: (180/820) 0.6/1.96E−05 L: (47/414) 0.59/1.76E−02 **Gender** Males C: (141/655) 0.62/1.29E−05	**All** C: (110/685) 0.56/2.85E−02 **Gender** Males C: (85/570) 0.58/8.53E−03	**All** C: (242/745) 1.13/3.59E−02 **Gender** Males C: (197/618) 1.16/2.19E−02	ADHD Alcohol Anxiety BP Depression Pain Stress SZ **Distant Output**	American Ginseng Antidepressants Carbamazepine Citalopram Clozapine Fluoxetine Imipramine Mirtazapine Psilocybin S-adenosyl methionine (SAM) Vortioxetine	26
AGO2 Argonaute RISC Catalytic Component 2	1561042_at	**(I)** 2 34.35%	10	1.70E−16/ 6	**All** L: (47/414) 0.61/7.69E−03 **Gender** Females L: (13/89) 0.66/2.99E−02	**Gender** Females L: (13/68) 0.8/4.60E−04	**Gender** Females L: (28/72) 3/1.92E−05	Addictions Alcohol Autism BP Dementia Depression Stress SZ	Indomethacin Magnesium Risperidone	26

Top CFE biomarkers (*n* = 30 genes) after Step 4 Testing in independent cohorts for state and trait predictive ability. For Step 4 Predictions, **C:** -cross-sectional (using levels from one visit), **L:** -longitudinal (using levels and slopes from multiple visits). In All and personalized by Gender, M-Males, F-Females. For Step 4 predictions scoring, 4 pts if significant in All, 2pts if significant in gender, 4 pts if significant in All, 2pts if significant in gender, capped at a maximum of 4pts. Underlined- best predictive biomarker in that category, gender and methodology.
BP Bipolar, *SZ* schizophrenia, *SZA* schizoaffective disorder.

analyses between biomarkers measures (cross-sectional, longitudinal) at a specific testing visit and future hospitalizations were performed as described above, based on assigning if subjects had been admitted to the hospital with suicidality or not. Additionally, a one tailed t-test with unequal variance was performed between groups of subject visits with and without future hospitalization with suicidality. Pearson R (one-tail) correlation was performed between hospitalization frequency (number of hospitalizations with suicidality divided by duration of follow-up) and marker levels. A Cox regression was performed using the time in days from the testing visit date to first hospitalization date in the case of patients who had been hospitalized, or 365 days for those who did not. The odds ratio was calculated such that a value greater than 1 always indicates increased risk for hospitalization, regardless if the biomarker is increased or decreased in expression.

We also conducted Cox regression and Pearson R analyses for all future hospitalizations with suicidality (365 subjects, 745 visits), including those occurring beyond one year of follow-up (up to 17.2 years, average: 7.8 years). The Cox regression was performed using the time in days from visit date to first hospitalization date in the case of patients who had hospitalizations with suicidality, or from visit date to last note date in the electronic medical records for those who did not. These calculations, unlike the ROC and t-test, account for the actual length of follow-up, which varied from subject to subject. The ROC and t-test might in fact, if used, under-represent the power of the markers to predict, as the more severe psychiatric patients are more likely to move geographically and/or be lost to follow-up.

Step 4 Predictions Scoring

Biomarkers that are nominally significant (for ROC AUC for State and First Year hospitalizations predictions, Cox Regression Odds Ratio for All Future Hospitalizations predictions) receive 4 points if they are predictive in all subjects in the cohort, and 2 points if they are only predictive within a gender. Scores are capped at 4, and the maximum score between cross-sectional and longitudinal predictions for each biomarker is taken moving forward. These points are then added to each biomarker's CFE score to create a final score (discovery + prioritization + validation + state predictions + trait predictions) indicative of each markers ability to track and predict suicidality (Table 1). The maximum possible CFE4 score is 36 $(6 + 12 + 6 + 12)$. We display in Table 1 the top biomarkers for suicidality with a CFE4 score of 26 and above ($n = 30$), chosen so the marker cannot have just maximum evidence from the first three steps, it has to have some evidence from Step 4 also. (Fig. 1, Table 1). This overall score provides a degree of certainty that the high-scoring biomarkers are indeed involved in the disease, while the performance might be better if the biomarkers were discovered, validated and tested separately by gender [3, 4]. In view of that, for the reports (Fig. 2), we chose a panel of the best predictive markers by gender (Table S4B, C).

Biological understanding

Pathway analyses. IPA (Ingenuity Pathway Analysis, version 107193442, Qiagen) and DAVID Functional Annotation Analysis (National Institute of Allergy and Infectious Diseases) (v2023q3) were used to analyze the biological roles, including top canonical pathways and diseases (Table 2A, B). We performed the pathway analyses for the 30 biomarkers for suicidality that were the top scoring CFE biomarkers after discovery, prioritization, validation, and testing.

Networks. For network analyses we performed STRING Interaction Network (https://string-db.org) by inputting the 30 genes into the search window, and performed Multiple Proteins Homo sapiens analysis (Fig. S3).

CFG beyond suicide: evidence for involvement in other psychiatric and related disorders. We also used a CFG approach to examine evidence from other psychiatric and related disorders, as exemplified for the list of top biomarkers after Steps 1- 4 (Table S3). This was not used to prioritize genes, but rather to understand the molecular basis of co-morbidities. We also calculated genomic co-morbidities % based on number of genes on our list that matched to different other disorders (Table 2D).

Therapeutics

Pharmacogenomics. We analyzed which of the top biomarkers for suicidality after Steps 1–4 ($n = 30$) are known to be changed in expression by existing drugs in a direction opposite to the one in disease, using our CFG databases (Table S4). These drugs and nutraceuticals are potential treatments and preventatives for patients with suicidality, and used in the prototype reports (described below) to demonstrate personalized medicine. Drugs are also listed individually by biomarker affected (Table 1 and S4).

Drug repurposing using the connectivity map. Following biomarker identification and validation, Connectivity Map was used in order to identify potential pharmaceuticals to alter the gene expression signature of the top biomarkers in a manner that opposes their alteration in suicidality. A Connectivity Map Query was performed using the selected top biomarkers ($n = 30$), performed using L1000 parameters with the latest dataset [19]. Results were converted into a matrix using cMapR [20]. The results from the query were analyzed and sorted based on normalized connectivity score. Drugs that are experimental were removed (Table S7).

Report generation

We present examples of how reports to doctors might look, using the above insights. We chose as case studies two patients who were tested by us, and who we learned subsequently had died by suicide. We used a panel of the top predictive biomarkers for state and trait, by gender.

Step 5 - Generalizability. All biomarkers that were nominally significant in the predictions in Step 4 Testing were retested for predictive ability on the whole database ($n = 1127$), consisting of male ($n = 893$) and female ($n = 234$) groups. Subsequently, to create our biomarker panel for the reports, for each gender we took the 12 best predictive biomarkers for state, first year hospitalizations, and all future hospitalization, resulting in a male and female panel of 36 biomarkers each.

Scores generation. The raw expression values of the biomarkers in our whole gene expression dataset ($n = 1127$) were Z-scored by gender. For state score, the Z-scored expression value of each increased biomarker was compared to the average value for the biomarker in the high suicidality group in the database, resulting in scores of 1 or 0 respectively, and 0.5 if it is in between. The reverse was done for decreased biomarkers. For trait chronic risk score, we calculated the average expression value for a biomarker in the first-year hospitalizations for suicidality group, and in the not hospitalized in the first-year group, and for all hospitalizations for suicidality group, and no future hospitalizations group. We then compared the biomarkers for the subject of interest to these reference levels. If a biomarker was higher than the average of the high group it got a 1, if it was below the average of the no group it got a 0, and if it was in between, it got a 0.5 for increased biomarkers. For decreased biomarkers, if it was lower than the average of the high group it got a 1, if it was higher than the average of the no group it got a 0, and if it was in between got a 0.5. These digitized scores for each biomarker are multiplied by the CFE4 score of each biomarker as a weight, to account for the totality of evidence, then summed into a polygenic risk score and then divided by a sum of all CFE4 scores.

The suicidality state risk score is the average score of all the state biomarkers multiplied by 100, generating 3 risk categories: high (red), intermediate (yellow), and low (green). The chronic suicidality risk score was calculated the same way using biomarkers for first year and for all future hospitalizations due to suicidality. These percentile scores of the patient are provided in the report (Fig. 2).

The digitized biomarkers are also used for matching with existing psychiatric medications and alternative treatments (nutraceuticals and others). We use our large datasets and literature databases to match biomarkers to medications that have effects on gene expression opposite to their expression in high suicidality. The gene expression data is from gene expression data in human and animal models. Each medication matched to a biomarker gets the biomarker's score of 1, 0.5 or 0. The scores for the medications are added, and divided by the number of biomarkers that were 1 or 0.5 in that patient, resulting in a percentile match. Thus, psychiatric medications are matched to the patient and ranked in order of impact on the panel.

