BMJ Open

Designing interventions guided by digital phenotype and pharmacogenetics in Spain for suicidal behaviour based on retrospective data: the multicentre SMARTomicS study protocol

Carmen Artés ,[1] Alejandro Porras-Segovia ,[2,3,4] Miguel Ruiz-Veguilla,[5,6] Lucas Giner,[6] Javier Garcia-Campayo,[7,8] Yolanda López del Hoyo,[8] Pilar Alejandra-Saiz,[9,10,11,12] Ainoa García-Fernández,[9,13,14] Juan José Martínez-Jambrina,[13,14,15,16] Rocío Villa-Díez,[9,12,13] Margalida Gili,[17] Miquel Roca,[17] Fernando de Andrés,[18,19] Víctor Pérez Sola,[20,21] Matilde E Elices ,[11,22] Iria Grande,[11,23,24,25] Ana Isabel Cebria ,[11,26,27] Maria Irigoyen-Otiñano,[11,28,29] Leonardo Galván-Santiago,[29,30] Eduardo Jesús Aguilar Garcia-Iturrospe,[11,31,32] Jesús Mesones-Peral,[33] Rafael Mora-Marín,[34,35] Pîlar Sierra,[31,36] Eva Penas-Lledo,[37,38] Adrian LLerena,[11,37,38] Jorge Lopez-Castroman,[11,39,40] Raquel Cruz,[41,42] Manuel Arrojo,[43,44] Santiago Ovejero,[3,44] Alejandro de la Torre-Luque,[11,45] Ángela Palao-Tarrero,[44,46] Concepción Vaquero-Lorenzo,[47] Claudio Toma,[44,48,49,50] Maria Luisa Barrigon,[11,51,52,53] Antonio Artes,[11,53,54,55] Marina Díaz-Marsá,[11,51,56,57] Wala Ayad-Ahmed,[51] Ángela Ibáñez,[11,58,59,60] Carmen Díaz-Sastre,[59,60] Mercè Salvador-Robert,[61,62] Raquel Álvarez-García,[62,63] Laura Mata-Iturralde,[3] Sergio Benavente-López,[3,64] Laura Muñoz-Lorenzo,[3] Ezequiel Fernando Di Stasio,[65] Rodrigo Carmona-Camacho,[3,45] Inmaculada Peñuelas-Calvo,[11,51,66] Miguel Valencia,[67,68] Purificación López Peña,[11,69,70,71] Lucía Albarracín-García,[1,3,4,64] Enrique Baca-Garcia [1,3,4,11,44,64,65]

To cite: Artés C, Porras-Segovia A, Ruiz-Veguilla M, *et al*. Designing interventions guided by digital phenotype and pharmacogenetics in Spain for suicidal behaviour based on retrospective data: the multicentre SMARTomicS study protocol. *BMJ Open* 2026;**16**:e106685. doi:10.1136/bmjopen-2025-106685

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2025-106685).

Received 17 June 2025
Accepted 01 April 2026

For numbered affiliations see end of article.

Correspondence to
Professor Enrique Baca-Garcia; research.fjd@gmail.com

ABSTRACT

Introduction Each year, suicide claims approximately 700 000 lives worldwide and generates a significant financial burden. Integrating genomic data, exposomic factors and digital phenotypes can enhance the development of short-term predictive models. Current knowledge and available tools provide the basis for designing personalised treatment strategies that incorporate real-time interventions to prevent suicide attempt recurrence cost-effectively. This study aims to develop a predictive algorithm for suicidal behaviour integrating psychiatric assessments, genetic risk markers, digital phenotypes and exposomic data.

Methods and analysis This protocol describes a retrospective multicentre study that will recruit participants with a clinical history of suicide across 25 hospitals across Spain with a catchment area of 8.6 million people (17.8% of Spain's population). Our sample target is over 5000 participants, aged over 12 years old, ensuring 93.5% statistical power for genetic analysis. Eligible participants must be over 12 years old. Data collection will include psychiatric assessments, biospecimen collections (DNA, RNA, plasma and serum), Google Takeout data for digital phenotyping, and a standardised set of administrative and

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study integrates digital phenotyping, clinical data and genetic profiles to characterise the exposome of suicide, a comprehensive mapping of environmental and biological exposures linked to suicide risk.
⇒ The inclusion of multi-omics data and its integration with medical and pharmacological patient history enables an innovative and comprehensive approach to identifying biological markers of suicide risk, treatment effectiveness and prognosis.
⇒ Passive and active digital phenotyping via smartphone and web-based data provides a novel, real-time assessment of behavioural factors linked to suicide risk.
⇒ Psychological distress and reluctance to share smartphone data may reduce participation rates, particularly in individuals with recent suicidal behaviour, which may potentially introduce sample bias.
⇒ Data from deceased individuals could not be collected directly, as data will be collected from individuals who survived suicide attempts, limiting the scope of digital and biological data acquisition in those most affected.

clinical data registered for each patient. Genotyping will be performed with the Axiom Spanish array (>750 000 markers), and genome-wide association studies (GWAS) will be performed after genetic imputation in a whole sample of >10 000 individuals (5000 suicide attempters; 5000 controls). Prescription and clinical history will also be retrospectively integrated, and codified data statistics forms will periodically be sent to the Government. Statistical analyses will combine traditional regression models and AI-based algorithms to identify predictive behavioural, genomic profiles, and digital markers of suicidal behaviour. Cost-effectiveness analyses of pharmacogenomic markers for antidepressant response will also be conducted.

By successfully implementing this project, we aim to help reduce suicide reattempts and lessen the emotional and economic burden on families and the healthcare system.

Ethics and dissemination This study has been approved by the Ethics Committee of the Fundación Jiménez Díaz (PIC301-24_FJD) and complies with the Declaration of Helsinki. It adheres to the GDPR (EU Regulation 2016/679), Spain's Organic Law 3/2018 on Personal Data Protection and Digital Rights, and Law 41/2002 on patient autonomy. All required data protection measures will be implemented, including those under Real Decreto 1718/2010 on prescriptions and treatment adherence. Underaged participants will require parental consent for participation. The results will be disseminated through publication in peer-reviewed scientific journals and presentations at psychiatric conferences.

Trial registration number NCT07422090.

