Multicentric Assessment of a Multimorbidity Adjusted Disability Score to stratify depression-related risks using temporal disease maps

Rubèn González-Colom, Kangkana Mitra, Emili Vela, Andras Gezsi, Teemu Paajanen, Zsófia Gál, Gabor Hullam, Hannu Mäkinen, Tamas Nagy, Mikko Kuokkanen, Jordi Piera-Jiménez, Josep Roca, Peter Antal, Gabriella Juhasz, Isaac Cano

Submitted to: Journal of Medical Internet Research
on: September 27, 2023

Table of Contents

Multicentric Assessment of a Multimorbidity Adjusted Disability Score to stratify depression-related risks using temporal disease maps

Rubèn González-Colom[1] MSc; Kangkana Mitra[1] MSc; Emili Vela[2, 3] MSc; Andras Gezsi[4] PhD; Teemu Paajanen[5] PhD; Zsófia Gál[6, 7] MSc; Gabor Hullam[4, 7] PhD; Hannu Mäkinen[5] PhD; Tamas Nagy[4, 6, 7] MSc; Mikko Kuokkanen[5, 8, 9] PhD; Jordi Piera-Jiménez[2, 3, 10] PhD; Josep Roca[1, 11, 12] MD, PhD; Peter Antal[4*] PhD; Gabriella Juhasz[6, 7*] PhD; Isaac Cano[1, 12*] PhD

[1]Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi i Sunyer (FRCB-IDIBAPS) Barcelona ES
[2]Digitalization for the Sustainability of the Healthcare (DS3) - IDIBELL Barcelona ES
[3]Catalan Health Service Barcelona ES
[4]Department of Measurement and Information Systems, Budapest University of Technology and Economics Budapest HU
[5]Department of Public Health and Welfare, Finnish Health and Welfare Institute Helsinki FI
[6]Department of Pharmacodynamics, Faculty of Pharmacy, Semmelweis University Budapest HU
[7]NAP3.0-SE, Neuropsychopharmacology Research Group, Hungarian Brain Research Program, Semmelweis University Budapest HU
[8]Department of Human Genetics and South Texas Diabetes and Obesity Institute, School of Medicine at University of Texas Rio Grande Valley Brownsville US
[9]Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki Helsinki FI
[10]Faculty of Informatics, Telecommunications and Multimedia, Universitat Oberta de Catalunya Barcelona ES
[11]Hospital Clínic de Barcelona Barcelona ES
[12]Universitat de Barcelona Barcelona ES
[*]these authors contributed equally

Corresponding Author:
Rubèn González-Colom MSc
Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi i Sunyer (FRCB-IDIBAPS)
C/Rosselló 149-153
Barcelona
ES

Abstract

Background: Multimorbidity management, a growing healthcare concern, necessitates precise health risk assessment (HRA) tools to increase the efficacy of its interventions and mitigate the disease burden. However, existing solutions often fall short of accurately predicting disease progression and the emergence of new comorbid conditions, hindering the implementation of preventive measures. In contrast, research on disease trajectories has provided valuable insights into the temporal patterns of disease occurrence, enabling the identification of causal relationships between concurrent diseases. The integration of these areas of study is crucial for developing next-generation health risk assessment tools that comprehensively consider the current burden of morbidity and the risk of multimorbidity progression based on disease trajectories.

Objective: Utilizing the major depressive disorder (MDD) as use case, the research aimed at generating a novel HRA tool to identify at-risk citizens. Allowing to: 1) Quantify the impact of MDD and its comorbidities on individuals and healthcare systems. And 2) Anticipate multimorbidity progression; thereby facilitating the development of preventive strategies.

Methods: In the EU project TRAJECTOME, we used a novel methodology for filtering disease-disease indirect associations and identifying temporal disease maps of depression and highly prevalent co-occurring disease conditions. This information was combined with disability weights established by the Global Burden of Disease Study 2019 to create a depression-related HRA tool, the Multimorbidity Adjusted Disability Score (MADS). MADS was used to independently stratify over one million cases from three different cohorts from Spain, UK and Finland; and evaluate the correspondence among the different risk strata and the impact on the mortality rates, utilisation of healthcare resources, pharmacological burden, healthcare expenditure and multimorbidity progression.

Results: Results indicate statistically significant associations between MADS risk strata and increased mortality rate (P <.001), heightened healthcare utilization (i.e. primary care visits P <.001; specialized care outpatient consultations P <.001; visits in

mental health specialized centres P <.001; emergency room visits P <.001; hospitalizations P <.001), increased pharmacological (P <.001) and non-pharmacological expenditures (P <.001), and a raised pharmacological burden (antipsychotics P <.001; anxiolytics P <.001; hypnotics and sedatives P <.001; antidepressants P <.001). The analysis revealed an augmented risk of disease progression within the high-risk groups, as indicated by a heightened incidence of new-onset depression-related illnesses within a 12-month period after MADS assessment.

Conclusions: MADS seems to be a promising approach to predict depression-related health risks, and estimate multimorbidity-adjusted risk of disease progression, which can be tested in other diseases; nevertheless, clinical validation is still necessary.

(JMIR Preprints 27/09/2023:53162)
DOI: https://doi.org/10.2196/preprints.53162

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Multicentric Assessment of a Multimorbidity Adjusted Disability Score to stratify depression-related risks using temporal disease maps.

Rubèn González-Colom[1], *Kangkana Mitra*[1], *Emili Vela*[2,3], *Andras Gezsi*[4], *Teemu Paajanen*[5], Zsofia Gal[6,7], Gabor Hullam[4,7], *Hannu Mäkinen*[5], Tamas Nagy[4,6,7], *Mikko Kuokkanen*[5,8,9], *Jordi Piera-Jiménez*[2,3,9], *Josep Roca*[1,11,12], *Peter Antal*[4,#], *Gabriella Juhasz*[6,7,#] *and Isaac Cano*[1,12,#]

1) Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi i Sunyer (FRCB-IDIBAPS). Barcelona, Spain.
2) Catalan Health Service. Barcelona, Spain
3) Digitalization for the Sustainability of the Healthcare (DS3) - IDIBELL. Barcelona, Spain
4) Department of Measurement and Information Systems, Budapest University of Technology and Economics. Budapest, Hungary
5) Department of Public Health and Welfare, Finnish Health and Welfare Institute. Helsinki, Finland.
6) Department of Pharmacodynamics, Faculty of Pharmacy, Semmelweis University. Budapest, Hungary.
7) NAP3.0-SE, Neuropsychopharmacology Research Group, Hungarian Brain Research Program, Semmelweis University. Budapest, Hungary
8) Department of Human Genetics and South Texas Diabetes and Obesity Institute, School of Medicine at University of Texas Rio Grande Valley, Brownsville, United States.
9) Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki. Helsinki, Finland.
10) Faculty of Informatics, Telecommunications and Multimedia, Universitat Oberta de Catalunya. Barcelona, Spain.
11) Hospital Clínic de Barcelona. Barcelona, Spain.
12) Universitat de Barcelona. Barcelona, Spain.

# - The authors contributed equally.

Corresponding Author: Rubèn González-Colom
C/Rosselló 149-153, 08036, Barcelona
+34 686626947
rgonzalezc@recerca.clinic.cat

ABSTRACT

Background: Comprehensive management of multimorbidity can significantly benefit from advanced health risk assessment tools that facilitate value-based interventions, allowing the assessment and prediction of disease progression. Our study proposes a novel methodology, the Multimorbidity Adjusted Disability Score (MADS), which integrates disease trajectory methodologies with advanced techniques for assessing interdependencies among concurrent diseases. This approach is designed to better assess the clinical burden of clusters of interrelated diseases and to enhance our ability to anticipate disease progression, thereby potentially informing targeted preventive care interventions.

Objective: This study aims to evaluate the effectiveness of MADS in stratifying patients into clinically relevant risk groups based on their multimorbidity profiles, which accurately reflect their clinical complexity and the probabilities of developing new associated disease conditions.

Methods: In a retrospective multicentric cohort study, we developed MADS by analyzing disease trajectories and applying Bayesian statistics to determine disease-disease probabilities combined with well-established disability weights. We used Major Depressive Disorder a primary case study for this evaluation. We stratified patients into different risk levels corresponding to different percentiles of MADS distribution. We statistically assessed the association of MADS risk strata with mortality, healthcare resources utilization, and disease progression across 1M individuals from Spain, the UK, and Finland.