Hierarchical clustering and subtypes

A two-way unsupervised hierarchical clustering (Fig. S2) was done using subjects in the discovery cohort with high suicidal ideation (HAMD-SI ≥ 2, $n = 103$). Clustering was done on 4 psychiatric dimensions, using quantitative instruments: stress (SSS4) [21], anxiety (SAS4) [17], mood (SMS7) [16], psychosis (PANSS Positive) [18]. Subjects measures were
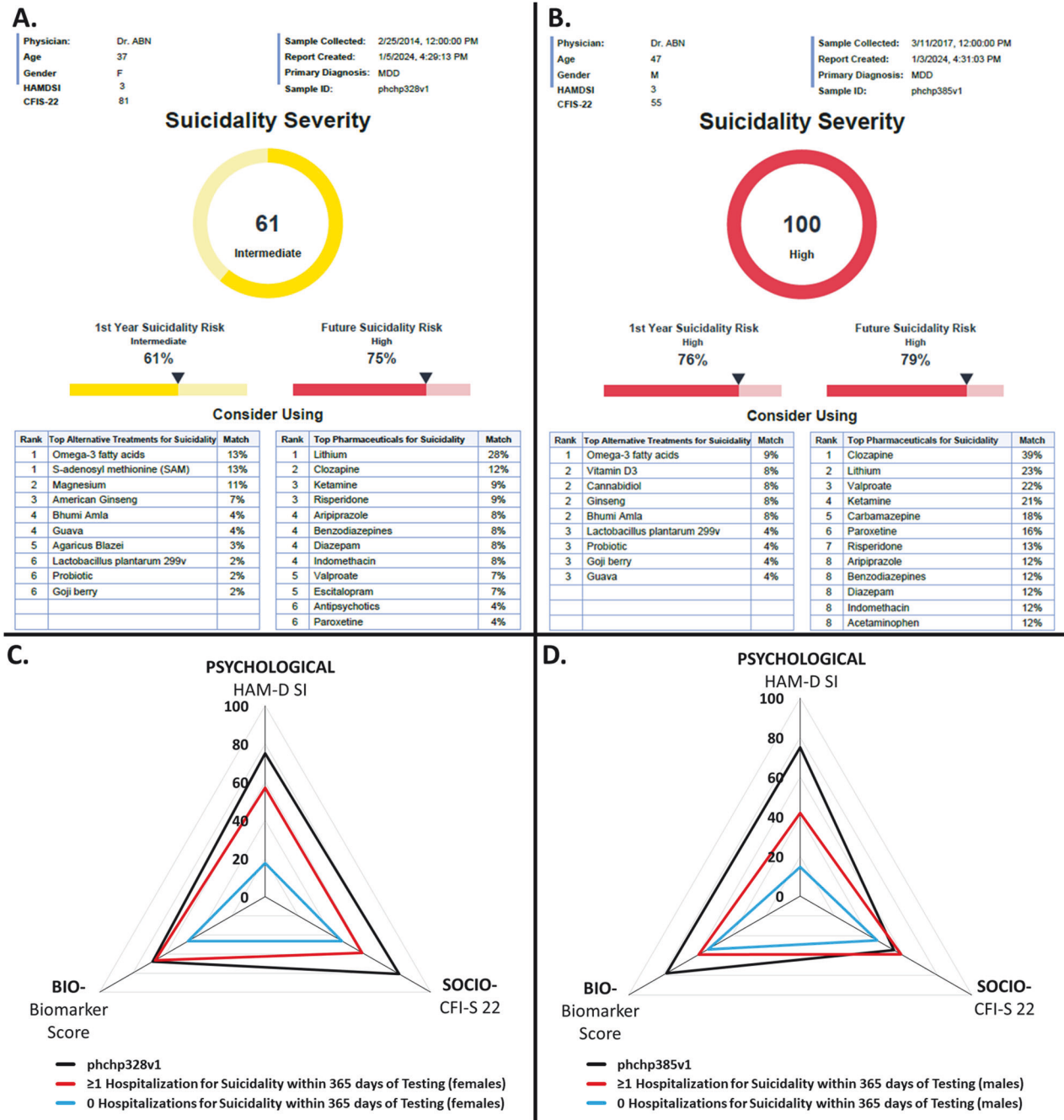
Fig. 2 Prototype reports and population Radar Plot. Subject phchp328 (female, 37 years old) died by suicide by overdose a year after being tested by us. Phchp385 (male, 47 years old) died by suicide by hanging three years after being tested by us. **A** Prototype Report for Phchp328v1. **B** Prototype Report for phchp385v1. Reports based on panels of top predictive biomarkers for that gender. **C, D** Radar plots of Hospitalizations in the First Year following testing. Our individual subject scores (black line), as well as average scores for high risk subjects (red, $n = 768$) and average scores for low risk subjects (blue, $n = 176$).

classified as high (red) above average, and low (blue) if below average for that scale. 16 distinct subtypes were revealed via the hierarchical clustering. The average hospitalization frequency for suicide related hospitalizations were then calculated for each subtype within the independent testing cohort (Fig. S2). It likely reflects a combination of self-perceived need for, and ability to seek help and get hospitalized.

Machine learning (ML) analyses
We compared different ML approaches for predicting occurrence of hospitalizations for suicidality in the year following testing, and time to first hospitalization. We used a comprehensive bio-socio- psychological input

into the models, consisting of panels of best predictive biomarkers by gender (the same ones used in the Reports in Fig. 2), as well as the CFI-S scale items, and the HAMD-SI item. The ML investigation is designed as follows: 1) the gene expression blood biomarkers, CFI-S items, overall CFI-S score, and HAMD-SI from each patient were considered as input features of the ML architectures; 2) the occurrence of hospitalization and the time to 1st hospitalization results were converted to binary indicators.

We used and compared the following ML approaches: 1. Support Vector Machine (SVM) is a supervised learning algorithm that finds the hyperplane which best separates different classes in the feature space; 2. Random Forest (RF) is an ensemble learning method that operates by constructing a multitude of decision trees at training time to output the class that is the

mode of the classes of the individual trees; 3. XGBoost stands for eXtreme Gradient Boosting and represents a scalable and accurate implementation of gradient boosting machines, which are used for supervised learning problems by optimizing differential loss functions; 4. Transformer is a deep learning model that uses self-attention mechanisms to process sequential data, and it is widely used in natural language processing tasks; and 5. Deep Neural Networks (DNN) refer to a deep learning technique consisting of several artificial neurons (i.e., non-linear computational units) organized in multiple layers (i.e., an input layer matching in size the input feature data, several hidden layers of various sizes that are decided based on a neural architecture search to minimize specific loss functions, and an output layer whose size is determined by the desired ML classification or prediction task) for extracting discriminative features from a given dataset.

In comparing the strengths and weaknesses of various ML algorithms, each brings unique advantages and challenges to the table. SVM approaches offer robust performance in high-dimensional spaces and are effective in cases where the number of dimensions exceeds the number of samples. However, they struggle with large datasets and require careful tuning of parameters. RF approaches are known for their simplicity and ability to run efficiently on large datasets, but they can overfit in cases of noisy data. XGBoost excels in handling sparse data and is faster and more efficient than traditional Gradient Boosting, but tuning its hyperparameters can be complex and time-consuming. Transformers show exceptional performance on sequential data, particularly in natural language processing tasks, due to their self-attention mechanism; however, they require significant amounts of data and computational power, potentially limiting their applicability in resource-constrained settings. DNNs are highly flexible and capable of learning complex patterns from large amounts of data, making them suitable for a wide range of applications, including image and speech recognition. Their primary drawbacks are the need for extensive computational resources and the risk of overfitting, which can be mitigated with techniques such as dropout.

The DNN approach turned out to be the most predictive, and was chosen to be used also for the subsequent feature importance analysis described below.

In more detail, we designed a comprehensive DNN framework for predicting the occurrence of hospitalization and the time to 1st hospitalization for suicidality. We provide details about the constructed gender-specific DNN models in the following paragraph.

Deep neural network training and testing. 261 subjects were designed as a training cohort, to train and tune the hyper-parameters of the DNN during the neural architecture search. Due to size and structure of the available data, the male and female models are trained and tested separately. The male model was trained with 217 subjects and tested on 570 subjects. The female model was trained with 44 subjects and tested on 115 subjects.

In females, for Occurrence of Hospitalizations in the First Year, the input layer had 32 neurons, with three hidden layers, with 128, 32 and 32 neurons each. For Time to First Hospitalization, the input layer had 32 neurons, with four hidden layers with 128, 128, 64 and 64 neurons each.

In males, for Occurrence of Hospitalizations in the First Year, the input layer had 32 neurons, with three hidden layers, with 128, 32 and 32 neurons each. For Time to First Hospitalization, the input layer had 32 neurons, with four hidden layers with 128, 64, 64 and 64 neurons each.

All the hidden layers were attached with a batch normalization layer and a dropout layer with 0.2 dropout rate and the ReLU activation function was applied on all dense layers, where both batch normalization and dropout layers are responsible for avoiding overfitting. The output layer has 2 output neurons with the "sigmoid" activation function. The DNN models used an optimizer with 0.001 learning rate and binary cross entropy loss function. Grid and random searches determined suitable hyperparameter values for all ML models in this work (e.g., batch size, kernel size, weight decay). An n-dimensional grid was defined to map the n hyperparameters and to identify their ranges. We examined all possible DNN configurations to identify optimal values for each hyperparameter. Of note, results of predictions in the testing cohorts were used for that. As such, unlike our four-step biomarker identification, prioritization, validation, and testing studies, the ML approach is not based on completely independent cohorts.

Feature importance analysis (saliency value). Next, we computed the saliency value for each input. For input X0 and a DNN model with a score function S(X), we ranked features (genotypes) in X0 based on their importance to S(X). We considered the linear score model $S(X){=}wTX + b$, where the input X, weight w and bias b are in one-dimensional (vectorized)

forms. Since the DNN model and score function are highly nonlinear functions of X, the linear score model cannot be applied directly. We approximate S(X) at X0 using the first-order Taylor series S(X0)w0TX0+b, where w0 = ∂S/∂X|X0 is the partial derivative of S(X) at X0 and b0 is the bias at X0, and w0 represents the saliency value for each input in X0.

RESULTS

In Step 1 Discovery, we identified candidate blood gene expression biomarkers that: 1. change in expression in blood between no and high suicidality states, 2. track the state across visits in a subject, and 3. track the suicidality states in multiple subjects. We used as a quantitative measure for suicidality the Suicidal Ideation item in the Hamilton Rating Scale for Depression (HAMD-SI) (Fig. S1). At a phenotypic level, this item quantifies suicidality state at a particular moment in time, and over the week prior to testing (Fig. S1).