INTRODUCTION

The impact of suicide

Each year, suicide claims approximately 700 000 lives worldwide, yet despite extensive prevention efforts, rates have not significantly declined.[1] For every suicide death, an estimated 20 attempts occur, underscoring the critical need for effective prevention strategies. The combined direct and indirect costs of suicide-related issues in Europe amount to €150 billion annually.[2] Moreover, suicide profoundly affects families, communities and society. In Spain, suicide is the leading cause of death among individuals aged 15–29 years, accounting for 129 481 years of potential life lost (YPLL).[2 3] In 2020 alone, suicide was responsible for 80 679 YPLL in Spain, surpassing the impact of COVID-19 (63 417 YPLL).[4]

Suicide is closely linked to mental health disorders, particularly depression.[5 6] Both are unresolved public health crises that impose a significant economic burden on healthcare systems.[7] The WHO emphasises the need for economic evaluations of suicide prevention strategies.[8] However, there is still limited data on the cost-effectiveness of these interventions.[9]

One of the strongest predictors of suicide is a prior attempt, with 20% of individuals who attempt suicide showing reattempts,[10 11] and a 12-month suicide fatality rate of 3.2% among US adult attempters.[12] Reattempt rates vary across studies, with systematic reviews estimating 15%–16% at 1 year[13] and 23%–37% at 5 years or longer[14]; notably, up to 88% of reattempts occur within the first 2 years, peaking during the initial 6–12 months following the index attempt.[15] Given the high recurrence rate, secondary prevention—targeting individuals with a history of suicide attempts—plays a crucial role in reducing suicide-related deaths.[16]

Advancing knowledge: detection and prognosis

Suicide prevention, as a critical public health issue, needs a comprehensive strategy that includes universal (population-wide), selective (targeting at-risk groups) and indicated (focusing on individuals exhibiting suicidal behaviours such as self-harm, ideation or attempts) interventions.[17] The success of these interventions depends on accurately identifying high-risk individuals—a task complicated by the multifactorial nature of suicide risk. Multiple interrelated factors, such as heritability (30%–50%), family history, biological markers, mental health disorders and various environmental and personal adversities, contribute to this risk.[18 19] Despite extensive research, our understanding of these factors has changed little over the past 50 years, as highlighted by Franklin *et al*.[20] Meta-analysis, and current risk stratification protocols often perform no better than chance.[21] Moreover, key factors—including genetic, behavioural and proximal determinants—remain inadequately characterised. Therefore, there is an urgent need for improved short-term prediction methods and personalised prognostic tools to enhance targeted suicide prevention interventions.[22]

Unexplored factors in suicide risk: genetics and sociodemographic factors

No single genetic cause of suicide has been identified, but twin and family studies[23] suggest that genetic factors, combined with environmental and psychological influences, may increase vulnerability to suicidal behaviour. Among the most promising genetic targets for understanding this psychiatric trait are genes involved in serotonin neurotransmission. Specifically, certain polymorphisms in the *SLC6A4* gene, which encodes the serotonin transporter, have been linked to an increased risk of suicide.[24]

However, suicide likely follows a highly polygenic model, driven by the cumulative effect of thousands of common variants with small effect sizes under an additive model. The largest genome-wide association studies (GWAS) meta-analysis to date, conducted by the Psychiatric Genomics Consortium, analysed 29 782 suicide attempters and 519 961 controls, identifying two genome-wide significant loci.[25] These findings highlight the complex genetic architecture underlying suicide and the need for large cohorts to improve statistical power. Additionally, other genetic variants, such as single nucleotide variants (SNVs) and copy number variants (CNVs), may contribute to suicide risk, although no large studies have been performed thus far. Future research should clarify how polygenic risk interacts with environmental and psychological factors to predispose individuals to suicidal behaviour.

Suicidal behaviour also exhibits sex and gender differences, with suicide rates being approximately three times

higher in men than in women, while suicide attempt rates are four times higher in women than in men.[3] Additionally, Lesbian, Gay, Bisexual, Transgender, Queer or Questioning, and others individuals show higher suicide rates compared with the general population.[26] However, gender is rarely considered a key variable in study designs, stratification or data analysis, highlighting a critical gap in suicide research.

Enhancing suicide behaviour characterisation: digital phenotype

Recent evidence indicates that lifestyle factors—particularly physical activity and sleep quality—play a significant role in reducing suicide risk. These factors are promising targets for real-time monitoring via digital phenotyping, an emerging approach that uses behavioural markers to establish a digital signature of mental health conditions.[22 27 28]

A straightforward way to capture these behavioural markers is through Google Takeout, a tool that allows users to download and export data stored in Google services. These data record movement patterns, distance travelled, locations visited and app interactions, offering objective insights into suicidal behaviour. Although exact figures are unavailable, Gmail is estimated to have 4.1 billion users, with 2.5 billion active[29] indicating the widespread use of Google accounts. Therefore, Google Takeout offers an efficient means to collect passive digital behavioural data in large population samples.

Previous studies have linked digital behaviour indicators to mental health (digital phenotype), with our research team finding an inverse relationship between physical activity during the COVID-19 lockdown and mood.[30] Recent findings from our research team have also demonstrated the predictive potential of passive data, achieving an area under the curve (AUC) of 0.77 for 1 week suicidal risk prediction.[31] While prior research has relied on ecological momentary assessment (EMA) or smartphone sensors,[32–34] Google Takeout remains an unexplored tool for digital phenotyping in this context.

Knowledge frontiers: treatment approaches

Effective treatments specifically targeting suicidal behaviour remain limited, largely due to the scarcity of clinical trials using suicide risk as a therapeutic endpoint. Most psychopharmacological trials exclude individuals with suicidal behaviour, limiting the available evidence on treatment efficacy.

Regarding pharmacological treatments, antidepressants provide short-term protective effects against suicidal ideation, yet further research is needed to understand the underlying mechanisms of selective serotonin reuptake inhibitors (SSRIs).[35] Growing evidence supports the potential benefits of ketamine and esketamine in rapidly reducing suicidal ideation.[36] Digital health records provide valuable insights into proximal risk factors, such as medication changes, which may influence suicidal behaviour.[37]

For non-pharmacological interventions, cognitive therapy, dialectical behaviour therapy (DBT), cognitive-behavioural therapy (CBT) and brief interventions show greater efficacy than standard care in preventing suicide attempts.[38] Further research is needed to optimise delivery methods (in-person or remotely), identify the most effective intervention strategies and assess their long-term impact.