Results: The results revealed significantly different distributions of the assessed outcomes across the MADS risk tiers, including mortality rates, primary care visits, specialized care outpatient consultations, visits in mental health specialized centres, emergency room visits, hospitalizations, pharmacological and non-pharmacological expenditures and dispensation of antipsychotics anxiolytics, sedatives, and antidepressants (all P<.001). Moreover, the results of the pairwise comparisons between adjacent risk tiers illustrate a significant and gradual pattern of increased mortality rate, heightened healthcare utilization, increased healthcare expenditures and a raised pharmacological burden as individuals progress from lower MADS risk tiers to higher risk tiers. The analysis also revealed an augmented risk of multimorbidity progression within the high-risk groups, aligned with a higher incidence of new onsets of MDD-related diseases.

Conclusions: MADS seems a promising approach for predicting health risks associated with multimorbidity. It might complement current risk assessment state-of-the-art tools by providing valuable insights for tailored epidemiologic impact analyses of clusters of interrelated diseases and by accurately assessing multimorbidity progression risks. The study paves the way for innovative digital developments to support advanced health risk assessment strategies. Further validation is required to generalize its use beyond the initial case study of MDD.

Keywords*:* Health Risk Assessment, Multimorbidity, Disease Trajectories, Major Depressive Disorder.

INTRODUCTION

The co-occurrence of multiple chronic diseases, known as multimorbidity[1], affects one in three adults. Its prevalence rises with age, affecting 60% of individuals aged between 65 and 74 years and escalating to 80% among those aged 85 years and older[2]. Due to its association with poor prognosis, functional impairment, and reduced quality of life, multimorbidity is considered a global healthcare challenge[3,4] tied to complex clinical situations, leading to increased encounters with healthcare professionals, hospitalizations, and pharmacological prescriptions, resulting in a substantial rise in healthcare costs[5]. The emergence of multimorbidity is not arbitrary and frequently aligns with shared risk factors and underlying pathophysiological mechanisms[6–8] that result from complex interactions between genetic and environmental factors throughout the lifespan[9]. Perceiving diseases not in isolation but as integral components of a more extensive, interconnected system within the human body has led to the emergence of network medicine[10,11]. The network-medicine analyses disease co-occurrence patterns, aiming to understand the complex connections between diseases to uncover biomarkers, therapeutic targets, and potential interventions[12,13]. The studies investigating the temporal patterns of disease concurrence, or disease trajectories[14,15] rely on a pragmatic approach of this concept to yield a better understanding of the time-dependent relationships among diseases and establish a promising landscape to identify disease-disease causal relationships.

According to this paradigm, a disease-centered approach might lead to suboptimal treatment of patients with multiple chronic conditions, triggering the need to implement new tools to enhance the effectiveness of health services[16]. In this regard, multimorbidity-adjusted Health Risk Assessment (HRA) tools[17–21], such as the morbidity groupers, are crucial for assessing the comprehensive health needs of multimorbid patients[22]. HRA utilizes algorithms and patient data to categorize individuals by risk, aiding healthcare professionals in customizing interventions, optimizing resource allocation, and enhancing patient outcomes through preventive care. HRA tools facilitate efficient case-finding and screening processes[23]. Case finding targets the most vulnerable individuals at high risk, which is crucial for specialized healthcare programs. Whereas, patient screening detects latent illnesses early, enabling cost-effective interventions to prevent disease progression and reduce healthcare demands.

However, despite their widespread utilization, prevailing population-based HRA tools such as the Adjusted Clinical Groups[24] (ACG), the Clinical Risk Groups[25] (CRG) or the Adjusted Morbidity Groups (AMG)[4,21] still do not incorporate information on disease trajectories in their calculations.

The AMG system is currently used in Catalonia (ES; 7M inhabitants), for health policy and clinical purposes. Adding disease-disease association information into the AMG (or other morbidity groupers) may open new avenues for implementing epidemiologic impact analyses concerning clusters of interrelated diseases. Additionally, it may facilitate the construction of risk groups that accurately represent probabilities of developing new associated disease conditions[26] susceptible to early prevention.

While acknowledging current limitations, the study seeks to explore the feasibility of incorporating procedures relevant to the study of disease trajectories[14,15] and novel techniques for analyzing dependency relationships between concomitant diseases[27,28] to improve the capabilities of the current morbidity groupers. This approach might better adjust the estimations of the burden of morbidity to clusters of diseases and improve the ability to anticipate the progression of multimorbidity.

We utilized Major Depressive Disorder (MDD) [F32-F33 ICD-10-CM[29]] as a use case due to its clinical relevance in multimorbidity management. However, the study pursues to showcase a methodology applicable beyond MDD, allowing the assessment of the impact of multimorbidity across different clusters of diseases.

The current observational retrospective multicentric cohort study describes the process of development and assessment of the Multimorbidity Adjusted Disability Score (MADS), showcasing a pioneering approach that integrates advanced techniques for analyzing disease associations, insights from the analysis of disease trajectories, and a comprehensive scoring method aimed at evaluating the disease burden. MADS was designed to stratify patients with different health needs according to 1) the disease burden caused by MDD and its comorbidities on individuals and health systems, and 2) the risk of morbidity progression and the onset of MDD comorbid conditions.

Based on the temporal disease maps among MDD and highly prevalent disease conditions[30] generated using Bayesian Direct Multimorbidity Maps (BDMMs)[27,28], a promising method for filtering indirect disease associations, in the context of the ERAPERMED EU project TRAJECTOME[31], we combined the probabilities of relevance (PR) among MDD and its comorbid conditions with the disability weights (DW)[32] documented in the 2019 revision of the Global Burden of Diseases study (GBD) to compute MADS. We used MADS to generate a risk pyramid and stratify the study population into five risk groups using different percentiles of MADS distribution. Finally,

we analyzed the correspondence between the MADS risk groups and health outcomes through a cross-sectional analysis of mortality and utilization of healthcare resources, and a longitudinal analysis of disease prevalence and incidence of new disease onsets. The clinical relevance of the identified risk groups was assessed through a multicentric assessment of the findings. To this end, MADS performance was analyzed using the data of three independent European cohorts from the UK, Finland, and Spain, including more than one million individuals.

METHODS

The development and evaluation of MADS involved the following steps (**Figure 1**):

Step 1 - Computing age-dependent disease-disease PR using the BDMM method in four age intervals (0-20, 0-40, 0-60, and 0-70 years). This analysis resulted in an inhomogeneous dynamic Bayesian network that determined the PR for MDD against the most prevalent co-occurring diseases in the three European cohorts considered in TRAJECTOME, namely: The Catalan Health Surveillance System (CHSS)[33], the UK Biobank (UKB)[34], and The Finnish National Institute for Health and Welfare cohort (THL)[35]. THL cohort amalgamates information from Finrisk[36] 1992, 1997, 2002, 2007, 2012, Finhealth[37] 2017 and Health[38] 2000/2011 studies.

Step 2 – Combining the PR of every disease condition assessed in the study with their corresponding DW, extracted from the GBD 2019 study, we estimated the morbidity burden caused by MDD and its comorbid conditions. MADS was computed following a multiplicative combination of PR and DW of all the disease conditions present in an individual.

Step 3 - Using MADS to stratify patients into different risk levels corresponding to different percentiles of the population-based risk pyramid of each patient cohort: 1) Very low risk ($\leq P_{50}$); 2) Low risk (P_{50}-P_{80}]; 3) Moderate risk (P_{80}-P_{90}]; 4) High risk (P_{90}-P_{95}]; 5) Very high risk ($> P_{99}$).

Step 4 - Finally, the correspondence between the MADS risk strata and health outcomes was analyzed through a cross-sectional analysis of utilization of healthcare resources, mortality, pharmacological burden, and healthcare expenditure, and a longitudinal analysis of disease prevalence and incidence of new disease onsets. The results were validated through a multicentric

replication of the findings in the three study cohorts, including 1,041,014 individuals.