For the discovery step, we used a powerful within –subject and then across-subject design in a longitudinally followed cohort of subjects (for the Affymetrix samples, $n = 68$ subjects, with 221 visits; for the RNAseq samples, $n = 22$ subjects, with 52 visits) who displayed at least a change in the suicidality measure (from 0 to 2 and above, and vice-versa) between at least two consecutive testing visits, to identify differentially expressed genes that track suicidality state.

The data from the two platforms was integrated as described in Methods. Using our 33% of maximum raw score threshold (internal score of 2 pt) [3, 4], there are 9184 unique probesets with corresponding transcripts.

In Step 2 Prioritization, we used a Convergent Functional Genomics (CFG) approach to prioritize the candidate biomarkers identified in the discovery step (33% cutoff, internal score of ≥2 pt.) by using prior published literature evidence (genetic, gene expression and proteomic), from human studies, for involvement in suicidality (Fig. 1 and Table S2). There were 2438 probesets with corresponding transcripts that had a total score (combined discovery score and prioritization CFG score) of 6 and above. These were carried forward to the validation step.

In Step 3 Validation, we validated the prioritized candidate biomarkers for change in suicide completers ($n = 101$, $n = 45$ Affymetrix and $n = 56$ RNAseq). We assessed which biomarkers were stepwise changed in expression from no suicidality in the discovery cohort, to high suicidality in the discovery cohort, to suicide completers (Fig. 1). Of the 2438 probesets after the prioritization step, 739 were nominally significant, and of these, 382 were Bonferroni significant.

Adding the scores from the first three steps into an overall convergent functional evidence (CFE) score (Fig. 1), we ended up with a list of 2340 top candidate biomarkers for suicidality, that had a CFE3 score ≥8, better than 33% of the maximum possible score of 24 after the first three steps, which we decided to use as an empirical cutoff. These top candidate biomarkers were tested in Step 4 for clinical utility/predictive ability in additional independent cohorts (Fig. 1 and Table 1).

Testing for clinical utility

In Step 4 Testing, we examined in independent cohorts from the ones used for discovery or validation whether the top candidate biomarkers after the first three steps can assess high suicidality states, as well as predict of future psychiatric hospitalizations due to suicidality (Fig. 1 and Table S1), using electronic medical records follow-up data of our study subjects (up to 17.2 years from initial visit at the time of the analyses). The gene expression data in the test cohorts was normalized (Z-scored)] by gender, before those groups were combined. This permits them to be combined, and reduces bias from larger groups. We used as predictors biomarker levels information cross-sectionally, as well as expanded long-itudinal information about biomarker levels at multiple visits. We

tested the biomarkers in all subjects in the independent test cohort, as well as in a more personalized fashion by gender (Fig. 1).

Convergent functional evidence (CFE)

For the top candidate biomarkers ($n = 2340$), we computed into a convergent functional evidence (CFE) score all the evidence from discovery (up to 6 points), CFG prioritization (up to 12 points), validation (up to 6 points), and testing (predicting state suicidality, first year hospitalizations with suicidality, all future hospitalizations with suicidality- up to 4 points each if it significantly predicts in all subjects, 2 points if in gender). The total score can be up to 36 points: 24 from our own new data, and 12 from literature data used for CFG. We weigh our new data more than the literature data, as it is functionally related to suicidality in 3 independent cohorts (discovery, validation, testing). The goal is to highlight, based on the totality of our data and of the evidence in the field to date, biomarkers that have all around evidence: track suicidality, have convergent evidence for involvement in suicidality, and predict suicidality state, and future clinical events (Table 1).

The 30 top blood biomarkers with the strongest overall convergent functional evidence (CFE) for tracking and predicting suicidality, after all four steps (Table 1) were, in order of CFE4 score: SLC6A4 (Solute Carrier Family 6 Member 4), TINF2 (TERF1 Interacting Nuclear Factor 2), INSR (Insulin Receptor), CLN5 (CLN5 Intracellular Trafficking Protein), PKP4 (Plakophilin 4), SLC49A4 (Solute Carrier Family 49 Member 4), SKP1 (S-Phase Kinase Associated Protein 1), ECHDC1 (Ethylmalonyl-CoA Decarboxylase 1), BCL2 (BCL2 Apoptosis Regulator), SELENOF (Selenoprotein F), SYNE2 (Spectrin Repeat Containing Nuclear Envelope Protein 2), NDFIP1 (Nedd4 Family Interacting Protein 1), VTI1B (Vesicle Transport Through Interaction With T-SNAREs 1B), E2F1 (E2F Transcription Factor 1), CTIF (Cap Binding Complex Dependent Translation Initiation Factor), MTCH2 (Mitochondrial Carrier 2), PRKAR2B (Protein Kinase CAMP-Dependent Type II Regulatory Subunit Beta), ANGPT1 (Angiopoietin 1), KLF12 (KLF Transcription Factor 12), CDH4 (Cadherin 4), APOE (Apolipoprotein E), MYH10 (Myosin Heavy Chain 10), UBL3 (Ubiquitin Like 3), CALD1 (Caldesmon 1), APC (APC Regulator Of WNT Signaling Pathway), MAP3K7 (Mitogen-Activated Protein Kinase Kinase Kinase 7), MAOA (Monoamine Oxidase A), LINC01432 (Long Intergenic Non-Protein Coding RNA 1432), S100A10 (S100 Calcium Binding Protein A10), and AGO2 (Argonaute RISC Catalytic Component 2).

SLC6A4, the overall top biomarker for suicidality in this study, is the serotonin transporter, which plays an essential role in the mechanism of action of serotonin-based anti-depressants. Abnormalities in serotonin may be central to suicidality [22]. SLC6A4, increased in expression in blood in high suicidality in our work, has previous, convergent evidence for involvement in suicidality. It is increased in expression in hippocampus in suicides [23]. There is also previous human genetic evidence [10, 24]. SLC6A4 in our studies modestly predicts high suicidality state in all patients in the independent testing cohort (AUC 59%, $p = 0.02$) using longitudinal analyses, with results being somewhat better in females (AUC 68%, $p = 0.02$). It also predicts future hospitalizations with suicidality in females, in the first year (AUC 79%, $p = 0.0007$), and in future years (OR 2.71, $p = 0.002$). SLC6A4 activity is blocked by antidepressants, but SLC6A4 may be increased in expression by SSRI treatment itself, in a physiological negative feedback loop. Medication non-compliance and withdrawal may expose the patient to high levels of SLC6A4, that decrease serotonin and increase impulsivity and suicidality. In particular short-half life SSRIs used in non-compliant populations like children and adolescents may increase the risk of suicidality. This may be why long-half life SSRIs like fluoxetine are safer and thus preferentially used in those under age 25.

TINF2 (TERF1 Interacting Nuclear Factor 2), another top biomarker, is a key component of the shelterin complex (telosome) that is involved in the regulation of telomere length and protection. A decrease in TINF2, as we see in suicidality, would lead to telomere shortening.

TINF2 in our studies predicts high suicidality state in females in the independent testing cohort (AUC 72%, $p = 0.005$) using longitudinal analyses. It also predicts future hospitalizations with suicidality in females, in the first year (AUC 67%, $p = 0.03$), and modestly predicts hospitalizations in all in future years (OR 1.2, $p = 0.005$). TINF2 has also been shown to be decreased in expression in blood in previous studies we did in stress [25], low memory [26], and hallucinations [26], suggestive of a stress-driven neuropathological component.

INSR (Insulin Receptor), another top biomarker, is a receptor tyrosine kinase which mediates the pleiotropic actions of insulin. In vivo and in several cell models, the expression of the insulin receptor and/or its mRNA is under positive regulation by glucocorticoid hormones and negative regulation by insulin. Glucocorticoid hormones stimulate receptor gene transcription and receptor protein synthesis. INSR is increased in expression in suicidality in our work, consistent with a high stress state. INSR in our studies modestly predicts high suicidality state in all patients in the independent testing cohort (AUC 60%, $p = 0.01$) as well as future hospitalizations (OR 1.38, $p = 0.002$), using longitudinal analyses, with results being somewhat better in females (OR 1.86, $p = 0.02$). It also has previous genetic evidence [27], and human postmortem brain evidence of being increased in the hippocampus in suicides [28]. INSR has also been shown to be decreased in expression in blood in previous studies we did in stress [25], anxiety [17], depression [16], low memory [26], and hallucinations [26], suggestive of a stress-driven neuropathological component. It is decreased in expression by lithium [29], valproate [30], and antidepressants [31].

CLN5 (CLN5 Intracellular Trafficking Protein), another top biomarker, is involved in the degradation of post-translationally modified proteins in lysosomes.CLN5 is decreased in suicidality in our work, consistent with a lysosomal accumulation. It modestly predicts high suicidality state in all patients in the independent testing cohort (AUC 66%, $p = 0.0002$), with results being somewhat better in females (AUC 74%, $p = 0.004$), using longitudinal analyses, as well as future hospitalizations in all (OR 1.85, $p = 0.004$). CLN5 is also decreased in the blood in our biomarker studies on stress [25], anxiety [17], and depression [16], as well as in excitatory neurons in dementia [32], suggestive of a stress-driven neuropathological component. It is increased in expression by lithium [29]. BCL2 (BCL2 Apoptosis Regulator) is involved in apoptosis. BCL2 was decreased in expression in suicidality in our work, consistent with an anti-survival mechanisms/ increased cell death. Suicidality may be a whole-organism apoptosis [2, 33], and we have described in our previous work suicidality as the opposite of longevity at a molecular level, the negative direction of a biological "Life Switch". BCL2 modestly predicts high suicidality state in all patients in the independent testing cohort (AUC 60%, $p = 0.01$), with results being somewhat better in females (AUC 77%, $p = 0.0008$), using longitudinal analyses, as well as future hospitalizations in the first year in females (AUC 72%, $p = 0.007$). It also has previous independent convergent evidence from human postmortem brain studies of being decreased in the pre-frontal cortex in suicide completers [34], particularly females [35], and in white blood cells [36]. BCL2 is also decreased in expression in the blood in our biomarker studies on pain [37], depression [16], and in studies by others on aging [38, 39], suggestive of an adversity-driven pathological component. BCL2 is increased in expression by lithium [40, 41], clozapine [41], as well as SNRIs (venlafaxine [42], duloxetine [43]), buproprion [44], and doxepin [45]. It is also increased in expression by the nutraceuticals omega-3 fatty acids [46], CoQ10 [47], curcumin [48], fisetin [49], and CBD [50].