The exposome

Beyond individual risk factors, their combined influence could approach us to an exposome—the totality of environmental and biological exposures that shape physical and mental health risks.[39] A characterisation of suicidal profiles and exposome determination could improve the identification of at-risk individuals to enhance treatment effectiveness. This approach has been explored by Visoki and colleagues[40] in youth cohorts using environmental and lifestyle factors and found that exposomic scores for suicide attempts offer a generalisable approach to risk classification. Exposomes have also been widely used in other health fields like cancer, age-related diseases and environmental health.[41–43] Most studies integrate high-resolution metabolomics and other multi-omics techniques with exposure analysis, bioinformatics, behavioural monitoring and monitoring devices.[44]

Integrating genetic data, exposomic factors and digital phenotypes can enhance the development of short-term predictive models, with gender inclusion playing a crucial role in improving accuracy. Current knowledge and available tools provide the foundation for designing personalised treatment strategies that incorporate real-time interventions to prevent suicide attempt recurrence cost-effectively.

Objectives

This study aims to develop a predictive algorithm for suicidal behaviour integrating genetic markers, digital phenotypes and exposome data. To achieve this, we will genotype patients with a history of suicide attempts to analyse their genetic profiles and identify associated risk factors. Blood samples will be processed for biospecimen collections and future genotyping, next-generation sequencing, methylation array and metabolic profiles to establish potential biomarkers via integrative omics analyses.

Secondary objectives include evaluating behavioural factors and digital phenotypes linked to suicide risk drawing from smartphone and web-based data. Additionally, we aim to investigate the relationship between medication changes, medical history and suicidal behaviour using data from electronic prescriptions and digital health records. Additionally, we will assess the impact of pharmacological interventions in conjunction with digital phenotyping.

By successfully implementing this project, we aim to reduce suicide rates, improve the quality of life for at-risk

individuals and lessen the emotional and economic burden on families and the healthcare system.

METHODS AND ANALYSIS

Study design and setting

This protocol describes an observational retrospective multi-site study. Patients will be recruited from 26 hospitals throughout Spain, which represent more than 8.6 million people, over 17.8% of the total Spanish population (see online supplemental material for the full list of participating hospitals and the 14 adjoining research entities). Patient recruitment will begin in three hospital units: emergency department, short-stay inpatient unit and outpatient mental health clinics for all hospitals.

Sample

Our study will recruit patients who attend mental health services and who have a history of at least one suicide attempt in their lifetime. Following informed consent, we will collect their electronic health records for further analysis.

Inclusion criteria are as follows:

1. At least one suicide attempt (without a fatal outcome) in their lifetime.
2. The attempt must have occurred when the participant was older than 12 years old.
3. Participants must be at least 18 years old or have parental consent if aged 12–17.

Individuals will be excluded if a medical contraindication prevents blood sample collection or if they are unable to provide informed consent to participate in the study.

Sample size calculation

This study is part of a consortium aiming to collect at least 5000 participants across Spain. Due to the participation of 19 centres across Spain, covering a catchment area of nearly 10 million people. Most centres have established suicidal treatment programmes, each following 300 to 4000 patients with a history of suicide attempts. Previous research on genetic markers and digital phenotypes suggests that this sample size is necessary to account for key variables such as gender, diagnosis and treatment response. Based on data from Li's[35] meta-analysis, which included 8315 cases across five cohorts, a sample of 5000 cases is projected to provide 93.5% statistical power at a 5×10^{-8} significance level, assuming a minor allele frequency (MAF) of 0.2 and an OR of 1.3.

Measures

A multifaceted array of information will be gathered from participants to build a complete patient profile and identify suicide risk factors. Baseline data will be collected from patients to assess the participant's initial state in a single visit or multiple sessions if needed. It is comprised of three measures: a genetic sample collection for genetic and metabolic analysis, a sociodemographic and health survey carried out by a mental health clinician and consent to Google Takeout data retrieval to analyse digital behaviour patterns.

To complete the health profiles of the participants, retrospective clinical data will be integrated into the study database. These data consist of the patient's electronic health record, pharmaceutical treatments and statistics from the National Institute of Statistics (Instituto Nacional de Estadística, INE) regarding socioeconomic and environmental factors. The participant's digital phenotype analysis will integrate geographically and chronologically referenced data from the INE to establish a digital phenotype. Data from the patient's electronic health record will include the Basic Minimum Set of Data (BMSD), a standardised set of administrative and clinical data registered for each patient encounter used in healthcare facilities.

Sociodemographic data

Participants will be asked questions about their gender identity, birthdate, nationality and marital status. Additional questions will examine employment status, education, living arrangements and whether the individual receives any form of financial assistance. We will also include questions on family composition, such as the number of biological and adopted children.

Medical and psychiatric information

Using validated tools to explore psychiatric history, current symptoms and risk factors, including:

- Medical history: Questions about chronic and acute medical conditions, categorised by cardiovascular, neurological, endocrine, respiratory, oncological and others.
- Psychiatric diagnoses following International Classification of Diseases, 10th Revision (ICD-10) criteria will be established by clinicians and drawn from electronic clinical records, including mood disorders, anxiety disorders, psychotic disorders, substance use disorders, eating disorders and personality disorders.
- Suicidal ideation and behaviour (Columbia-Suicide Severity Rating Scale—C-SSRS)[45]: This questionnaire assesses past and recent suicidal thoughts, intent, planning and attempts. It evaluates frequency, severity and potential risk factors, including protective elements (such as family support or religious beliefs) and preparatory behaviours (eg, acquiring means or writing suicide notes).
- Stressful life events[46]: This questionnaire identifies major stressors, such as serious illnesses, financial crises, legal issues, job loss or bereavement.

This evaluation provides a multidimensional understanding of the participant's profile, identifying risk factors and patterns associated with suicidal behaviour.

Biospecimen collection and genomic data

Blood samples will be collected from all participants for biospecimen collection. Blood samples enable DNA and RNA extraction and allow the isolation of plasma and

serum for lipidomics, metabolomics and pharmacological assays.