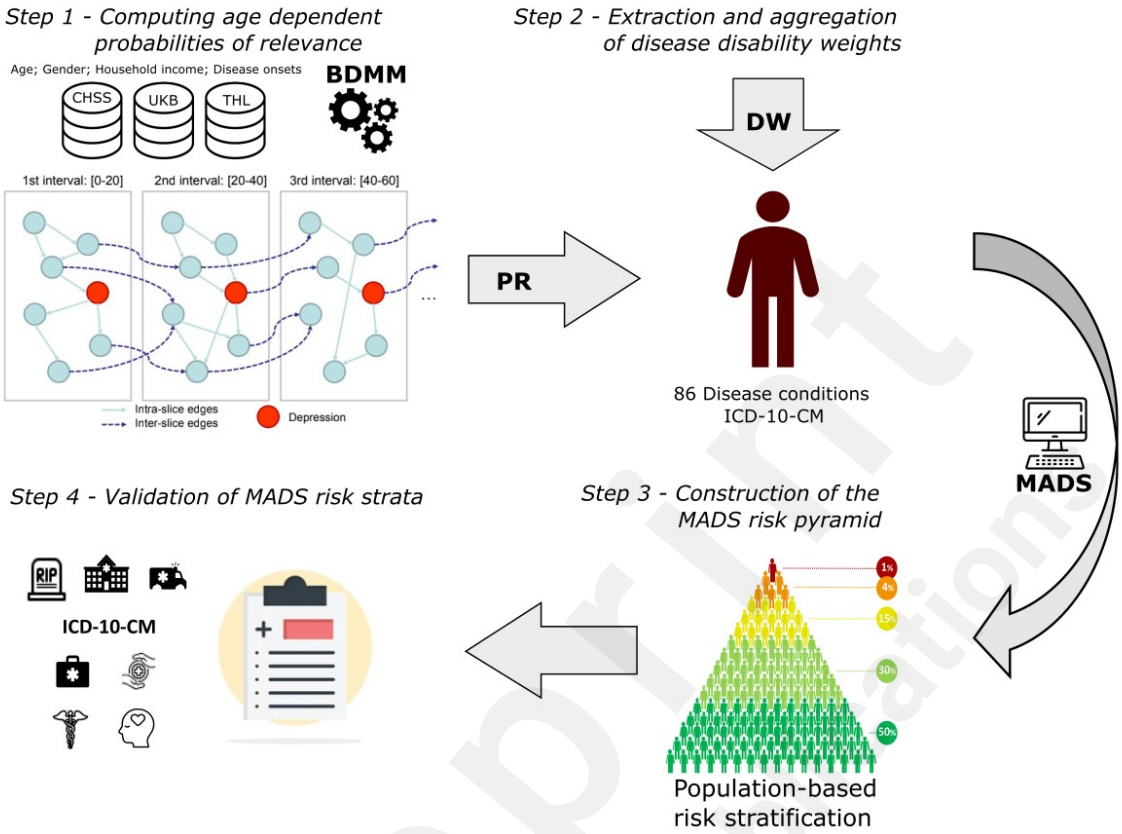


Figure 1 - Workflow for building and assessing MADS. *BDMM stands for Bayesian Direct Multimorbidity Maps, PR for Probabilities of Relevance, and DW for Disability Weights.*

STEP 1- Computing age-dependent probabilities of relevance

BDMMs were used to assess direct and indirect associations between MDD and 86 potential comorbid conditions. The set of 86 disease conditions considered in the study had a prevalence greater than 1% in all the study cohorts. The list of diseases and their associated ICD-10-CM[29] codes are displayed in the **Supplementary material – Appendix 3**.

This step considered information on: 1) **Disease diagnosis**: Disease conditions were catalogued using the first three characters of ICD-10-CM codes; 2) **Age at disease onset time**: The age at disease onset corresponds to the first diagnosis in a lifetime for each ICD-10-CM code; 3) **Sex**; and, 4) **Socio-economic status**: annual average total household income (before tax with co-payment exemption) as a categorical variable with three categories: a) Less than 18,000; b) 18,000 to 100,000; c) Greater than 100,000. Thresholds are given in EUR.

BDMM analysis resulted in an inhomogeneous dynamic Bayesian network, which was utilized to compute temporal PR, ranging from 0 (no association) to 1 (strong association), for MDD in conjunction with sex, socio-economic status, and the set of 86 predetermined consensual diseases[30]. To construct the trajectories, the PR was calculated in four different age ranges: 0-20, 0-40, 0-60, and 0-70 years of age. The PR calculated and utilized for MADS computation are reported in the **Supplementary material – Appendix 4**. Further details regarding the core analysis conducted in TRAJECTOME can be found in [30].

STEP 2 - Extraction and aggregation of disease DWs

MADS was developed by weighting the DWs of single diseases according to their estimated PR against MDD. DWs indicate the degree of health loss based on several health outcomes and are used to indicate of the total disability caused by a certain health condition or disease. Often, the DWs present specific disability scores tailored to the severity of the disease. The disease categories, severity distribution and associated DWs utilized in this study were extracted from the GBD studies 2019 and reported in the **Supplementary material – Appendix 3**.

DWs were extracted and aggregated as follows: 1) We considered only the DW of MDD and the set of 86 disease codes; 2) We considered the DW of all the chronic conditions diagnosed in patients' lifetime, whereas, since the disability caused by acute illnesses is transitory, the DWs for the acute diseases diagnosed more than 12 months before the MADS assessment were arbitrary set to 0 (no disability); 3) Due to the unavailability of information on the severity of diagnoses, we determined the DWs of each disease condition by calculating the weighted mean of the DWs associated to the disease severity categories and their prevalence. In instances where the severity distribution was not available, we computed the arithmetic mean of the DWs of each severity category; 4) We finally weighted the DWs according to the PR of each disease condition with respect to MDD. The PR were adjusted according to the age of disease onset, discretized in the following intervals 0-20, 20-40, 40-60, and >60 years old.

Since the DWs do not account for multimorbidity in their estimates, the utilization of DW independently can cause inaccuracies in the burden of disease estimations, particularly in ageing populations that include large proportions of persons with two or more disabling disease conditions[39].

Consequently, we combined the DW and the PR for all the disease conditions present in one individual following a multiplicative approach (**Eq. 1**)[40], aggregating several DW in a single score that accounts for the overall disability caused by numerous concurrent chronic conditions in which every comorbid disease increases the utility loss of a patient, though it is less than the sum of the utility loss of both diseases independently.

(Eq. 1)

$$MADS = 1 - \prod_{k=i}^{n} (1 - PR_i * DW_i)$$

"DW" stands for Disability Weight, "PR" Probability of Relevance and "n" is the number of diseases present in one individual.

The MADS pseudocode is reported in the **Supplementary material – Appendix 5**

STEP 3 - Construction of the MADS risk pyramid

Once calculated, MADS was utilized to stratify patients in different levels of risk according to the percentiles (P) of its distribution in the source population, producing the following risk pyramid: 1) Very low risk ($\leq P_{50}$); 2) Low risk (P_{50}-P_{80}]; 3) Moderate risk (P_{80}-P_{90}]; 4) High risk (P_{90}-P_{95}]; 5) Very high risk ($> P_{99}$).

STEP 4 - Evaluation of MADS risk strata

The clinical relevance of the risk strata was assessed through two interconnected analyses: 1) A cross-sectional analysis of health outcomes; and 2) a longitudinal analysis of disease prevalence and incidence of new onsets.

Cross-sectional analysis of health outcomes and use of healthcare resources

To validate the results of MADS, we conducted a cross-sectional analysis of clinical outcomes within the 12 months following the MADS assessment. The burden of MDD and its comorbidities on patients and healthcare providers, corresponding to each risk group of the MADS risk pyramid, was assessed using the following features (the parameters evaluated in each cohort may vary depending on the availability of the requested information in the source databases):

1) **Prescriptions of psycholeptic and psychoanaleptic drugs** (Information available in all the databases) The prescribed medication was catalogued using the first four characters from ATC[41] codes, resulting in the following categories: *Antipsychotics (N05A), Anxiolytics (N05B), Hypnotics and sedatives (N05C) and Antidepressants (N06A).*
2) **Cost of the pharmacological prescriptions in €** *(*Information available only in *CHSS and THL).*
3) **Mortality rates** *(*Information available only in *CHSS and THL).*
4) **Contacts and encounters with healthcare professionals** *(*Information available only in *CHSS)* Encompassing: i) primary care visits; ii) specialized care outpatient visits; iii) ambulatory visits in mental health centers; iv) emergency room visits; v) planned and unplanned hospital admissions; and vi) admissions in mental health centers.
5) ***Total healthcare expenditure*** *(*Information available only in *CHSS)* Including: i) direct healthcare delivery costs; ii) pharmacological costs; and iii) other billable healthcare costs, such as non-urgent medical transportation, ambulatory rehabilitation, domiciliary oxygen therapy, and dialysis.

We assessed the effect of sex and age, replicating the analyses disaggregated by sex and age. The age ranges were discretized into the following categories: 0-20, 20-40, 40-60, and >60 years.