APOE (Apolipoprotein E) is involved in lipid transport, as well as the repair and regeneration of neurons. APOE was decreased in

Table 2. <u>Biological Analyses</u> A. Pathways. B. Diseases. C. Upstream Regulators. D. Genomic Co-Morbidity for Top Biomarkers. E. Therapeutics for Top Biomarkers.

A.

DAVID Functional Annotation Biological Processes

Top CFE BioM (n = 30 genes)	#	Term	Count	%	P-Value
	1	**positive regulation of apoptotic process**	4	13.3	1.10E−02
	2	**negative regulation of neuron apoptotic process**	3	10	2.10E−02
	3	**anoikis**	2	6.7	1.70E−02
	4	neuron maturation	2	6.7	2.20E−02
	5	positive regulation of receptor internalization	2	6.7	3.80E−02

Ingenuity Pathways

Top Canonical Pathways	P-Value	Overlap
Neurotransmitter clearance	6.22E−05	20.0% 2/10
Ovarian Cancer Signaling	3.68E−05	2.5% 4/158
Autophagy	1.26E−04	1.8% 4/217
Insulin Secretion Signaling Pathway	3.11E−04	1.5% 4/275
Synaptogenesis Signaling Pathway	5.20E−04	1.3% 4/315

B.

David Genetic Association Disease

Top CFE BioM (n = 30 genes)	#	Term	Count	%	P-Value
	1	Major depressive disorder	4	13.3	1.40E−04
	2	Epilepsy, Temporal Lobe	3	10	1.20E−04
	3	Sleep disorders	3	10	2.30E−04
	4	Brain Injuries	3	10	3.20E−04
	5	Sudden Infant Death	3	10	1.00E−03

Ingenuity Pathways Disease

Diseases and Disorders	P-Value	# Molecules
Organismal Injury and Abnormalities	3.60E−03 to 1.03E−07	28
Neurological Disease	3.60E−03 to 7.04E−07	26
Skeletal and Muscular Disorders	3.60E−03 to 7.04E−07	18
Cardiovascular Disease	3.60E−03 to 1.03E−07	16
Psychological Disorders	3.60E−03 to 8.86E−07	12

C. Upstream Regulators Top CFE BioM

Regulator	P-Value
Prednisolone	3.05E−07
3-methyladenine	7.65E−07
thyroid hormone	1.86E−06
KLF3	2.57E−06
Simvastatin	5.36E−06

D. Genomic Co-Morbidities Top CFE BioM

Comorbidity	Percentile
Alcohol	90%
Depression	83.3%
Stress	80%
Dementia	76.7%
Schizophrenia	66.7%
Aging	50%
Bipolar	50%
Pain	46.7%
Addictions	46.7%
Anxiety	40%
ASD	26.7%
PTSD	23.3%
ADHD	6.7%
Neurological	6.7%

Table 2. continued

D. Genomic Co-Morbidities Top CFE BioM

Comorbidity	Percentile
Delusions	6.7%

E. Therapeutics

Top Overall Biomarkers (n = 30 genes)

Treatment	Percentile
Lithium	26.7%
Clozapine	23.3%
Ketamine	20%
Valproate	16.7%
Citalopram	13.3%
Omega-3 fatty acids	13.3%
Vortioxetine	13.3%
Carbamazepine	10%
Escitalopram	10%
Fluoxetine	10%
Ginseng	10%
Imipramine	10%
Indomethacin	10%
Magnesium	10%
Risperidone	10%

Top Predictive Biomarkers in Males (n = 23 genes)

Treatment	Percentile
Clozapine	34.8%
Lithium	30.4%
Valproate	30.4%
Carbamazepine	21.7%
Ketamine	21.7%
Acetaminophen	17.4%
Escitalopram	13%
Indomethacin	13%
Paroxetine	13%
Bhumi Amla	8.7%
Probiotic	8.7%
Risperidone	8.7%
Sertraline	8.7%
Aripiprazole	4.3%
Brexanolone	4.3%

Top Predictive Biomarkers in Females (n = 26 genes)

Treatment	Percentile
Lithium	26.9%
Ketamine	23.1%
Clozapine	19.2%
Magnesium	15.4%
Risperidone	11.5%
Valproate	11.5%
Escitalopram	7.7%
Fluoxetine	7.7%
Ginseng	7.7%
Indomethacin	7.7%
Nortriptyline	7.7%
Omega-3 fatty acids	7.7%
Probiotics	7.7%
S-adenosyl methionine (SAM)	7.7%
Aripiprazole	3.8%

expression in suicidality in our work, consistent with decreased neuronal repair and growth. Suicidality may be a form of dementia. APOE predicts high suicidality state in all patients in the independent testing cohort (AUC 74%, $p = 1.30E-22$), with results being somewhat better in females (AUC 88%, $p = 7.35E-06$). It also predicts first year hospitalizations in all (AUC 67%, $p = 9.04E-09$,) with results better in females (AUC 88%, $p = 1.07E-05$), and all future hospitalizations in all (OR 1.38, $p = 1.26E-05$,) with results again better in females (OR 1.9, $p = 1.60E-03$). It also has previous independent convergent evidence from human postmortem brain studies of being decreased in the pre-frontal cortex in suicide completers [51], particularly females [35]. APOE is also decreased in expression in aging [52] and increased in longevity [53], potentially being involved in the Life Switch described by us [33]. It is decreased in expression in the blood in our biomarker studies on depression [16],as well as psychosis [54], memory disorders [26], and in studies by others on ASD [55], suggestive of a cognitive component. APOE is increased in expression by escitalopram [56] and by nortryptiline [56], as well as the nutraceuticals ginseng [57] and magnesium [58].

MAOA (Monoamine Oxidase A), a biomarker increased in expression in our studies, is an enzyme involved in neurotransmitter degradation, and is the target of a class of broad-spectrum antidepressants that inhibit it. MAOA modestly predicts high suicidality state in females (AUC 65%, $p = 0.04$) in longitudinal analyses, as well as future hospitalizations in the first year (AUC 74%, $p = 0.003$), and all future years (OR 2.42, $p = 0.001$). It also has previous independent convergent evidence from genetic studies [59, 60], and of being increased in expression in human neuronal progenitor cells studies of suicidality [61]. MAOA is also increased in expression in the blood in our biomarker studies on stress [25], pain [37], anxiety [62], depression [16], and in studies by others on panic disorders [63] and on depression [64], consistent with an adversity -driven pathological component. MAOA is inhibited/ decreased in expression by antidepressants [65–67], as well as by ketamine [68], the nutraceutical olive extract [67], and by psychotherapy [69].

Biological understanding
Biological pathways. We carried out biological pathway analyses using the list of top biomarkers for suicidality ($n = 30$ genes). The top pathways were related to apoptosis and neurotransmitter clearance (Table 2A). Major depressive disorder and sleep disorders were top diseases identified by the pathway analyses using DAVID, pointing out to a molecular underpinning for these well-known clinical co-morbidities (Table 2B).

Networks and interactions. We carried out a STRING analysis (Fig. S3) of the top candidate biomarkers that revealed groups of interacting proteins. BCL2 is at the nexus of three networks: one containing SLC6A4, MAOA, and APOE; one containing INSR and ANGPT1; and one containing AGO2 and E2F1. These networks may have biological significance and could be targeted therapeutically.

Therapeutics
Overall, lithium (26.7%) had the best evidence for broad efficacy in suicidality (Table 2E), followed by clozapine (23.3%) and ketamine (20%). Interestingly, these are the only medications in psychiatry approved for suicidality, and their emergence out of our empirical, hypothesis-driven work, is a strong validation of our approach. Omega-3 fatty acids (13.3%) was the top nutraceutical and may be a widely deployable preventive treatment, with minimal side-effects, including in women who are or may become pregnant.

A number of individual top biomarkers are known to be modulated by medications in current clinical use for treating schizophrenia and suicidality such as by clozapine, mood disorders and suicidality such as lithium, ketamine, as well as the nutraceutical omega-3 fatty acids (Tables 1and S4). This is of potential utility in pharmacogenomics approaches matching suicidality patients to the right medications, and monitoring response to treatment.

Best predictive biomarkers
In Step 4, we identified best predictive biomarkers for suicidality state (HAMDSI ≥ 2) and trait (first year, and all future hospitalizations, for suicidality), using cross-sectional and longitudinal methodology. In an additional Step 5, all the nominally significant biomarkers were tested for ability to predict using the whole population used in the study ($n = 1127$), to avoid an overfit to the testing cohort. The best predictive biomarkers in all, and for each gender, male and female, can be combined in panels to generate reports for doctors, as shown in Fig. 2.

Machine learning
Machine learning (ML) models (e.g., XGB, RF, and SVM) were trained on the discovery cohort, and tested in the independent test cohort for predicting hospitalizations in the first year following testing. Deep Neural Networks (DNN) performed best. The model used a combination of the top biomarkers, CFI-S, and HAMD-SI. It's predictive ability was compared to and shown to be less than to our simple additive model of the three scores (biomarker panel score, CFI-S score, and HAMD-SI score) (Table 3). In general, the model predicted better in females than in males, consistent with all of our other results. The top predictive features for occurrence of hospitalizations for suicidality in the first year following testing were HAMD-SI (suicidality intensity) and JOSD1 in females, and JOSD1 and THY1 in males. For imminence of hospitalizations, the top features were Feeling Useless and HAMD-SI in females, and Medical Problems and Age in males. These represent readily addressable targets for preventive approaches.