With this purpose, four tubes will be drawn and stored at 4°C for a maximum of 24 hours before biospecimen processing. The sample includes three ethylenediaminetetraacetic acid (EDTA) tubes (two 5 mL and one 10 mL EDTA K2 or K3) and a serum-separator tube, facilitating the isolation of DNA, RNA, plasma, proteins and serum. DNA genotyping will be performed at healthcare centres, and samples will be sent to designated laboratories for genotyping. This will take place through the Axiom Spanish array (758 740 markers, including single nucleotide polymorphisms (SNPs) and a subset of 114 898 SNVs in the Spanish population). This complete set of genome-wide SNPs will be used to calculate a Polygenic Risk Score (PRS), which is a numeric value representing the additive contribution of hundreds of risk alleles. The PRS offers a measure for the genetic burden of common variants (SNPs) using risk alleles already identified in the GWAS summary statistics from the Psychiatric Genomics Consortium (PGC) for suicide attempts.[26] These per-individual data will be integrated with clinical outcomes in future combined analyses. Additionally, ancestry-specific associations in the Spanish population will be performed through a GWAS, comparing allelic frequencies between >5000 suicide attempters and 5000 controls. 9–10 million SNPs (MAF >0.01) obtained after imputation in TOPMed Imputation Server will be analysed. The control population includes healthy donors from the Banco Nacional de ADN and the Madrid Manic Group Cohort.[47 48] Pharmacological profiles will be studied to provide a comprehensive approach to help uncover genetic factors influencing suicidal behaviour.

Digital phenotype

To build the digital phenotype of participants, clinicians will ask them to share their Google Takeout data. Google Takeout is a Google service that allows users to export and download personal data stored across various Google services, such as Google Maps, Gmail, Google Contacts or Google Calendar. This tool provides objective, indirect insights into user behaviour patterns through their interactions with these platforms. This study will extract the following data categories from Google Takeout: phone audio, Chrome browsing history, Gmail, Google Fit, location history (timeline), Google Maps, YouTube and YouTube Music.

Among these sources, geolocation and time-stamped (chrono-referenced) data derived from Google Maps Timeline constitute the primary digital phenotyping indicators. These data enable the quantification of mobility patterns (eg, daily distance travelled, time spent at home, variability in routine) and temporal behavioural rhythms. Importantly, geospatial and temporal markers will allow linkage with external environmental datasets, including climatological and atmospheric exposure records, facilitating integration with exposomic variables. Additional extracted metadata (eg, browsing categories, search terms related to mental health, physical activity logs and communication frequency indicators) will be operationalised into predefined behavioural features reflecting activity levels, sleep–wake patterns, social interaction proxies and health-related information-seeking behaviour.

To ensure privacy, all collected data will be anonymised before analysis. Patients will personally select the data they consent to share, with support from a research assistant. The selected data will be temporarily transferred to a Dropbox folder, where a script will apply pseudonymisation. The processed data will then be securely stored on servers managed by the Universidad Carlos III, maintaining pseudonymisation protocols to ensure participant confidentiality.

Analysing these patterns may help identify correlations between digital behaviour and mental health conditions. For example, a decrease in mobility data or an increase in searches related to mental health symptoms could serve as early indicators of psychological distress.

Electronic health records

Information from the electronic health records will be gathered retrospectively, focusing on diagnosis, general consultations, urgency visits and hospitalisation instances.

We will use this tool to collect the BMSD data set, which includes sex, age and diagnoses, coded according to the International Classification of Diseases (ICD). Additionally, data on psychiatry and psychology consultations, hospitalisations, emergency department visits and suicide attempts treated in the emergency department (including their dates) will be recorded. The Clinical Global Impression (CGI) scale, which is integrated into the Mental Health form and completed by clinicians,[49] will also be included.

Digital pharmacological history

Prescription history from the different Spanish autonomous communities will be used. These records allow authorised healthcare professionals to access a patient's prescription history, ensuring continuous monitoring of medication initiation, suspension, dosage adjustments and treatment changes. Information gathered will include patients' identification and health history codes, dates for treatment start and end, diagnosis and drug description, including doses. Moreover, frequency and administration route will be recorded, and whether patients collect their medications from the pharmacy will be tracked, providing an objective measure of treatment adherence.

Exposome

We will draw information from the INE, which provides geo-referenced and time-referenced data on several variables, such as crime rates and mortality, broken down by neighbourhoods and other territorial divisions. By combining this dataset with the patient's digital footprints, researchers can analyse individual behavioural patterns in relation to aggregated national or local statistics. An aggregated dataset of the environmental,

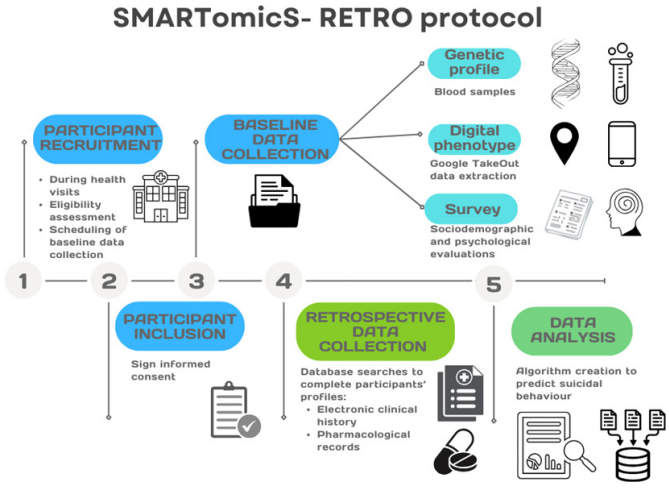

Figure 1 Study diagram.

sociodemographic and genetic factors for each patient will be generated to build the exposomes.

Procedure

The study will be divided into five phases (described in figure 1): recruitment, patient inclusion, baseline data collection, retrospective data collection and an algorithm creation. The three first phases may occur on the same day as the participant's healthcare visit or be scheduled for a later date in case of logistical constraints. This will be followed by retrospective data collection from the electronic health records to further characterise the participants and data analysis.

The study will take place over 2 years, starting in 2025, and the workflow is depicted in figure 2. Before patient recruitment, the healthcare professionals' preparation and formation will occur. Recruitment and sample collection will be followed by database processing to merge baseline and retrospective data before data analysis and scientific diffusion.

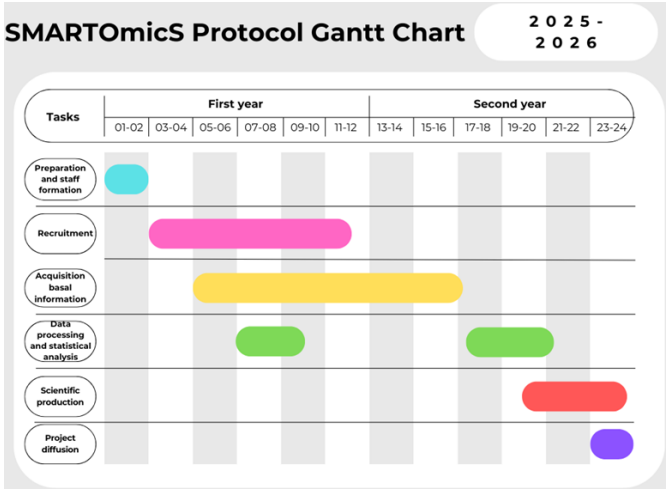

Figure 2 Gantt chart of the workflow of the study.