Longitudinal analysis of disease prevalence and incidence of new onsets

To address the age-dependency of disease onsets, we performed a longitudinal analysis of the prevalence of a target disease and the incidence of new diagnostics within the five years following the MADS assessment.

We iteratively computed MADS in five-year intervals throughout the patients' lives. Within each interval, the population was stratified based on the MADS distribution. Subsequently, within each risk tier, the prevalence of the target disease and the incidence of new disease onset over the subsequent five years were calculated. Only individuals with complete information for the next interval at each time point of the analysis were included.

In the analysis, we considered only the chronic disease conditions with a PR against MDD ≥ 0.80 in

at least one of the four age intervals assessed, namely: 0-20, 0-40, 0-60 and 0-70**.** It was resulting in the following set of mental diseases: *MDD (F32-F33), schizophrenia (F20), bipolar disorder (F31), anxiety related disorders (F40-F41), stress related disorders (F43), mental disorders related to alcohol abuse (F10).* And the following somatic diseases: *irritable bowel syndrome (K58), overweight and obesity (E66) and gastro-esophageal reflux (K21).*

Data sources

The study was conducted utilizing data from three public health cohorts, namely:

1) **The Catalan Health Surveillance System (CHSS)** - The main cohort used in MADS development was extracted from the CHSS. Operated by a single-public payer (CatSalut)[42] since 2011, the CHSS gathers information across healthcare tiers on the utilization of public healthcare resources, pharmacological prescriptions, and patients' basic demographic data, including registries of 7.5 million citizens from the entire region of Catalonia (ES). Nevertheless, for MADS development purposes we considered only registry data from the citizens resident in the entire Health District of Barcelona-Esquerra (AISBE) between the 1st of January 2011 and the 31st of December 2019 (n=654,913). To validate the results of MADS, we retrieved additional information from CHSS corresponding to the 12 months posterior to the MADS assessment, from the 1st of January 2020 to the 31st of December 2020. It is to note that, all the deceased patients in addition to those who moved their residence outside of AISBE district between 2011 and 2019 were discarded from the MADS assessment analysis; the remaining subset of patients comprises 508,990 individuals.
2) **The United Kingdom Biobank (UKB)** – The UKB data considered in this study contained medical and phenotypic data from participants aged between 37-93 years. Recruitment was based on NHS patient registers, and initial assessment visits were carried out between the 3rd of March 2006 and the 1st of October 2010 (n = 502,504). The analyzed data included disease diagnosis and onset time, medication prescriptions, and socio-economic descriptors.
3) **The Finnish National Institute for Health and Welfare biobank (THL)** - THL cohort integrates information from Finrisk[36] 1992, 1997, 2002, 2007, 2012, Finhealth[37] 2017 and Health[38] 2000/2011 studies. For the consensual clustering, 41,092 participants were used from Finnish population surveys. After data cleaning, 30,961 participants remained from Finnish population surveys. These participants aged 20-100 were chosen at random from the Finnish

population and represented different parts of Finland.

Demographic information on the study cohorts is displayed in the results section (**Table 1**).

Ethical approval

As a multicentric study, TRAJECTOME accessed data from multiple cohorts, all subject to the legal regulations of their respective regions of origin and obtained the necessary approvals from the corresponding ethics committees.

For the CHSS cohort, the Ethical Committee for Human Research at Hospital Clinic de Barcelona approved the core study of TRAJECTOME on the 24th of March 2021 (HCB/2020/1051) and subsequently approved the analysis for the generation and the assessment of MADS on the 25th of July 2022 (HCB/2022/0720).

UK Biobank received ethical approval from the National Research Ethics Service Committee Northwest–Haydock (ref. 11/NW/0382).

The THL cohort integrates information from the Finrisk databases: 1997 (Ethical committee of National Public Health Institute. Statement 38/96. 30.10.1996), 2002 (Helsinki University Hospital, Ethical committee of epidemiology and public health, Statement 87/2001. Reference 558/E3/2001. 19.12.2001), 2007 (Helsinki University Hospital, Coordinating ethics committee, Dnro HUS 229/EO/2006, 20.6.2006) and 2012 (Helsinki University Hospital, Coordinating ethics committee, Dnro HUS 162/13/03/11, 1.12.2011); the FinHealth 2017 (Helsinki University Hospital, Coordinating ethics committee, 37/13/03/00/2016 22.3.2016) and the Health 2000-2011 databases (Ethical committee of National Public Health Institute, 8/99/12. Helsinki University Hospital, Ethical committee of epidemiology and public health, 407/E3/2000. 31.05.2000 and 17.06.2011).

The ethics committees exempted the requirement to obtain informed consent for the analysis and publication of retrospectively acquired and fully anonymized data in the context of this non-interventional study.

All the data was handled in compliance with the General Data Protection Regulation 2016/679,

which safeguards data protection and privacy for all individuals in the European Union. The study was conducted in conformity with the Helsinki Declaration (Stronghold Version, Brazil, October 2013) and in accordance with the protocol and the relevant legal requirements (Biomedical Research Act 14/2007 of 3 July).

Statistical analysis

The results of the cross-sectional analysis of health outcomes and use of healthcare resources were evaluated through various metrics. Mortality rates were summarized as cases per 1,000 inhabitants. In contrast, numeric health outcomes variables were described by the average number of cases per person, per 100 inhabitants or 1,000 inhabitants according to their prevalence. Average healthcare expenditures are reported in € per person. Kruskal-Wallis, supplemented by Bonferroni-adjusted post-hoc right-tailed Dunn's tests, and pairwise Fisher's exact tests were employed to evaluate changes in the target outcomes across the risk pyramid tiers. Statistical significance was determined by considering a P-value less than .05 in all analyses.

The results of the longitudinal analysis on disease prevalence and the incidence of new disease onsets of MDD and nine mental and somatic MDD-related chronic conditions (PR >.80) were expressed in percentages (%) and in per thousand (‰), respectively.

All the data analyses were performed using R[43], version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria). The MADS algorithm was fully developed and tested in the CHSS database and transferred to the other sites through an R programming executable script.

The study is reported according to the STROBE[23] guidelines for observational studies.

RESULTS

Sociodemographic characteristics of the study cohorts

One of the first results is the characterization of the three study cohorts and compared the sociodemographic attributes of their MADS risk groups (**Table 1**). All the individuals were classified

into distinct risk strata based on quantiles of MADS distribution within the source population, resulting in the formation of the subsequent risk pyramid: Very low risk tier (≤ P_{50}); Low risk tier (P_{50}-P_{80}]; Moderate risk tier (P_{80}-P_{90}]; High risk tier (P_{90}-P_{95}]; Very high risk tier (> P_{99}).

It is imperative to underscore the fundamental distinctions in the cohorts under study to comprehend the inherent sociodemographic disparities across them. Specifically, the THL and UKB cohorts predominantly consist of data derived from biobanks, specifically focusing on the middle-aged and elderly population. In contrast, the CHSS cohort represents a population-based sample encompassing the entire population spectrum.

It is worth noting that a common pattern is observed among all the cohorts in the age distribution of the citizens at risk. Although MADS is an additive morbidity grouper, it is not monotonically increasing with age. Remarkably, a notable proportion of high-risk cases were observed within the age range of 40 to 60 years, when depression typically manifests for the first time on average.

A divergence in the sex distribution across the risk strata is observable and especially noticeable in CHSS and UKB cohorts, where the morbidity burden associated with depression and its related diseases is amplified in women (P<.001). Likewise, the disability caused by depression and its comorbidities is larger in families with fewer economic resources (P<.001). Overall, the prevalence of MDD is greater in UKB than in the other cohorts. However, upon analyzing the allocation of the population afflicted with depression in the risk pyramid, a total of 22,238 individuals (57.79% of those diagnosed with MDD) are categorized in the "high" and "very high" risk tiers in the CHSS cohort, whereas the number of individuals diagnosed with MDD that are allocated at the tip of the risk pyramid is 920 (40.22%) in THL and 23,409 (43.78%) in UKB.