JOSD1 (Josephin Domain Containing 1) is involved in deubiquitination. It is decreased in expression in blood in high suicidality in our studies, and also has convergent evidence of being decreased in expression postmortem brain studies from suicide completers [35]. The decrease in JOSD1 may be associate with increased ubiquitination and apoptosis, which is cellular suicide [70].

DISCUSSION
We describe novel and comprehensive efforts to advance precision medicine approaches for suicidality. The top blood biomarkers were discovered, validated and tested in independent cohorts to evaluate predictive ability and clinical utility. These biomarkers also open a window into understanding the biology of suicidality, as well as indicate new and more precise therapeutic approaches.

Current clinical practice and the need for biomarkers
Assessing a persons' internal subjective perceptions and thoughts, along with more objective external ratings of actions and behaviors, are used in clinical practice to assess suicidality. Such an approach is insufficient, and lagging those used in other medical specialties. Moreover, individuals do not always report or want to share when they are suicidal, leading to missed opportunities to intervene and help. Blood biomarkers related to suicidality, if used as part of routine mental health clinical visits and even primary care annual exams, would provide a critical objective measurement to inform clinical assessments and treatment decisions.

Advantages of biomarkers
Blood biomarkers offer real-world clinical practice advantages. As the brain cannot be readily biopsied in live individuals, and CSF is less easily accessible than blood, we have endeavored over the years to identify blood biomarkers for neuropsychiatric disorders.

Table 3. Genomic and clinical predictions comparison and integration.

Gender	Predictor	Events/Total Events	ROC AUC	AUC *p*-value
Females	Bio-Socio-Psychological (Biomarker Panel + CFI-S + HAMD-SI)	24/113	0.87	1.97E−08
Females	CFI-S + HAMD-SI	24/113	0.82	7.28E−07
Females	Best Predictive Biomarker (FGF18)	25/115	0.80	2.28E−06
Females	HAMD-SI	25/115	0.79	1.86E−07
Females	Biomarker Panel	25/115	0.79	4.66E−06
Females	CFI-S	24/113	0.77	2.62E−05
Females	Machine Learning (DNN)	113	0.76	1.24E−04
Males	Bio-Socio-Psychological (Biomarker Panel + CFI-S + HAMD SI)	84/569	0.78	9.26E−17
Males	CFI-S 22 + HAMD-SI	84/569	0.77	3.17E−15
Males	HAMD-SI	85/570	0.75	1.53E−21
Males	Machine Learning (DNN)	569	0.75	2.53E−04
Males	CFI-S	84/569	0.71	1.66E−10
Males	Best Predictive Biomarker (SFN)	85/570	0.67	1.94E−07
Males	Biomarker Panel	85/570	0.59	5.48E−03

Predictions of occurrence of hospitalizations for suicidality in the year following testing were performed the independent Testing Cohort, in males ($n = 570$) and females ($n = 115$) separately. Deep Neural Network (DNN). Predictors performed better in females than in males. The simple additive Bio-Socio-Psychological integration was synergistic and performed best in both genders, including better than the best ML approach (DNN).

A whole–blood approach facilitates field deployment of sample collection. The assessment of gene expression changes focuses our approach on immune cells. The ability to identify peripheral gene expression changes that reflect brain activities is likely due to the fact that the brain and immune system have developmental commonalities, marked by shared reactivity and ensuing gene expression patterns. There is also a bi-directional interaction between the brain and immune system. Not all changes in expression in peripheral cells are reflective of or germane to brain activity. By carefully tracking a phenotype with our within-subject design in the discovery step, and then using convergent functional genomics prioritization, we are able to extract the peripheral changes that do track and are relevant to the brain activity studied, in this case suicidality. Subsequent validation and testing in independent cohorts narrow the list to the best markers. In the end, we do not expect to recapitulate in the blood all that happens in the brain. We just want to have good accessible peripheral biomarkers- "liquid biopsies", as they are called in cancer.

Comprehensiveness

In this current work, we carried out extensive blood gene expression studies in male and female subjects with major psychiatric disorders, an enriched population in terms of co-morbidity with suicidality. The potential molecular-level co-morbidity between other psychiatric disorders and suicidality is underlined by the fact that medications for mood disorders (lithium) and psychosis (clozapine) are also used to treat suicidality. Our primary goal was to discover and validate biomarkers for suicidality, that are transdiagnostic. Secondarily, we aimed to understand their universality vs. their specificity by gender.

Our studies were arranged in a stepwise fashion. First, we endeavored to discover blood gene expression biomarkers for suicidality using a longitudinal design, looking at differential expression of genes in the blood of male and female subjects with major psychiatric disorders (bipolar disorder, major depressive disorder, schizophrenia/schizoaffective, and post-traumatic stress disorder (PTSD)), high risk populations prone to suicidality, which constitute and enriched pool in which to look for biomarkers. We compared no suicidality states to high suicidality states using a powerful within-subject design [2–4, 71], to generate a list of differentially expressed genes. Second, we used a comprehensive Convergent Functional Genomics (CFG) approach with the whole body of knowledge in the field to prioritize from the list of differentially expressed genes/biomarkers of relevance to suicidality. CFG integrates multiple independent lines of evidence-genetic, gene expression, and protein data, from brain and periphery, from human studies, as a Bayesian strategy for identifying and prioritizing findings, reducing the false-positives and false-negatives inherent in each individual approach. Third, we examined if the expression levels of the top biomarkers identified by us as tracking suicidality state are changed even more strongly in blood samples from an independent cohort of subjects who died by suicide, to validate these biomarkers. Fourth, the biomarkers thus discovered, prioritized, and validated were tested in corresponding independent cohorts of psychiatric subjects. Fifth, we used the biomarkers to match to existing psychiatric medications, as well as to identify and potentially repurposed drugs for suicidality treatment and prevention, using bioinformatics analyses. Sixth, we used bio-socio-psychological approaches, including with and without machine learning, to identify best predictors. The series of studies was a systematic and comprehensive approach to move the field forward towards precision medicine.

Power

We used a systematic discovery, prioritization, validation, and testing approach, as we have done over the years for suicidality and other disorders [6, 16, 17, 21, 37, 72]. For discovery, we used a hard to accomplish but powerful within-subject design, with an N of 90 subjects with 273 visits. A within-subject design factors out genetic variability, as well as some medications, lifestyle, and demographic effects on gene expression, permitting identification of relevant signal with Ns as small as 1 [71]. Another benefit of a within-subject design may be accuracy/consistency of psychiatric symptoms ("phene expression"), as it is the same person reporting different states. This is similar in rationale to the signal detection benefits it provides in gene expression.

Based on our work of over two decades in genetics and gene expression, along with the results of others in the field, we estimate that using a quantitative phenotype is up to 1 order of magnitude more powerful than using a categorical diagnosis. The within-subject longitudinal design, by factoring out all genetic and

some environmental variability, is up to 3 orders of magnitude more powerful than an inter-subject case-control cross-sectional design. Moreover, gene expression, by integrating the effects of many SNPs and environment, is up to 3 orders of magnitude more powerful than a genetic study. Combined, our approach may be up to 6 orders of magnitude more powerful than a GWAS study, even prior to the CFG literature-based prioritization step, which encompasses all the independent work in the field prior to our studies, which may add up to 1 order of magnitude as well. In addition, the Validation and the Testing steps add additional 1 order of magnitude power each. As such, our approach might be up to 10 orders of magnitude more powered to detect signal than most current genetic study designs as used in GWAS.

Reproducibility
We reproduced and expanded our earlier biomarker findings [6]. 86% of our candidate biomarkers from the discovery step in the current work reproduce those in our 2017 study. Moreover, 29 out of our 30 top biomarkers after Step 4 in the current study passed the discovery step in the 2017 study. 26 out of our 30 current top biomarkers were Bonferroni significant after the validation step, compared to 4 of them in the earlier smaller 2017 study.

Additionally, there is reproducibility of our candidate biomarkers from discovery with findings generated by other independent studies as part of the Step 2 Prioritization using Convergent Functional Genomics (see Table S2). This independent reproducibility of findings between our studies and these other studies, which are done in independent cohorts from ours, with independent methodologies, is reassuring, and provides strong convergent evidence for the validity and relevance of our approach and of their approaches. Our work also provides functional evidence for some of their top genetic hits.

Pathophysiology
Top biological pathways have to do with apoptosis (cellular suicide) and neurotransmitter clearance (Table 2A). Suicidality may be a whole-body apoptosis in response to an adverse environment.

The majority of top blood biomarkers we have identified have prior evidence in human brain data from suicides, which indicates their relevance to the pathophysiology of suicidality (Table S2). The co-directionality of blood changes in our work and brain changes reported in the literature needs to be interpreted with caution, as it may depend on brain region.

The top candidate biomarkers also had prior evidence of involvement in other psychiatric and related disorders (Table 1 and S3), providing a molecular basis for co-morbidity, and the possible predisposing effects of some these disorders on suicidality. In particular, over 80% of top biomarkers identified by us overlap with genes implicated in alcoholism, depression, and stress (Table 1, Table 2D), consistent with the known clinical co-morbidity. These are common, treatable and preventable disorders.

6 of the top 30 genes (20%) are involved in the circadian clock, an enrichment over the 7% of the genome that is involved in circadian mechanisms [16]. Circadian clock genes in general are core to mood and to levels of activity of the organism. Thyroid hormones were among the top upstream regulators (Table 2C). Thyroid hormones ae known to have a profound impact on mood and levels of activity.