Patient recruitment

Patient recruitment will take place in person at healthcare facilities. It will be conducted by psychiatrists, clinical psychologists and mental health nurses in outpatient mental health clinics, emergency departments and short-stay hospitalisation units. When patients attend the designated health service, clinicians will check for eligibility and, if feasible, explain the project and data protection policies and offer them to participate.

Data storage

Clinicians will store survey data in the MEmind digital environment—a digital platform (https://frontend.memind.net) used to record survey responses for both sociodemographic and psychological assessments. Google Takeout data will be transferred to Dropbox anonymously for data analysis. Genetic sampling will be labelled according to the participant, who will be then offered two options regarding any excess sample: the leftover samples can either be destroyed or stored for future research use in the 'National Biobank Registry, Collections Section' which is managed by Dr Baca (C.0000278). For further use in research studies, approval by the Ethics Committee will be required, along with new informed consent if the research is unrelated to the original study.

Data confidentiality

Pseudonymisation involves assigning each participant a coded identifier that removes any direct identifying elements. To ensure data confidentiality, the research team responsible for pseudonymisation will be separate from the coordinating team. Only the designated team will retain the information necessary for potential re-identification, ensuring compliance with privacy regulations. Patient data will remain strictly confidential throughout the study. A secure tool will link participant codes to medical record numbers via an external, access-restricted file stored separately from the main study database.

These security measures ensure strict compliance with data protection regulations, maintaining confidentiality and integrity throughout the study.

Outcomes

The primary outcomes are suicidal behaviour, psychological and behavioural assessments, and pharmacological treatment changes. Suicidal behaviour will be measured through the clinical history of the patients in terms of visits to the emergency department, suicide attempts codified in the hospital's discharge reports, and deaths by suicide. Psychological and behavioural assessments will be measured using the C-SSRS and Brugha Scales,[45 46] and the digital phenotype, which will be obtained through Google Takeout reports. Pharmacological treatment information will be collected from the hospital's electronic health records, and treatment decisions will be guided by the genetic report.

Our secondary outcomes include pharmacological treatment efficacy and the cost-effectiveness of genetic markers to guide antidepressant treatments.

Statistical analysis

Traditional statistics

SPSS v. 29.0 software will be used to carry out the traditional analysis. Logistic regressions will be performed to explore the factors associated with our outcomes, and a multivariate regression model will be built to explore independent associations. All tests will be two-tailed with statistical significance set at $p<0.05$, with 95% CIs.

Non-traditional statistics

Advanced statistical and machine learning methods will be employed to identify digital behavioural biomarkers. Longitudinal trajectories and delayed environmental effects will be analysed using generalised linear mixed models, including distributed lag non-linear models (DLNM) to account for time-dependent exposomic influences. Time-to-reattempt will be modelled using a parametric survival model (Weibull or Gompertz, based on model fit).

To integrate high-dimensional multi-omics and digital phenotyping data, deep neural network architectures will be implemented alongside conventional statistical models to allow comparison of predictive performance and interpretability.

Feature selection will follow a structured pipeline. Initially, random forest models with permutation-based Gini importance will be used to identify candidate predictors. Subsequently, least absolute shrinkage and selection operator (LASSO) regularisation will be applied to reduce dimensionality, followed by correlation analyses (Pearson or Spearman, as appropriate) to remove multicollinearity.

The cohort (n=5000) will be divided into training (60%), validation (20%) and test (20%) sets. Model development will incorporate nested fivefold cross-validation repeated 10 times, with bootstrapping procedures (1000 resamples) to optimise hyperparameters and estimate internal stability. Model performance will be assessed using area under the curve – receiver operating characteristic (AUC-ROC), sensitivity, specificity, and F1-score, aiming for an AUC-ROC of >0.80 and an F1-score of above 75%.

Distinct suicide-related personalised profiles will be described and compared using genetic, metabolic, medical and digital phenotyping data. Finally, a personalised anonymised exposome will be constructed, incorporating geo-temporal data from Google Takeout, along with information from public sociodemographic databases (INE).

Limitations

Despite the innovative and interdisciplinary approach of this study, which integrates multifactorial data for suicide risk prediction, certain limitations must be considered.

First, participant sampling may be subject to selection bias, as recruitment relies on medical staff offering patients the opportunity to participate. However, training sessions and standardised guidelines are provided to healthcare professionals to ensure adherence to the study protocol. Additionally, the completeness of patient records may vary across participating healthcare centres. Given that psychological distress is associated with lower study participation rates,[50] this factor may influence the study sample, particularly given participants' history of suicidal behaviour. Nevertheless, previous psychiatric studies using smartphone-based monitoring have reported participation rates of 80%.[28]

Another potential limitation is the requirement for participants to have a Google account, as well as possible reluctance to share their Google Takeout data, despite strict protocols ensuring data security and anonymity. Furthermore, data from individuals who have died by suicide cannot be collected, introducing an inherent sampling bias. However, these cases will be accounted for through available clinical records.

PATIENT AND PUBLIC INVOLVEMENT

Patients and the public will not be involved in its design, recruitment or conduct. However, findings will be disseminated to relevant stakeholders, including patient advocacy groups, through institutional websites and scientific meetings.

ETHICS AND DISSEMINATION

Ethical considerations

This study will be conducted in accordance with the Declaration of Helsinki and all applicable ethical and legal regulations, and the protocol has been registered in clinicaltrials.gov. This study has been approved by the Ethics Committee of the Fundación Jiménez Díaz (PIC301-24_FJD). Compliance is ensured with the GDPR (EU Regulation 2016/679), Spain's Organic Law 3/2018 on Personal Data Protection and Digital Rights, and the provisions outlined in Law 41/2002 on patient autonomy, rights and clinical documentation. Additionally, all security and data protection measures required by relevant legislation will be implemented.

The national requirements for general prescriptions (BOE 1718/2010, Real Decreto 1718/2010, 17 de diciembre) allow for monitoring information regarding pharmacological prescriptions, treatments and adherence. The protocol for this study has been registered in ClinicalTrials.gov.