Table 1 *– Demographic characteristics of each stratum of the MADS risk pyramid in the three study cohorts: CHSS[33], UKB[34] and THL[35].*

Risk Pyramid Tiers	**N**			**Age, mean (SD)**			**Sex, n (%)** **M = Male** **F = Female**			**Household Income, n (%)** **L = Low (< 18k €)** **M = Medium (18-100k €)** **H = High (> 100k €)**			**Major Depressive Disorder Prevalence, n (%)**		
	CHSS	**THL**	**UKB**	**CHSS**	**THL**	**UKB**	**CHSS**	**THL**	**UKB**	**CHSS**	**THL**	**UKB**	**CHSS**	**THL**	**UKB**
All cases	507,549	30,	502,	45.36	64.	61	M:	M:	M:	L:	L:	L:	38,	2,2	53,

		961	504	(23.07)	27 (14.28)	.48 (9.31)	237,598 (46.81) F: 269,951 (53.19)	14,435 (46.62) F: 16,526 (53.38)	229,122 (45.60) F: 273,382 (54.40)	262,753 (51.77) M: 223,369 (44.01) H: 21,427 (4.22)	11,489 (37.11) M: 10,025 (32.38) H: 9,447 (30.51)	117,737 (23.43) M: 358,492 (71.34) H: 26,275 (5.23)	479 (7.58)	87 (7.39)	466 (10.64)
P value	N.A.	N.A.	N.A.	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Very high risk $> P_{99}$	5,651	310	5,026	55.74 (18.83)	68.83 (14.86)	61.7 (8.75)	M: 2,322 (41.09) F: 3,329 (58.91)	M: 129 (41.61) F: 181 (58.39)	M: 2,207 (43.91) F: 2,819 (56.09)	L: 4,343 (76.85) M: 1,251 (22.14) H: 57 (1.01)	L: 191 (61.61) M: 77 (24.84) H: 42 (13.55)	L: 2,285 (45.46) M: 2,620 (52.13) H: 121 (2.41)	3,870 (68.48)	186 (60.00)	4,370 (86.95)
High risk ($P_{95} – P_{99}$]	22,894	1,238	20,084	60.08 (20)	65.12 (15.10)	63.2 (8.74)	M: 7,170 (31.32) F: 15,724 (68.68)	M: 559 (45.15) F: 679 (54.85)	M: 7,545 (37.57) F: 12,539 (62.43)	L: 14,568 (63.63) M: 7,946 (34.71) H: 380 (1.66)	L: 690 (55.74) M: 327 (26.42) H: 221 (17.85)	L: 7,626 (37.97) M: 12,003 (59.76) H: 455 (2.27)	18,368 (80.23)	734 (59.29)	19,039 (94.8)
Moderate risk ($P_{80} – P_{95}$]	84,371	4,644	75,378	54.56 (21.87)	68.86 (14.77)	63.6 (9.02)	M: 34,462 (40.85) F: 49,909 (59.15)	M: 2,201 (47.41) F: 2,441 (52.59)	M: 34,282 (45.48) F: 41,096 (54.52)	L: 49,818 (59.05) M: 32,822 (38.9) H: 1,731 (2.05)	L: 2,285 (49.22) M: 1,437 (30.96) H: 920 (19.82)	L: 23,208 (30.79) M: 49,684 (65.91) H: 2,486 (3.3)	16,241 (19.25)	1,367 (29.45)	25,776 (34.2)
Low risk ($P_{50} – P_{80}$]	162,170	9,266	150,759	47.66 (24.2)	66.16 (14.15)	62.2 (9.39)	M: 77,082 (47.53) F: 85,088 (52.47)	M: 4,132 (44.58) F: 5,137 (55.42)	M: 70,550 (46.80) F: 80,209 (53.20)	L: 85,936 (52.99) M: 71,429 (44.05) H: 4,805 (2.96)	L: 3,623 (39.09) M: 3,081 (33.24) H: 2,565 (27.67)	L: 36,773 (24.39) M: 106,441 (70.6) H: 7,545 (5)	0 (0)	0 (0.00)	2,002 (1.33)

Very low risk ≤ P_{50}	232,463	15,503	251,257	38.72 (20.72)	61.62 (13.55)	60.3 (9.22)	M: 116,562 (50.14) F: 115,901 (49.86)	M: 7,414 (47.83) F: 8,088 (52.17)	M: 114,538 (45.59) F: 136,719 (54.41)	L: 108,088 (46.5) M: 109,921 (47.29) H: 14,454 (6.22)	L: 4,700 (30.32) M: 5,103 (32.92) H: 5,699 (36.76)	L: 47,845 (19.04) M: 187,744 (74.72) H: 15,668 (6.24)	0 (0)	0 (0.00)	2,279 (0.91)

The prevalence of depression was calculated considering both F32 and F33 ICD-10-CM diagnostic codes. Kruskal-Wallis tests were used to assess changes in the target outcomes according to the risk pyramid tiers (Statistical significance = P-value < .05; H_0 = "all MADS risk groups have the same outcome distribution"; H_1 = "at least one MADS risk group has a different outcome distribution than the others"). Abbreviations: CHSS: Catalan Health Surveillance System cohort; THL: The Finnish National Institute for Health and Welfare biobank cohort; UKB: UK biobank cohort.

Assessment of the MADS risk groups

Assessment of the PRs

Analyzing the relationship between MDD and the morbidities assessed in the study is essential to interpreting the MADS risk strata. This analysis revealed various relevant connections between MDD, and the diseases investigated, encompassing both acute and chronic conditions, with the latter being particularly noteworthy due to their non-transient nature. Notably, the cluster of mental and behavioral disorders showed the highest average PRs in depression. However, relevant associations also emerge among MDD and specific chronic somatic diseases affecting multiple organic systems (**Figure 2**).

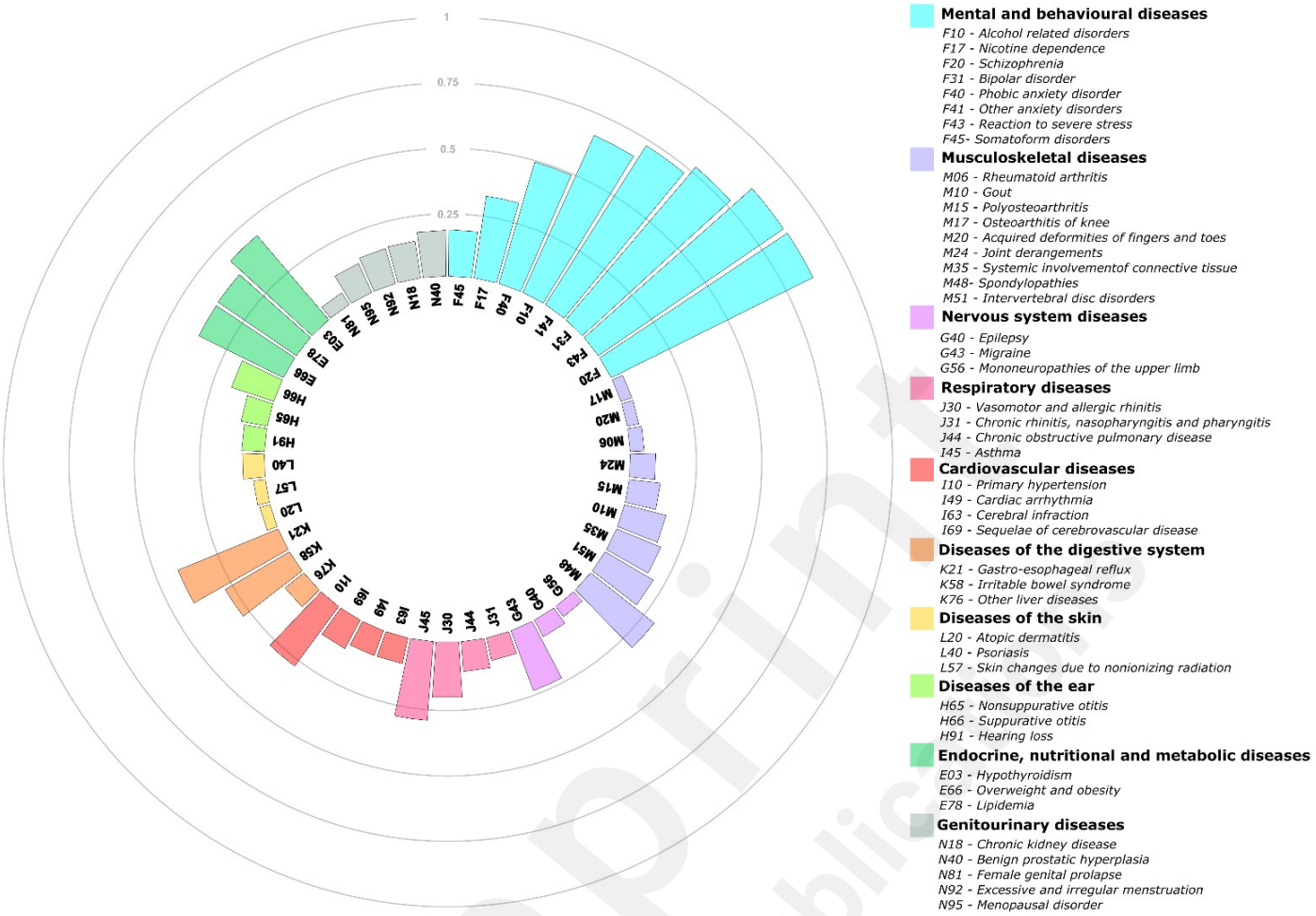


Figure 2 – Average probabilities of relevance between Major Depressive Disorder and 45 chronic conditions utilized to compute MADS.