Phenomenology
In addition to using the standard HAMD-SI item, we had previously developed and used in this study a scale to assess psycho-social determinants of suicidality risk. Convergent Functional Information for Suicidality (CFI-S) scale is a 22-items checklist of risk factors and social impairment, that notably does not ask about suicidal ideation. The CFI-S has been shown in

previous studies [3, 4, 6, 12, 13], and confirmed in the current one, to be predictive of suicidality state and trait.

We have also looked at subtypes of suicidality in the subjects from the discovery cohort while they were in a high suicidality state. We identified 16 subtypes, based on two-way unsupervised hierarchical clustering on measures of stress, anxiety, mood and psychosis (Fig. S2). The three subtypes with the most subsequent hospitalizations in the year following testing had in common high anxiety.

Biomarkers vs. Scales
In general, the best predictive biomarkers were better than the rating scales (CFI-S, HAMD-SI) at predicting trait suicidality in all patients, in males, and in females. In females, the biomarkers were better than the scales for state suicidality also. This may reflect the fact that these are difficult phenotypes to assess by clinicians, and reinforces the need for using objective blood biomarkers to assess suicidality (Table S6). The biomarkers and scales are also synergistic (Table 3).

Diagnostics
For the biomarkers identified by us, combining all the available evidence from this current work into a convergent functional evidence (CFE) score, brings to the fore biomarkers that have clinical utility for objective assessment and risk prediction for suicidality (Table 1). These biomarkers should be tested individually as well as tested as polygenic panels of biomarkers in future clinical studies and practical clinical applications in the field. They may permit to distinguish, upon an initial clinical presentation of suicidality, whether the person is in fact severely suicidal and at chronic risk (Fig. 2). The integration of phenomic data, such as CFI-S done yearly, and repeated measures of HAMD-SI (perhaps via a phone app in a daily fashion), can further substantiate and elucidate suicidality risk, distinguishing between an intermittent type such as transient suicidal ideation, and continuous type such as chronic suicidality. We demonstrated by using a bio-socio-psychological approach that there was synergy between the components. Machine learning did not perform better, but it highlighted the importance of certain individual features (markers). Predictions of occurrence and imminence of hospitalizations for suicidality were stronger in women than in men (Fig. 3 and Table 3).

In general, our predictive results with biomarkers were stronger in females than in males, by an order of 10–20% points on AUCs. While some of it may be biological, in terms of immune system reactivity and brain-blood interplay being perhaps higher in women, it is also possible that men are not as accurate as women in terms of reporting suicidality symptoms (affecting our results on state predictions), and do not seek help as much (affecting our results on future hospitalizations predictions). If so, this under-reporting makes the use of objective biomarker tests in men even more necessary. Of note, death by suicide is four times higher in men than in women.

In regard to how our biomarker discoveries might be applied in clinical laboratory settings, we suggest that panels of top biomarkers for suicidality risk be used (Fig. 2). In practice, every new patient tested would be normalized against the database of similar patients already tested, and compared to them for ranking and risk prediction purposes, regardless if a platform like microarrays, RNA sequencing, or a more targeted one like PCR is used in the end clinically. As databases get larger, normative population levels can and should be established, similar to any other laboratory measures. Moreover, longitudinal monitoring of changes in biomarkers within an individual, measuring most recent slope of change, maximum levels attained, and maximum slope of change attained in the past, may be even more informative than simple cross-sectional comparisons of levels within an individual with normative populational levels, as we
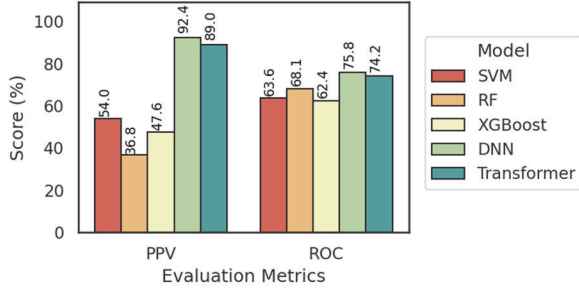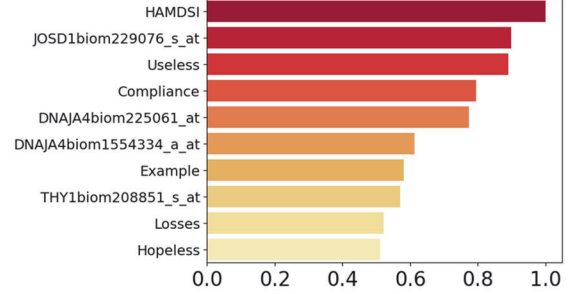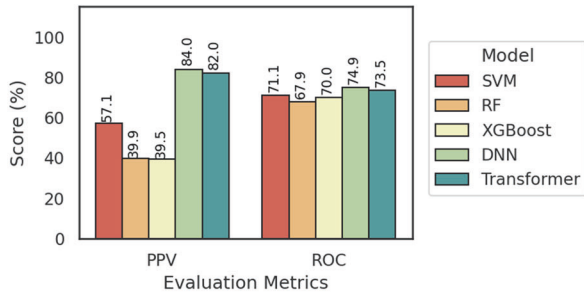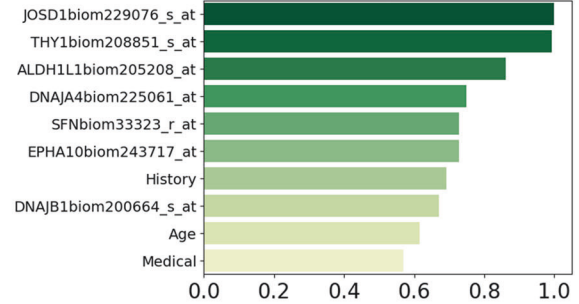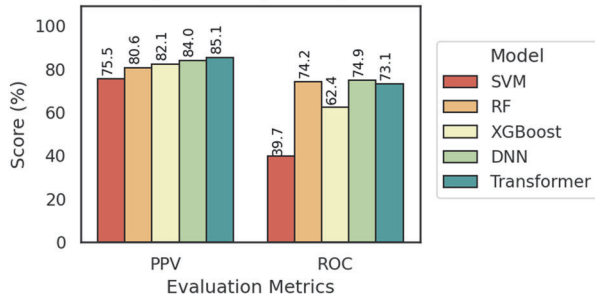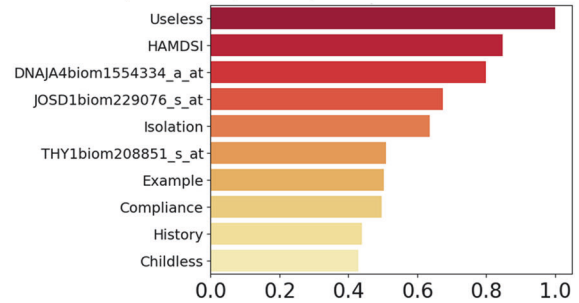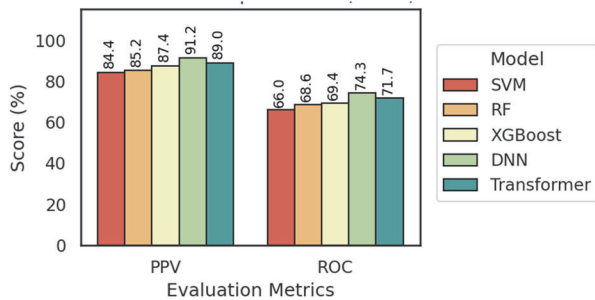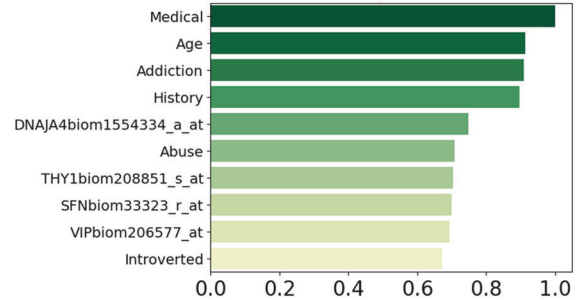
Fig. 3 Machine learning analysis. A, C, E, G Positive predictive value and ROC AUC of occurrence of hospitalizations as well as time to first hospitalization for various machine learning models utilizing all three aspects of the biopsychosocial model (biomarkers, CFIS, and HAMD-SI). **B, D, F, H** Salience analysis of which features of the model are the most important.

have shown in our studies. For future point of care approaches, research and development should focus on top individual biomarkers, including at a protein level. One might look at a combination of the best universal biomarkers (that are predictive in all), for reliability, and of the best personalized biomarkers (that are predictive by gender, and even diagnosis), for higher accuracy.

Treatment
Biomarkers may also be useful for matching patients to medications and measuring response to treatment

(pharmacogenomics) (Fig. 2, Tables 2E and S4), as well as new drug discovery clinical trials, and drug repositioning (Table S7). From the pharmacogenomics analyses, lithium was a top hit, second was clozapine. Other interesting matches were ketamine, valproate, magnesium, omega-3 fatty acids, citalopram, escitalopram. All these drugs and nutraceuticals are relatively safe if used appropriately, and have been used in clinical practice for other indications for decades, which facilitates the direct translation to clinical practice of our findings. The fact that the top medication matches were lithium, clozapine, and ketamine is striking, as the

first two are FDA-approved for suicidality, and ketamine has been a recent addition to the list of medications studied for this indication.

Drug repurposing analyses identified inhibitors of the renin-angiotensin system (lisinopril, losartan, ramipril), and of the cyclooxygenase system (celecoxib pranoprofen, tenoxicam), as potential choices, as well as the SSRI sertraline. Our earlier 2017 study identified metformin as a potential anti-suicidal compound, and this study identified phenformin, a predecessor of metformin with more side-effects.