Informed consent

After their eligibility is confirmed, patients will receive a detailed explanation of the project. Underaged participants will require parental consent. Those interested will be given time to review the information sheet, ask questions and sign the informed consent. The information

sheet contains details regarding their participation and what it entails, biological sample collection (blood), the benefits and risks of the study, data protection and confidentiality. Participation is entirely voluntary, and consent can be withdrawn at any time. No financial incentives will be provided. After the informed consent has been signed, participants will be scheduled for the interview and sample collection.

Dissemination plans

Advancements in technology have enhanced the feasibility of effective suicide prevention—a critical societal goal. This study aims to develop a predictive algorithm for suicide risk by integrating genetic, metabolic, environmental and digital phenotype characteristics associated with suicidal behaviour. The findings will be analysed and disseminated through peer-reviewed publications and presentations at scientific conferences and forums.

Author affiliations

[1]Translational Psychiatry Research Group, Instituto de Investigacion Sanitaria de la Fundacion Jimenez Diaz, Madrid, Community of Madrid, Spain
[2]Instituto de Investigacion Sanitaria de la Fundacion Jimenez Diaz, Madrid, Spain
[3]Department of Psychiatry, Hospital Universitario Fundacion Jimenez Diaz, Madrid, Community of Madrid, Spain
[4]Department of Psychiatry, Hospital Universitario Rey Juan Carlos, Móstoles, Community of Madrid, Spain
[5]Psychiatry, Hospital Universitario Virgen del Rocio, Sevilla, Spain
[6]Department of Psychiatry, Universidad de Sevilla, Seville, Spain
[7]Hospital Universitario Miguel Servet, Zaragoza, Aragon, Spain
[8]Universidad de Zaragoza, Zaragoza, Aragon, Spain
[9]Department of Psychiatry, Universidad de Oviedo, Oviedo, Asturias, Spain
[10](ISPA), Instituto de Investigación Sanitaria del Principado de Asturias, Oviedo, Asturias, Spain
[11](CIBERSAM), Carlos III Institute of Health, Centro de Investigación Biomédica en Red de Salud Mental, Madrid, Community of Madrid, Spain
[12](SESPA), Servicio de Salud del Principado de Asturias, Oviedo, Spain
[13]Instituto de Investigación Sanitaria del Principado de Asturias, Oviedo, Asturias, Spain
[14](INEUROPA), Instituto de Neurociencias del Principado de Asturias, Oviedo, Spain
[15]Hospital San Agustín de Avilés, Aviles, Asturias, Spain
[16]Servicio de Salud del Principado de Asturias, Oviedo, Spain
[17]Instituto de Investigación Sanitaria de las Islas Baleares (IDISBA), Universitat de les Illes Balears, Palma, Spain
[18]Departamento de Química Analítica y Tecnología de Alimentos, Facultad de Farmacia, Universidad de Castilla-La Mancha, Ciudad Real, Spain
[19]Instituto Regional de Investigación Científica Aplicada (IRICA), Universidad de Castilla-La Mancha, Ciudad Real, Spain
[20]Hospital del Mar, Hospital del Mar Medical Research Institute, Barcelona, Spain
[21]Universitat Pompeu Fabra, Barcelona, Spain
[22]Hospital del Mar Medical Research Institute, Barcelona, Spain
[23]Bipolar and Depressive Disorders Unit, Hospital Clinic de Barcelona, Barcelona, Spain
[24]Departament de Medicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Barcelona, Spain
[25](IDIBAPS), Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain
[26]Universitat Autònoma de Barcelona, Barcelona, Spain
[27]Consorci Corporació Sanitària Parc Taulí, Sabadell, Spain
[28]Hospital Universitario de Santa María, Lleida, Spain
[29](IRB Lleida), Institut de Recerca Biomèdica de Lleida Fundació Dr Pifarré, Lleida, Spain
[30]Departamento de Farmacia, Regio Sanitaria Lleida, Lleida, Spain
[31]Universidad de Valencia, Valencia, Valencian Community, Spain
[32]INCLIVA, Hospital Clinico Universitario, Valencia, Valencian Community, Spain
[33]Hospital Universitario de Torrevieja, Torrevieja, Valencian Community, Spain
[34]Consorcio Hospitalario Provincial de Castellón, Hospital Provincial Castellón, Castellón de la Plana, Valencian Community, Spain
[35]Universitat Jaume I, Castellón de la Plana, Valencian Community, Spain
[36]Hospital Politécnico y Universitario La Fe, Valencia, Valencian Community, Spain
[37]Facultad de Medicina y Ciencias de la Salud, Universidad de Extremadura, Badajoz, Spain
[38]Instituto Universitario de Investigación Biosanitaria de Extremadura (INUBE), Badajoz, Spain
[39]Universidade de Santiago de Compostela, Santiago de Compostela, Spain
[40]University Hospital Centre Nimes, Nimes, Occitanie, France
[41]Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, Spain
[42](CIBERER), Centro de Investigación Biomédica en Red de Enfermedades Raras, Valencia, Valencian Community, Spain
[43](SERGAS), Servicio Gallego de Salud, Santiago de Compostela, Spain
[44]Universidad Autónoma de Madrid, Madrid, Community of Madrid, Spain
[45]Universidad Complutense de Madrid, Madrid, Spain
[46]Hospital Universitario La Paz, Madrid, Community of Madrid, Spain
[47](IUBM-UAM), Instituto de Biología Molecular, Madrid, Community of Madrid, Spain
[48](CBMSO), Centro de Biología Molecular Severo Ochoa, Madrid, Community of Madrid, Spain
[49](CSIC), Consejo Superior de Investigaciones Cientificas, Madrid, Community of Madrid, Spain
[50]Neuroscience Research Australia, Randwick, New South Wales, Australia
[51]Universidad Complutense de Madrid, Madrid, Community of Madrid, Spain
[52]Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, Madrid, Community of Madrid, Spain
[53](IISGM), Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Community of Madrid, Spain
[54]Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Leganés, Madrid, Spain
[55]Evidence-Based Behavior, Madrid, Community of Madrid, Spain
[56]Hospital Clínico San Carlos, Madrid, Community of Madrid, Spain
[57]Sociedad Española de Psiquiatría y Salud Mental, Madrid, Community of Madrid, Spain
[58]Universidad de Alcalá de Henares, Alcala de Henares, Spain
[59]Hospital Universitario Ramón y Cajal, Madrid, Community of Madrid, Spain
[60](IRYCIS), Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Community of Madrid, Spain
[61]Hospital Universitario de Móstoles, Móstoles, Community of Madrid, Spain
[62]Universidad Rey Juan Carlos, Móstoles, Community of Madrid, Spain
[63]Hospital Universitario Rey Juan Carlos, Móstoles, Community of Madrid, Spain
[64]Hospital Universitario Infanta Elena, Valdemoro, Community of Madrid, Spain
[65]Department of Psychiatry, Hospital Universitario General de Villalba, Collado Villalba, Spain
[66]Department of Child and Adolescent Psychiatry, Hospital Universitario 12 de Octubre, Madrid, Community of Madrid, Spain
[67]IdiSNA, Instituto de Investigacion Sanitaria de Navarra, Pamplona, Navarre, Spain
[68]CIMA Universidad de Navarra, Pamplona, Navarre, Spain
[69]Hospital Universitario Araba, Vitoria-Gasteiz, Spain
[70]Instituto de Investigación Sanitaria Bioaraba, Vitoria-Gasteiz, Spain
[71]Universidad del País Vasco, Leioa, Spain