Utilization of healthcare resources

The impact of MADS risk groups on healthcare systems was evaluated by investigating the correlation between the MADS risk categories and the utilization of health resources over the 12-month period following the MADS assessment within the CHSS cohort (**Table 2**). The results revealed significantly different distributions of the assessed outcomes across the MADS risk tiers, including primary care visits (P<.001), specialized outpatient visits (P<.001), emergency room visits (P<.001), hospital admissions (P<.001) and ambulatory visits in mental health centers (P<.001) as well as the pharmacological burden (P<.001). Furthermore, the results of the pairwise comparisons between adjacent risk tiers illustrate a significant and gradual pattern of increased healthcare utilization as individuals progress from lower MADS risk to higher risk tiers, reflecting an escalation in healthcare needs and requirements. Overall, patients with higher MADS scores exhibit a greater likelihood of experiencing morbidity-related adverse events, which subsequently leads to recurrent interactions with healthcare systems across multiple levels.

Table 2 *–Utilization of healthcare resources over 12 months in each stratum of the MADS risk pyramid for the CHSS cohort.*

Risk Pyramid Tiers	**Primary Care visits** (visits/ person)	**Specialized Outpatient visits** (visits/ person)	**Emergency Room visits** (visits/100 inhabitants)	**Hospital admissions** (admissions/100 inhabitants)	**Mental Health visits** (visits/ 100 inhabitants)	**Number of prescriptions** (prescriptions/ person)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Very high risk $> P_{99}$	12.50	3.07 ***	135.00 ***	28.50 ***	554.00 ***	8.02 ***
High risk $(P_{95} – P_{99}]$	11.90 ***	2.56 ***	87.20 ***	20.60 ***	136.00 ***	7.48 ***
Moderate risk $(P_{80} – P_{95}]$	9.03 ***	1.82 ***	61.90 ***	14.50 ***	44.20 ***	5.11 ***
Low risk $(P_{50} – P_{80}]$	6.21 ***	1.21 ***	42.40 ***	8.87 ***	15.10 ***	3.20 ***
Very low risk $\leq P_{50}$	2.96	0.50	23.40	3.25	5.96	1.07

*Kruskal-Wallis tests were used to assess changes in the target outcomes according to the risk pyramid tiers (P value). Subsequent pairwise comparisons between each risk tier and the next consecutive level of lesser risk were conducted using right-tailed Dunn's post hoc test. These specific outcome differences are denoted within the table using *, **, *** to represent p-values of < .05, < .01, < .001, respectively. (Statistical significance = P-value < .05)*

Mortality and healthcare expenditure

We performed a cross-sectional analysis investigating mortality rates and the healthcare expenditure within the 12 months following the MADS assessment, expressed as the average healthcare expenditure per capita and differentiating among pharmaceutical and non-pharmaceutical costs within the CHSS and THL cohorts (**Table 3**). Significant variations in mortality rates were observed across the risk pyramid tiers (P < .001), with rates in the high-risk strata being markedly elevated—ranging from 5 to 20 times depending upon the cohort—compared to low-risk individuals. Furthermore, the distribution of average healthcare expenditures per person was significantly different among the risk tiers, with both pharmacological and non-pharmacological expenses demonstrating disparities (P < .001). Pairwise comparisons further indicated that individuals at the highest risk tier incurred substantially greater healthcare costs than those at the lowest tier, reflecting a gradient of financial impact correlated with increased risk levels.

Table 3: *Mortality rates and pharmacological and non-pharmacological healthcare expenditure in €, over 12 months, in each stratum of the MADS risk pyramid in CHSS[33] and THL[35].*

Risk Pyramid Tiers	**Mortality** (cases/1k inhabitants)		**Pharmacological expenditure in €** (average expenditure per person)		**Hospitalization expenditure in €** (average expenditure per person)		**Total expenditure in €** (average expenditure per person)
	CHSS	**THL**	**CHSS**	**THL**	**CHSS**	**THL**	**CHSS**
P value	N.A.	N.A.	<.001	<.001	<.001	<.001	<.001
Very high risk $> P_{99}$	46.2 ***	36.0 ***	1,214 ***	966	539 ***	270	12,517 ***
High risk $(P_{95} – P_{99}]$	41.5 ***	33.7 ***	772 ***	1,131 ***	383 ***	340 ***	8,404 ***
Moderate risk $(P_{80} – P_{95}]$	25.5 ***	32.2 ***	485 ***	1,077 ***	270 ***	254 ***	5,209 ***
Low risk $(P_{50} – P_{80}]$	11.5 ***	14.8 ***	292 ***	810 ***	165 ***	185 ***	3,075 ***
Very low risk $\leq P_{50}$	2.57	7.3	99	363	60	123	1,192

*Kruskal-Wallis tests were used to assess changes in the target outcomes according to the risk pyramid tiers (P value). Subsequent pairwise comparisons between each risk tier and the next consecutive level of lesser risk were conducted using right-tailed Dunn's post hoc test. These specific outcome differences are denoted within the table using *, **, *** to represent p-values of < .05, < .01, < .001, respectively. Pairwise comparisons of Fisher exact tests were used to assess changes in mortality rates, outcome differences are denoted within the table using the same scheme. Statistical significance = P-value < .05. Abbreviations: CHSS: Catalan Health Surveillance System cohort; THL: The Finnish National Institute for Health and Welfare biobank cohort.*

Pharmacological burden

The study also examined the pharmacological burden on individuals after 12 months following the MADS assessment (**Table 4**). The data analysis reveals distinct patterns of medication utilization across the risk tiers, with significant differences in the use of antidepressants, antipsychotics, anxiolytics, and sedatives (all P<.001). This trend, consistently observed across the three cohorts, is further emphasized by pairwise comparisons between adjacent risk levels, which reveal a strong positive correlation between higher risk strata and increased pharmaceutical consumption. This upward trend in medication usage forms a clear gradient, demonstrating that individuals in progressively higher risk tiers face substantially greater pharmaceutical needs.

Table 4: *Prescription of depression related pharmacological treatments over 12 months in each stratum of the MADS risk pyramid in CHSS[33], UKB[34] and THL[35].*

Risk Pyramid Tiers	Antipsychotic (N05A) (prescriptions/person)			Anxiolytic (N05B) (prescriptions/person)			Hypnotics and sedatives (N05C) (prescriptions/person)			Antidepressant (N06A) (prescriptions/person)		
	CHSS	**THL**	**UKB**	**CHSS**	**THL**	**UKB**	**CHSS**	**THL**	**UKB**	**CHSS**	**THL**	**UKB**
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Very high risk $> P_{99}$	0.75 ***	0.60 ***	0.33 ***	0.47	0.21 ***	0.27 ***	0.15 ***	0.14 ***	0.24 ***	0.79 ***	0.43 ***	0.80 ***
High risk $(P_{95} – P_{99}]$	0.20 ***	0.27 ***	0.18 ***	0.46 ***	0.19 ***	0.20 ***	0.10 ***	0.12 ***	0.19 ***	0.66 ***	0.41 ***	0.71 ***
Moderate risk $(P_{80} – P_{95}]$	0.07 ***	0.08 ***	0.15 ***	0.28 ***	0.08 ***	0.16 ***	0.05 ***	0.10 ***	0.18 ***	0.27 ***	0.27 ***	0.54 ***
Low risk $(P_{50} – P_{80}]$	0.03 ***	0.03 ***	0.13 ***	0.14 ***	0.04 ***	0.12 ***	0.02 ***	0.07 ***	0.13 ***	0.08 ***	0.11 ***	0.36 ***
Very low risk $\leq P_{50}$	0.01	0.01	0.11	0.04	0.02	0.09	0.01	0.04	0.10	0.02	0.06	0.26

*For recurrently dispensed medication only the first prescription was considered in the analysis. Kruskal-Wallis tests were used to assess changes in the target outcomes according to the risk pyramid tiers (P value). Subsequent pairwise comparisons between each risk tier and the next consecutive level of lesser risk were conducted using right-tailed Dunn's post hoc test. These specific outcome differences are denoted within the table using *, **, *** to represent p-values of < .05, < .01, < .001, respectively. Statistical significance = P-value < .05. Abbreviations: CHSS: Catalan Health Surveillance System cohort; THL: The Finnish National Institute for Health and Welfare biobank cohort; UKB: UK biobank cohort.*

To evaluate the influence of age and sex on the outcomes examined in this section, we replicated all the previously presented results, categorizing the outcomes by sex and age and reported them in the **Supplementary material – Appendix 1**. The results suggest that the morbidity burden in individuals might be a primary driver influencing the occurrence of adverse health events and the heightened utilization of healthcare resources.