CONCLUSIONS

Overall, this work is a major step forward towards better understanding, diagnosing, and treating suicidality. Taken together, our data supports the possibility that biologically, suicidality is an extreme stress-driven form of active aging/death. Stress needs to be actively addressed and mitigated in high-risk individuals and circumstances, in both men [73], and women [74]. We hope that our trait biomarkers for future risk may be useful in preventive approaches, before full-blown suicidality manifests itself (or re-occurs). The two cases of subjects who completed out testing and died later by suicide illustrates the power of our approach to identify risk (Fig. 2). Prevention could be accomplished with biological interventions (i.e., early targeted use of medications or nutraceuticals), social measures to help with integration in society using the risks identified by CFI-S, and psychological support. Given the fact that suicidality is on the increase in the US and worldwide, that suicidality can severely affect quality of life and lead to shortened lifespans, and that not all patients respond to current treatments, the need for and importance of efforts such as ours cannot be overstated.

DATA AVAILABILITY

The data that support the findings of this study are available from the senior author (ABN), upon reasonable request.

REFERENCES

1. Sussell A, Peterson C, Li J, Minino A, Scott KA, Stone DM. Suicide Rates by Industry and Occupation - National Vital Statistics System, United States, 2021. MMWR Morb Mortal Wkly Rep. 2023;72:1346–50.
2. Le-Niculescu H, Levey DF, Ayalew M, Palmer L, Gavrin LM, Jain N, et al. Discovery and validation of blood biomarkers for suicidality. Mol Psychiatry. 2013;18:1249–64.
3. Niculescu AB, Levey DF, Phalen PL, Le-Niculescu H, Dainton HD, Jain N, et al. Understanding and predicting suicidality using a combined genomic and clinical risk assessment approach. Mol Psychiatry. 2015;20:1266–85.
4. Levey DF, Niculescu EM, Le-Niculescu H, Dainton HL, Phalen PL, Ladd TB, et al. Towards understanding and predicting suicidality in women: biomarkers and clinical risk assessment. Mol Psychiatry. 2016;21:768–85.
5. Niculescu AB, Le-Niculescu H. Dissecting Suicidality Using a Combined Genomic and Clinical Approach. Neuropsychopharmacology. 2017;42:360.
6. Niculescu AB, Le-Niculescu H, Levey DF, Phalen PL, Dainton HL, Roseberry K, et al. Precision medicine for suicidality: from universality to subtypes and personalization. Mol Psychiatry. 2017;22:1250–73.
7. Niculescu AB, Le-Niculescu H. Convergence of recent GWAS data for suicidality with previous blood biomarkers: independent reproducibility using independent methodologies in independent cohorts. Mol Psychiatry. 2020;25:19–21.
8. Mamdani F, Weber MD, Bunney B, Burke K, Cartagena P, Walsh D, et al. Identification of potential blood biomarkers associated with suicide in major depressive disorder. Transl Psychiatry. 2022;12:159.
9. Levey DF, Polimanti R, Cheng Z, Zhou H, Nunez YZ, Jain S, et al. Genetic associations with suicide attempt severity and genetic overlap with major depression. Transl Psychiatry. 2019;9:22.
10. Mullins N, Kang J, Campos AI, Coleman JRI, Edwards AC, Galfalvy H, et al. Dissecting the Shared Genetic Architecture of Suicide Attempt, Psychiatric Disorders, and Known Risk Factors. Biol Psychiatry. 2022;91:313–27.
11. Docherty AR, Mullins N, Ashley-Koch AE, Qin X, Coleman JRI, Shabalin A, et al. GWAS Meta-Analysis of Suicide Attempt: Identification of 12 Genome-Wide Significant Loci and Implication of Genetic Risks for Specific Health Factors. Am J Psychiatry. 2023;180:723–38.
12. Brucker K, Duggan C, Niezer J, Roseberry K, Le-Niculescu H, Niculescu AB, et al. Assessing Risk of Future Suicidality in Emergency Department Patients. Acad Emerg Med. 2019;26:376–83.
13. Cheng M, Roseberry K, Choi Y, Quast L, Gaines M, Sandusky G, et al. Polyphenic risk score shows robust predictive ability for long-term future suicidality. Discov Ment Health. 2022;2:13.
14. Quinn TP, Hess JL, Marshe VS, Barnett MM, Hauschild AC, Maciukiewicz M, et al. A primer on the use of machine learning to distil knowledge from data in biological psychiatry. Mol Psychiatry. 2024;29:387–401.
15. Ramos A, Ishizuka K, Hayashida A, Namkung H, Hayes LN, Srivastava R, et al. Nuclear GAPDH in cortical microglia mediates cellular stress-induced cognitive inflexibility. Mol Psychiatry. 2024. https://doi.org/10.1038/s41380-024-02553-1.
16. Le-Niculescu H, Roseberry K, Gill SS, Levey DF, Phalen PL, Mullen J, et al. Precision medicine for mood disorders: objective assessment, risk prediction, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2021;26:2776–804.
17. Roseberry K, Le-Niculescu H, Levey DF, Bhagar R, Soe K, Rogers J, et al. Towards precision medicine for anxiety disorders: objective assessment, risk prediction, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2023;28:2894–912.
18. Hill MD, Gill SS, Le-Niculescu H, MacKie O, Bhagar R, Roseberry K, et al. Precision medicine for psychotic disorders: objective assessment, risk prediction, and pharmacogenomics. Mol Psychiatry. 2024.
19. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017;171:1437–1452.e1417.
20. Enache OM, Lahr DL, Natoli TE, Litichevskiy L, Wadden D, Flynn C, et al. The GCTx format and cmapPy, R, M, J packages: resources for optimized storage and integrated traversal of annotated dense matrices. Bioinformatics. 2019;35:1427–9.
21. Le-Niculescu H, Roseberry K, Levey DF, Rogers J, Kosary K, Prabha S, et al. Towards precision medicine for stress disorders: diagnostic biomarkers and targeted drugs. Mol Psychiatry. 2019.
22. Underwood MD, Kassir SA, Bakalian MJ, Galfalvy H, Dwork AJ, Mann JJ, et al. Serotonin receptors and suicide, major depression, alcohol use disorder and reported early life adversity. Transl Psychiatry. 2018;8:279.
23. Gross-Isseroff R, Israeli M, Biegon A. Autoradiographic analysis of tritiated imipramine binding in the human brain post mortem: effects of suicide. Arch Gen Psychiatry. 1989;46:237–41.
24. Schneider E, El Hajj N, Muller F, Navarro B, Haaf T. Epigenetic Dysregulation in the Prefrontal Cortex of Suicide Completers. Cytogenet Genome Res. 2015;146:19–27.
25. Le-Niculescu H, Roseberry K, Levey DF, Rogers J, Kosary K, Prabha S, et al. Towards precision medicine for stress disorders: diagnostic biomarkers and targeted drugs. Mol Psychiatry. 2020;25:918–38.
26. Niculescu AB, Le-Niculescu H, Roseberry K, Wang S, Hart J, Kaur A, et al. Blood biomarkers for memory: toward early detection of risk for Alzheimer disease, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2020;25:1651–72.
27. Sokolowski M, Wasserman J, Wasserman D. Polygenic associations of neurodevelopmental genes in suicide attempt. Mol Psychiatry. 2016;21:1381–90.
28. Glavan D, Gheorman V, Gresita A, Hermann DM, Udristoiu I, Popa-Wagner A. Identification of transcriptome alterations in the prefrontal cortex, hippocampus, amygdala and hippocampus of suicide victims. Sci Rep. 2021;11:18853.
29. Palmos AB, Duarte RRR, Smeeth DM, Hedges EC, Nixon DF, Thuret S, et al. Lithium treatment and human hippocampal neurogenesis. Transl Psychiatry. 2021;11:555.
30. Herteleer L, Zwarts L, Hens K, Forero D, Del-Favero J, Callaerts P. Mood stabilizing drugs regulate transcription of immune, neuronal and metabolic pathway genes in Drosophila. Psychopharmacology. 2016;233:1751–62.
31. Bigio B, Mathe AA, Sousa VC, Zelli D, Svenningsson P, McEwen BS, et al. Epigenetics and energetics in ventral hippocampus mediate rapid antidepressant action: Implications for treatment resistance. Proc Natl Acad Sci USA. 2016;113:7906–11.
32. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. Single-cell transcriptomic analysis of Alzheimer's disease. Nature. 2019;570:332–7.
33. Rangaraju S, Levey DF, Nho K, Jain N, Andrews KD, Le-Niculescu H, et al. Mood, stress and longevity: convergence on ANK3. Mol Psychiatry. 2016;21:1037–49.
34. Smalheiser NR, Lugli G, Rizavi HS, Torvik VI, Turecki G, Dwivedi Y. MicroRNA expression is down-regulated and reorganized in prefrontal cortex of depressed suicide subjects. PLoS ONE. 2012;7:e33201.
35. Cabrera-Mendoza B, Fresno C, Monroy-Jaramillo N, Fries GR, Walss-Bass C, Glahn DC, et al. Sex differences in brain gene expression among suicide completers. J Affect Disord. 2020;267:67–77.
36. Amidfar M, Kim YK, Scaini G, Quevedo J. Evidence for additionally increased apoptosis in the peripheral blood mononuclear cells of major depressive patients with a high risk for suicide. Am J Med Genet B Neuropsychiatr Genet. 2018;177:388–96.

37. Niculescu AB, Le-Niculescu H, Levey DF, Roseberry K, Soe KC, Rogers J, et al. Towards precision medicine for pain: diagnostic biomarkers and repurposed drugs. Mol Psychiatry. 2019;24:501–22.

38. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. Nature communications. 2015;6:8570.

39. Harris SE, Riggio V, Evenden L, Gilchrist T, McCafferty S, Murphy L, et al. Age-related gene expression changes, and transcriptome wide association study of physical and cognitive aging traits, in the Lothian Birth Cohort 1936. Aging (Albany NY). 2017;9:2489–503.