Acknowledgements We gratefully acknowledge the support of CIBER—Consorcio Centro de Investigación Biomédica en Red—(CB/07/09/0025), the Instituto de Salud Carlos III with the support of the European Regional Development Fund (ISCIII PI23/00614; PMP24/00026), by Fundació La Marató de TV3 (202226-31) and by CaixaResearch Health 2023 LCF/PR/HR23/52430033.

Contributors EB-G, MLB, CV-L and EMP-L were in charge of conceptualisation and funding, and CA and AP-S of writing the original draft. Supervision, revision and editing were carried out by the previously mentioned authors and the remaining authors: MR-V, LG, JG-C, YL-H, PA-S, AG-F, JJM-J, RV-D, MG, MR-B, FA, VP-S, ME, IG, AIC-M, MI-O, LG-S, EJA-GI, JM-P, RM-M, PS, AL, JL-C, RC, MA, SO, AT-L, AP-T, CT, AA-R, MD-M, WA-A, AI, CD-S, MS-R, RA-G, LM-I, SB-L, LM-L, EF-S, RC-C, IP-C, MV, PL-P and LA-G. Guarantor is EB-G. The corresponding author, EB-G, on behalf of all authors, confirms the contribution of all authors and that the manuscript has been read and approved by all authors, and that the authorship requirements established by the ICMJE have been met.

Funding This research was supported by CIBER -Consorcio Centro de Investigación Biomédica en Red- (CB/07/09/0025), the Instituto de Salud Carlos III

with the support of the European Regional Development Fund (ISCIII PI23/00614; PMP24/00026), by Fundació La Marató de TV3 (202226-31) and by CaixaResearch Health 2023 LCF/PR/HR23/52430033.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Carmen Artés https://orcid.org/0009-0006-8284-5614
Alejandro Porras-Segovia https://orcid.org/0000-0003-2019-9099
Matilde E Elices https://orcid.org/0000-0002-7034-4242
Ana Isabel Cebria https://orcid.org/0000-0002-2632-8130
Enrique Baca-Garcia https://orcid.org/0000-0002-6963-6555