Multimorbidity progression

We analyzed the prevalence and incidence of new MDD-associated diagnoses and the relevant comorbid conditions in 5-year intervals after MADS assessment for depression throughout the patients' lifespan (**Table 5)**, allowing for a comprehensive examination of multimorbidity progression over time.

Table 5 displays the current disease prevalences (red) expressed in % and the incidence (blue) of new disease onsets along an interval of 5 years after MADS assessment expressed in ‰. **Table 5**

showcases the results for MDD, and nine mental and somatic MDD-related (PR >.80) chronic conditions, assessed independently in the three study cohorts, namely CHSS, THL and UKB; and in four time points, that is 20, 40 60 and 70 years old, corresponding to the intervals in which the PRs were recalculated. A continuous assessment of these outcomes is reported in the **Supplementary material – Appendix 2**.

In general, both MDD and the comorbid conditions investigated in this study exhibit a positive correlation between the MADS risk tiers and the current prevalence and incidence of new disease onsets within a subsequent 5-year interval. This is evident from the colored cells within the table. Notably, the highest disease prevalence and incidence values, consistently appear in the high and very high-risk tiers. Additionally, there is a discernible pattern of well-stratified values across these risk tiers within the same age ranges, underlining significantly elevated prevalence rates of the studied diseases compared to the population average within the high-risk groups. Age also emerges as a pivotal determinant influencing disease onset, delineating unique patterns across various disorders. Notably, conditions such as gastro-esophageal reflux and overweight consistently exhibit ascending trends in both incidence and prevalence throughout individuals' lifespans. Conversely, severe afflictions such as schizophrenia, bipolar disorder, and alcohol abuse reach their zenith in prevalence and incidence during middle-aged adulthood, followed by a decline, possibly indicating an association with premature mortality. Moreover, anxiety and stress-related disorders show their highest incidence rates during youth and early adulthood.

The consistency of the findings illustrated in **Table 5** remains robust across all three study cohorts despite their significant demographic differences described in **Table 1**. These heterogeneities result in disease prevalence discrepancies among cohorts, as vividly portrayed in **Table 5**. Among the most relevant cases, there exists an elevated prevalence of schizophrenia in the THL cohort in comparison with CHSS and UKB. In this particular case, schizophrenic patients integrate 100% of the very high-risk group in adulthood. Such differences in disease prevalence among cohorts may influence distinct health outcomes, particularly for the citizens allocated at the apex of the Finnish risk pyramid, as observed in the pharmacological and hospitalization expenditure outcomes reported in **Table 3**.

Table 5- Longitudinal analysis of disease prevalence (red) and incidence (blue) of new disease onsets in CHSS[33], UKB[34] and THL[35].

	CHSS	THL	UKB

Disease	Risk Pyramid Tiers	Age								Age								Age							
		Prevalence (%)				Incidence (‰)				Prevalence (%)				Incidence (‰)				Prevalence (%)				Incidence (‰)			
		20	40	60	70	20	40	60	70	20	40	60	70	20	40	60	70	20	40	60	70	20	40	60	70
Major Depressive Disorder (F32-33)	Very High Risk	24	68	67	73	29	36	31	26	17	43	45	44	20	44	58	15	78	91	89	89	29	37	50	67
	High Risk	0	21	99	97	15	58	133	48	0	26	88	71	8	51	37	8	0	68	98	95	9	65	49	33
	Moderate Risk	0	0	7	24	5	28	58	55	-	0	7	5	-	20	17	17	0	0	20	28	6	27	29	24
	Low Risk	-	0	0	0	-	7	33	40	-	-	0	0	-	-	8	7	-	0	0	0	-	15	19	17
	Very Low Risk	0	0	0	0	2	7	19	25	0	0	0	0	2	6	4	5	0	0	0	0	4	10	13	13
Schizophrenia (F20)	Very High Risk	7	36	40	33	7	8	3	3	8	72	100	97	0	13	-	0	3	11	14	15	2	2	1	3
	High Risk	0	0	0	0	3	5	1	2	0	0	12	9	0	10	8	5	0	0	0	0	0	1	1	1
	Moderate Risk	0	0	0	0	1	2	1	1	-	0	0	0	-	2	3	3	0	0	0	0	0	0	0	0
	Low Risk	-	0	0	0	-	1	1	0	-	-	0	0	-	-	1	1	-	0	0	0	-	0	0	0
	Very Low Risk	0	0	0	0	1	1	1	1	0	0	0	0	1	2	1	1	0	0	0	0	0	0	0	0
Bipolar Disorder (F31)	Very High Risk	2	5	8	7	5	19	13	8	2	11	18	16	3	12	31	20	0	2	5	7	1	5	9	17
	High Risk	1	1	3	2	2	5	7	2	0	3	5	2	0	6	11	4	0	0	1	1	0	2	3	3
	Moderate Risk	0	1	2	1	1	2	2	2	-	0	1	0	-	2	1	3	0	0	0	1	0	0	1	2
	Low Risk	-	0	0	0	-	1	1	1	-	-	0	0	-	-	1	0	-	0	0	0	-	0	0	0
	Very Low Risk	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Anxiety Related Disorders (F40-41)	Very High Risk	36	44	35	32	141	137	98	96	14	21	16	11	30	14	23	9	1	16	38	33	5	45	47	54
	High Risk	18	54	36	35	78	130	99	79	0	15	15	13	8	31	19	10	1	5	17	28	1	21	49	49
	Moderate Risk	0	15	47	31	47	80	77	67	-	0	5	3	-	5	6	1	0	2	11	10	1	9	21	27
	Low Risk	-	0	1	13	-	29	62	53	-	-	0	1	-	-	4	3	-	0	0	4	-	4	12	16
	Very Low Risk	0	0	0	0	10	21	29	38	0	0	0	0	2	2	2	2	0	0	0	0	1	3	7	13
Stress Related Disorders (F43)	Very High Risk	34	41	34	36	29	48	30	20	10	16	9	3	22	26	7	0	2	9	31	18	1	29	11	2
	High Risk	7	25	51	67	20	48	46	30	0	11	14	10	4	19	14	7	0	9	7	12	1	13	15	5
	Moderate Risk	0	3	17	17	9	27	30	26	-	0	3	3	-	5	4	3	0	0	9	8	0	9	9	2
	Low Risk	-	0	0	2	-	10	24	19	-	-	0	0	-	-	1	2	-	0	0	1	-	4	6	2
	Very Low Risk	0	0	0	0	2	8	13	14	0	0	0	0	2	2	1	2	0	0	0	0	0	2	3	1
Mental Disorders Related to Alcohol Abuse (F10)	Very High Risk	7	12	22	13	26	38	20	15	8	21	23	16	4	14	33	0	4	6	12	14	1	16	14	11
	High Risk	3	15	5	7	12	18	18	11	0	27	23	17	2	30	12	5	0	4	4	6	0	5	11	11
	Moderate Risk	0	0	10	8	5	12	12	10	-	0	16	17	-	10	9	6	0	0	3	6	0	3	9	8
	Low Risk	-	0	0	1	-	2	11	7	-	-	0	0	-	-	5	5	-	0	0	1	-	1	5	7
	Very Low Risk	0	0	0	0	1	3	4	4	0	0	0	0	1	4	6	3	0	0	0	0	0	1	4	6
Irritable Bowel Syndrome (K58)	Very High Risk	1	2	4	4	6	7	6	4	0	1	3	2	3	0	0	0	2	11	16	14	12	34	13	11
	High Risk	0	2	4	6	3	5	8	7	1	1	3	5	0	8	5	7	2	7	16	14	5	23	22	13
	Moderate Risk	2	2	2	3	2	4	6	6	-	2	1	3	-	3	5	3	1	7	8	9	3	16	13	9
	Low Risk	-	1	2	2	-	3	4	6	-	-	2	2	-	-	5	6	-	6	8	7	-	10	13	8
	Very Low Ris	0	0	0	1	1	1	4	4	0	0	0	0	1	1	2	2	0	0	2	4	2	6	8	6