40. Chen G, Zeng WZ, Yuan PX, Huang LD, Jiang YM, Zhao ZH, et al. The mood-stabilizing agents lithium and valproate robustly increase the levels of the neuroprotective protein bcl-2 in the CNS. J Neurochem. 1999;72:879–82.

41. Bar-Yosef T, Hussein W, Yitzhaki O, Damri O, Givon L, Marom C, et al. Mitochondrial function parameters as a tool for tailored drug treatment of an individual with psychosis: a proof of concept study. Sci Rep. 2020;10:12258.

42. Tamasi V, Petschner P, Adori C, Kirilly E, Ando RD, Tothfalusi L, et al. Transcriptional evidence for the role of chronic venlafaxine treatment in neurotrophic signaling and neuroplasticity including also Glutamatergic [corrected] - and insulin-mediated neuronal processes. PLoS ONE. 2014;9:e113662.

43. Engel D, Zomkowski AD, Lieberknecht V, Rodrigues AL, Gabilan NH. Chronic administration of duloxetine and mirtazapine downregulates proapoptotic proteins and upregulates neurotrophin gene expression in the hippocampus and cerebral cortex of mice. J Psychiatr Res. 2013;47:802–8.

44. Sasaki K, Linh TN, Hirano A, Tominaga K, Nukaga S, Nozaki H, et al. Microalgae extract induces antidepressant-like activity via neuroinflammation regulation and enhances the neurotransmitter system. Food Chem Toxicol. 2022;170:113508.

45. Reisi P, Eidelkhani N, Rafiee L, Kazemi M, Radahmadi M, Alaei H. Effects of doxepin on gene expressions of Bcl-2 family, TNF-alpha, MAP kinase 14, and Akt1 in the hippocampus of rats exposed to stress. Res Pharm Sci. 2017;12:15–20.

46. El-Agawy MSE, Badawy AMM, Rabei MR, Elshaer MMA, El Nashar EM, Alghamdi MA. et al. Methotrexate-Induced Alteration of Renal Aquaporins 1 and 2, Oxidative Stress and Tubular Apoptosis Can Be Attenuated by Omega-3 Fatty Acids Supplementation. Int J Mol Sci. 2022;23:12794.

47. Ruiz-Conca M, Gardela J, Mogas T, Lopez-Bejar M, Alvarez-Rodriguez M. Apoptosis and glucocorticoid-related genes mRNA expression is modulated by coenzyme Q10 supplementation during in vitro maturation and vitrification of bovine oocytes and cumulus cells. Theriogenology. 2022;192:62–72.

48. Campisi A, Sposito G, Pellitteri R, Santonocito D, Bisicchia J, Raciti G, et al. Effect of Unloaded and Curcumin-Loaded Solid Lipid Nanoparticles on Tissue Transglutaminase Isoforms Expression Levels in an Experimental Model of Alzheimer's Disease. Antioxidants (Basel) 2022;11.

49. Ghoneim FM, Abo-Elkhair SM, Elsamanoudy AZ, Shabaan DA Evaluation of Endothelial Dysfunction and Autophagy in Fibromyalgia-Related Vascular and Cerebral Cortical Changes and the Ameliorative Effect of Fisetin. Cells. 2021;11.

50. Khaksar S, Bigdeli M, Samiee A, Shirazi-Zand Z. Antioxidant and anti-apoptotic effects of cannabidiol in model of ischemic stroke in rats. Brain Res Bull. 2022;180:118–30.

51. Kim S, Choi KH, Baykiz AF, Gershenfeld HK. Suicide candidate genes associated with bipolar disorder and schizophrenia: An exploratory gene expression profiling analysis of post-mortem prefrontal cortex. BMC genomics. 2007;8:413.

52. Jiang CH, Tsien JZ, Schultz PG, Hu Y. The effects of aging on gene expression in the hypothalamus and cortex of mice. Proc Natl Acad Sci USA. 2001;98:1930–4.

53. Lu JY, Simon M, Zhao Y, Ablaeva J, Corson N, Choi Y, et al. Comparative transcriptomics reveals circadian and pluripotency networks as two pillars of longevity regulation. Cell Metab. 2022;34:836–856.e835.

54. Kurian SM, Le-Niculescu H, Patel SD, Bertram D, Davis J, Dike C, et al. Identification of blood biomarkers for psychosis using convergent functional genomics. Mol Psychiatry. 2011;36:37–58.

55. Hu VW, Frank BC, Heine S, Lee NH, Quackenbush J. Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes. BMC genomics. 2006;7:118.

56. Marchetti L, Lauria M, Caberlotto L, Musazzi L, Popoli M, Mathe AA, et al. Gene expression signature of antidepressant treatment response/non-response in Flinders Sensitive Line rats subjected to maternal separation. Eur Neuropsychopharmacol. 2020;31:69–85.

57. Yan J, Ma Y, Zhao F, Gu W, Jiao Y. Identification of immunomodulatory signatures induced by american ginseng in murine immune cells. Evid Based Complement Alternat Med. 2013;2013:972814.

58. Almousa LA, Salter AM, Castellanos M, May ST, Langley-Evans SC The Response of the Human Umbilical Vein Endothelial Cell Transcriptome to Variation in Magnesium Concentration. Nutrients. 2022;14.

59. Brunner HG, Nelen MR, van Zandvoort P, Abeling NG, van Gennip AH, Wolters EC, et al. X-linked borderline mental retardation with prominent behavioral disturbance: phenotype, genetic localization, and evidence for disturbed monoamine metabolism. Am J Hum Genet. 1993;52:1032–9.

60. Ho LW, Furlong RA, Rubinsztein JS, Walsh C, Paykel ES, Rubinsztein DC. Genetic associations with clinical characteristics in bipolar affective disorder and recurrent unipolar depressive disorder. Am J Med Genet. 2000;96:36–42.

61. Labonte B, Abdallah K, Maussion G, Yerko V, Yang J, Bittar T, et al. Regulation of impulsive and aggressive behaviours by a novel lncRNA. Mol Psychiatry. 2021;26:3751–64.

62. Azadmarzabadi E, Haghighatfard A, Mohammadi A. Low resilience to stress is associated with candidate gene expression alterations in the dopaminergic signalling pathway. Psychogeriatrics. 2018;18:190–201.

63. Domschke K, Tidow N, Kuithan H, Schwarte K, Klauke B, Ambree O, et al. Monoamine oxidase A gene DNA hypomethylation - a risk factor for panic disorder? Int J Neuropsychopharmacol. 2012;15:1217–28.

64. Melas PA, Wei Y, Wong CC, Sjoholm LK, Aberg E, Mill J, et al. Genetic and epigenetic associations of MAOA and NR3C1 with depression and childhood adversities. Int J Neuropsychopharmacol. 2013;16:1513–28.

65. Cabrera-Mendoza B, Martinez-Magana JJ, Monroy-Jaramillo N, Genis-Mendoza AD, Fresno C, Fries GR, et al. Candidate pharmacological treatments for substance use disorder and suicide identified by gene co-expression network-based drug repositioning. Am J Med Genet B Neuropsychiatr Genet. 2021;186:193–206.

66. Bond CM, Johnson JC, Chaudhary V, McCarthy EM, McWhorter ML, Woehrle NS. Perinatal fluoxetine exposure results in social deficits and reduced monoamine oxidase gene expression in mice. Brain Res. 2020;1727:146282.

67. Tariq U, Butt MS, Pasha I, Faisal MN. Neuroprotective effects of Olea europaea L. fruit extract against cigarette smoke-induced depressive-like behaviors in Sprague-Dawley rats. J Food Biochem. 2021;45:e14014.

68. Kim JJ, Sapio MR, Vazquez FA, Maric D, Loydpierson AJ, Ma W, et al. Transcriptional Activation, Deactivation and Rebound Patterns in Cortex, Hippocampus and Amygdala in Response to Ketamine Infusion in Rats. Front Mol Neurosci. 2022;15:892345.

69. Ziegler C, Richter J, Mahr M, Gajewska A, Schiele MA, Gehrmann A, et al. MAOA gene hypomethylation in panic disorder-reversibility of an epigenetic risk pattern by psychotherapy. Transl Psychiatry. 2016;6:e773.

70. Wu X, Luo Q, Zhao P, Chang W, Wang Y, Shu T, et al. JOSD1 inhibits mitochondrial apoptotic signalling to drive acquired chemoresistance in gynaecological cancer by stabilizing MCL1. Cell Death Differ. 2020;27:55–70.

71. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Miriami E, et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell. 2012;148:1293–307.

72. Niculescu AB, Le-Niculescu H, Roseberry K, Wang S, Hart J, Kaur A, et al. Blood biomarkers for memory: toward early detection of risk for Alzheimer disease, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2019.

73. Khan AJ, Campbell-Sills L, Sun X, Kessler RC, Adler AB, Jain S, et al. Association Between Responsibility for the Death of Others and Postdeployment Mental Health and Functioning in US Soldiers. JAMA Netw Open. 2021;4:e2130810.

74. Niwa M, Lockhart S, Wood DJ, Yang K, Francis-Oliveira J, Kin K, et al. Prolonged HPA axis dysregulation in postpartum depression associated with adverse early life experiences: A cross-species translational study. Nat Ment Health. 2024;2:593–604.

AUTHOR CONTRIBUTIONS

ABN designed the study and wrote the manuscript. RB, SG, HLN, KR, AG analyzed the data. EP, JC, CT, ZT organized, conducted, and scored testing in psychiatric subjects. DB, MY, GS assisted with coroner's cases. JM assisted with bioinformatics pipeline. MS assisted with sample report generation. AS assisted with data interpretation. SMK oversaw microarray and RNA sequencing experiments. CY and PB conducted machine learning analyses. All authors discussed the results and commented on the manuscript.

COMPETING INTERESTS

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-024-03071-y.

Correspondence and requests for materials should be addressed to A. B. Niculescu.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.