REFERENCES

1 World Health Organization. Suicide. WHO. Available: https://www.who.int/news-room/fact-sheets/detail/suicide [Accessed 12 Mar 2025].
2 Zalsman G, Hawton K, Wasserman D, *et al*. Suicide prevention strategies revisited: 10-year systematic review. *Lancet Psychiatry* 2016;3:646–59.
3 Binder EB. New Insights Into Suicidal Behavior From Large Multi-Ancestry Genetic Meta-Analysis. *AJP* 2023;180:705–7.
4 Merayo-Cano JM, Porras-Segovia AA, Baca-García E. COVID-19 impact vs. suicide impact in Spain. *Span J Psychiatry Ment Health* 2023;16:129–30.
5 Brådvik L. Suicide Risk and Mental Disorders. *Int J Environ Res Public Health* 2018;15:2028.
6 Oliva V, Roberto N, Andreo-Jover J, *et al*. Anxious and depressive symptoms and health-related quality of life in a cohort of people who recently attempted suicide: A network analysis. *J Affect Disord* 2024;355:210–9.
7 Henry M. Suicide prevention: A multisectorial public health concern. *Prev Med* 2021;152:106772.
8 World Health Organization. *National suicide prevention strategies: progress, examples and indicators*. Geneva: World Health Organization, 2018.
9 Martínez-Alés G, Cruz Rodríguez JB, Lázaro P, *et al*. Cost-effectiveness of a Contact Intervention and a Psychotherapeutic Program for Post-discharge Suicide Prevention. *Can J Psychiatry* 2021;66:737–46.
10 de la Torre-Luque A, Pemau A, Ayad-Ahmed W, *et al*. Risk of suicide attempt repetition after an index attempt: A systematic review and meta-analysis. *Gen Hosp Psychiatry* 2023;81:51–6.
11 Pemau A, de la Torre-Luque A, Marin-Martin C, *et al*. Symptomatic networks in suicide attempt and reattempt: Relevance of psychiatric comorbidity. *Eur Psychiatr* 2025;68:e4.
12 Han B, Kott PS, Hughes A, *et al*. Estimating the rates of deaths by suicide among adults who attempt suicide in the United States. *J Psychiatr Res* 2016;77:125–33.
13 Owens D, Horrocks J, House A. Fatal and non-fatal repetition of self-harm. Systematic review. *Br J Psychiatry* 2002;181:193–9.
14 Nicolau-Subires E, Irigoyen-Otiñano M, Arenas-Pijoan L, *et al*. Five-Year Outcomes of First Suicide Attempts: Insights on Lethality, Recurrence, and Mortality. *J Clin Psychiatry* 2025;86:24m15754.
15 Parra-Uribe I, Blasco-Fontecilla H, Garcia-Parés G, *et al*. Risk of re-attempts and suicide death after a suicide attempt: A survival analysis. *BMC Psychiatry* 2017;17:163.
16 Ghanbari B, Malakouti SK, Nojomi M, *et al*. Suicide Prevention and Follow-Up Services: A Narrative Review. *Glob J Health Sci* 2015;8:145–53.
17 Pirkis J, Gunnell D, Hawton K, *et al*. A Public Health, Whole-of-Government Approach to National Suicide Prevention Strategies. *Crisis* 2023;44:85–92.
18 Turecki G, Brent DA, Gunnell D, *et al*. Suicide and suicide risk. *Nat Rev Dis Primers* 2019;5:74.
19 Sher L, Oquendo MA. Suicide. *Medical Clinics of North America* 2023;107:119–30.
20 Franklin JC, Ribeiro JD, Fox KR, *et al*. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull* 2017;143:187–232.
21 Mughal F, Burton FM, Fletcher H, *et al*. New guidance for self-harm: an opportunity not to be missed: commentary, Mughal et al. *Br J Psychiatry* 2024;225:340.
22 Barrigon ML, Courtet P, Oquendo M, *et al*. Precision Medicine and Suicide: an Opportunity for Digital Health. *Curr Psychiatry Rep* 2019;21:131.
23 DiBlasi E, Kang J, Docherty AR. Genetic contributions to suicidal thoughts and behaviors. *Psychol Med* 2021;51:2148–55.
24 Bhagar R, Gill SS, Le-Niculescu H, *et al*. Next-generation precision medicine for suicidality prevention. *Transl Psychiatry* 2024;14:362.
25 Mullins N, Kang J, Campos AI, *et al*. Dissecting the Shared Genetic Architecture of Suicide Attempt, Psychiatric Disorders, and Known Risk Factors. *Biol Psychiatry* 2022;91:313–27.
26 Pfister A. Suizidprävention für LGBTQ+-Jugendliche: Notwendigkeit, Modell und Zugänge. *Präv Gesundheitsf* 2025;20:13–9.
27 Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. *JAMA* 2017;318:1215–6.
28 Porras-Segovia A, Díaz-Oliván I, Barrigón ML, *et al*. Real-world feasibility and acceptability of real-time suicide risk monitoring via smartphones: A 6-month follow-up cohort. *J Psychiatr Res* 2022;149:145–54.
29 Statista. Number of e-mail users worldwide from 2017 to 2028. Statista; 2024. Available: https://www.statista.com/statistics/255080/number-of-e-mail-users-worldwide/ [Accessed 12 Mar 2025].
30 Norbury A, Liu SH, Campaña-Montes JJ, *et al*. Social media and smartphone app use predicts maintenance of physical activity during Covid-19 enforced isolation in psychiatric outpatients. *Mol Psychiatry* 2021;26:3920–30.
31 Barrigon ML, Romero-Medrano L, Moreno-Muñoz P, *et al*. One-Week Suicide Risk Prediction Using Real-Time Smartphone Monitoring: Prospective Cohort Study. *J Med Internet Res* 2023;25:e43719.
32 Barrigon ML, Porras-Segovia A, Courtet P, *et al*. Smartphone-based Ecological Momentary Intervention for secondary prevention of suicidal thoughts and behaviour: protocol for the SmartCrisis V.2.0 randomised clinical trial. *BMJ Open* 2022;12:e051807.
33 Abascal-Peiró S, Peñuelas-Calvo I, Alacreu-Crespo A, *et al*. Digital Platform for the Prevention of Suicidal Behaviour and Non-Suicidal Self-Injuries in Adolescents: The SmartCrisis-Teen Study Protocol. *Behav Sci* 2024;14:740.
34 Mournet AM, Ruork AK, Kleiman EM, *et al*. An actigraphy case series: Sleep disturbance associations with suicidal ideation, self-harm and desire to live. *Psychiatry Res Case Rep* 2025;4:100244.
35 Li QS, Shabalin AA, DiBlasi E, *et al*. Genome-wide association study meta-analysis of suicide death and suicidal behavior. *Mol Psychiatry* 2023;28:891–900.
36 Wang Y-T, Wang X-L, Lei L, *et al*. A systematic review and meta-analysis of the efficacy of ketamine and esketamine on suicidal ideation in treatment-resistant depression. *Eur J Clin Pharmacol* 2024;80:287–96.
37 Jeon SM, Lim H, Cheon H, *et al*. Assessing the Labeling Information on Drugs Associated With Suicide Risk: Systematic Review. *JMIR Public Health Surveill* 2024;10:e49755.
38 Hu F-H, Xu J, Jia Y-J, *et al*. Non-pharmacological interventions for preventing suicide attempts: A systematic review and network meta-analysis. *Asian J Psychiatr* 2024;93:103913.
39 Rejek M, Misiak B. The Associations of Exposome Score with Various Domains of Psychopathology: A Network Analysis in a Non-Clinical Sample. *Brain Sci* 2024;14:242.
40 Visoki E, Moore TM, Zhang X, *et al*. Classification of Suicide Attempt Risk Using Environmental and Lifestyle Factors in 3 Large Youth Cohorts. *JAMA Psychiatry* 2024;81:1020.

41 Vineis P, Chadeau-Hyam M, Gmuender H, *et al*. The exposome in practice: Design of the EXPOsOMICS project. *Int J Hyg Environ Health* 2017;220:142–51.
42 Jones DP, Cohn BA. A vision for exposome epidemiology: The pregnancy exposome in relation to breast cancer in the Child Health and Development Studies. *Reprod Toxicol* 2020;92:4–10.
43 Ding E, Wang Y, Liu J, *et al*. A review on the application of the exposome paradigm to unveil the environmental determinants of age-related diseases. *Hum Genomics* 2022;16:54.
44 Haddad N, Andrianou XD, Makris KC. A Scoping Review on the Characteristics of Human Exposome Studies. *Curr Pollution Rep* 2019;5:378–93.
45 Posner K, Brown GK, Stanley B, *et al*. The Columbia–Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings From Three Multisite Studies With Adolescents and Adults. *Am J Psychiatry* 2011;168:1266–77.
46 Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand* 1990;82:77–81.
47 de Coo A, Cruz R, Quintela I, *et al*. Genome-wide association study of stage III/IV grade C periodontitis (former aggressive periodontitis) in a Spanish population. *J Clin Periodontol* 2021;48:896–906.
48 Toma C, García LA, Ortiz IG, *et al*. F46. The Madrid manic group (MadManic): a novel Spanish cohort for clinical and genomic studies in bipolar disorder and suicide. *Eur Neuropsychopharmacol* 2024;87:230.
49 Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont)* 2007;4:28–37.
50 Tsubota-Utsugi M, Yonekura Y, Suzuki R, *et al*. Psychological Distress in Responders and Nonresponders in a 5-year Follow-up Health Survey: The RIAS Study. *J Epidemiol* 2022;32:527–34.