	k																							
Overweight and Obesity (E66)	Very High Risk	10	16	26	34	21	70	91	89	2	3	6	4	3	22	7	26	0	3	12	16	0	22	41	35
	High Risk	5	10	22	30	12	44	80	84	3	2	7	5	2	10	21	7	0	1	9	13	0	8	43	41
	Moderate Risk	14	9	18	31	8	41	73	83	-	3	4	5	-	6	8	11	0	2	7	17	0	9	31	37
	Low Risk	-	3	19	20	-	15	73	82	-	-	3	3	-	-	8	6	-	1	5	8	-	4	26	36
	Very Low Risk	0	0	2	2	2	10	36	48	0	0	0	0	0	1	3	2	0	0	1	1	0	3	14	25
Gastro Oesophageal Reflux (K21)	Very High Risk	1	2	5	6	7	15	23	26	1	3	5	6	7	11	7	36	1	8	20	25	2	37	39	60
	High Risk	1	2	4	5	4	14	29	31	0	4	8	9	0	14	31	21	1	2	17	24	2	22	65	61
	Moderate Risk	0	1	3	5	4	9	25	27	-	2	6	8	-	11	17	15	0	5	12	20	1	18	42	47
	Low Risk	-	0	3	3	-	4	19	23	-	-	4	8	-	-	20	17	-	0	9	17	-	9	47	47
	Very Low Risk	0	0	0	1	0	2	8	14	0	0	1	1	0	3	7	9	0	0	4	10	1	6	33	41

Prevalences are expressed as % and incidences are calculated considering a 5-year period after MADS assessment and expressed as ‰. The highest prevalence and incidence values per disease and cohort appear in the colored cells. The analysis considered only the chronic disease conditions with a PR against MDD ≥ 0.80 in at least one of the four age intervals assessed, namely: 0-20, 0-40, 0-60 and 0-70: MDD (F32-F33), schizophrenia (F20), bipolar disorder (F31), anxiety related disorders (F40-F41), stress related disorders (F43), mental disorders related to alcohol abuse (F10), irritable bowel syndrome (K58), overweight and obesity (E66) and gastro-esophageal reflux (K21).

DISCUSSION

Main findings

MADS seems to provide a novel and more comprehensive understanding of the complex nature of depression-related multimorbidity. This approach recognizes that individuals with depression often experience a range of comorbid conditions that may manifest and evolve differently over time. By capturing this dynamic aspect, MADS offers a nuanced assessment beyond a mere checklist of discrete disorders. The novelty of the MADS approach lies in its capability to serve as the first morbidity grouper that incorporates information on disease trajectories while improving the filtering of indirect disease associations using BDMMs.

In addition to capturing disease-disease associations, MADS endeavors to gauge their impact within the system by leveraging well-established DWs. However, despite achieving success in fulfilling the study's objectives, it is crucial to acknowledge that this approach carries inherent limitations, as will be elaborated upon in the subsequent section of the discussion.

In the current investigation, we have unearthed robust correlations between the MADS risk strata and the extent of deleterious impact caused by MDD and its comorbid conditions. Such associations indicate the presence of specific health risks and an escalated utilization of healthcare resources. Furthermore, a positive association has emerged between the levels of pharmacological and non-pharmacological healthcare expenditures and the different tiers of MADS risk. Also, the analysis has revealed an augmented risk of disease progression within the high-risk groups (high and very high-risk), as indicated by a heightened incidence of new-onset depression-related illnesses within a 12-month period after MADS assessment. Similarly, mortality rates have exhibited elevated values in these high-risk groups.

The findings presented in this study are underpinned by the complementary studies conducted within the TRAJECTOME project[30] that have established a better understanding of the complex multimorbidity landscape associated with MDD across an individual's lifespan, encompassing modifiable and genetic risk factors.

Limitations of the current approach

Despite meeting expectations and validating the hypothesis by which the study was conceived, the authors acknowledge a series of limitations leading to suboptimal results and limited potential for adaptation and generalization that should be undertaken to bring MADS, or an indicator derived from it to short-term real-world implementation.

In the current research, using of estimations of mean DW[44] to assess the burden of disease conditions has achieved desirable results, and is conceptually justified, but it undoubtedly exhibits significant limitations. In an ideal clinical scenario, each disease diagnosis indicated in the patient’s electronic medical record should be characterised by three key dimensions: i) severity of the diagnosis, ii) rate of disease progression, and iii) impact on disability. However, the degree of maturity for characterizing the last two dimensions, disease progression and disability, is rather poor because of the complexities involved in their assessment. In other words, the authors acknowledge the weakness associated with the current use of DW. However, they stress the importance of incorporating such dimensions in future evolutions of MADS.

A noteworthy aspect that should be acknowledged is that factors such as the advancements of

diagnostic techniques, the digitization of medical records, and the modifications in disease taxonomy and classification over time have contributed to a more exhaustive documentation of the disease states in the most recent health records. Consequently, this fact could lead to imprecisions in estimating the disease onset ages in older individuals.

Insights and potential impact of MADS in multimorbidity management.

The results reported in this study not only reaffirm the well-established link between multimorbidity and adverse outcomes such as a decline in functional status, compromised quality of life, and increased mortality rates[45] but also shed light on the significant burden imposed on individuals and healthcare systems. From the population-based HRA perspective, the strain on resource allocation and overall healthcare spending is a pressing concern that necessitates effective strategies for addressing and managing multimorbidity[46]. In this context, assessing individual health risks and patient stratification emerges as crucial approaches that enable the implementation of predictive and preventive measures in healthcare.

While population-based HRA tools like ACGs, CRGs, or AMGs have traditionally addressed this aspect, MADS is designed to complement rather than replace those tools. The current study aims to test a method to refine existing HRA tools by aligning them with the principles of network medicine, thereby merging traditional HRA with the practical application of network medicine insights. This innovative approach holds the promise of unlocking new potential advantages and capabilities.

The strength of the MADS approach lies in utilizing of disease-disease associations drawn from the analysis of temporal occurrence patterns among concurrent diseases. This virtue allows MADS to refine the analysis of the morbidity burden by focusing on clusters of correlated diseases, which, in turn, can aid in developing more tailored epidemiological risk-related studies. This refined analysis might also assist resource allocation and inform healthcare policies for targeted patient groups with specific needs. Moreover, this approach holds promise for potential extrapolation to other non-communicable disease clusters like diabetes, cardiovascular ailments, respiratory diseases, or cancer clinics. By leveraging this targeted approach, MADS can be adapted to other disease clusters with shared characteristics, enabling a more precise assessment of disease burden and comorbidity patterns and thereby generating multiple disease-specific indices.

Notably, when considering information derived from disease co-occurrence patterns, the presence or

absence of certain diseases seems to correlate with the risk of developing related comorbid conditions, as elucidated in **Table 5.** This highlights the potential for a nuanced understanding of disease relationships and their impacts on health outcomes and to implement preventive interventions to mitigate their effect. Moreover, the findings of this study highlight the potential of preventive strategies targeted at mental disorders, including substance abuse disorders, depressive disorders, and schizophrenia, to reduce the incidence of negative clinical outcomes in somatic health conditions. These important implications for clinical practice call for a comprehensive and interdisciplinary approach that bridges the gap between psychiatric and somatic medicine. By developing cross-specialty preventive strategies, healthcare professionals can provide more holistic and effective care for individuals with complex health needs, ensuring that their mental and physical health are adequately addressed[47].

The current study provided good prospects using of the disease trajectories to enhance the performance of existing state-of-the-art morbidity groupers, such as AMG. Recognized for its transferability across EU regions by the EU joint Action on implementation of digitally enabled integrated person-centered care (JADECARE)[48], AMG stands out due to its stratification capabilities, adaptability, and distribution as open-source software, providing several advantages over its commercial counterparts. The AMG system employs disease-specific weighting derived from statistical analysis incorporating mortality and healthcare service utilization data. This method addresses the primary drawback identified in the MADS approach inherent to the utilization of DW, while enabling the development of adaptable tools that align with the unique characteristics of each healthcare system. Consequently, it allows for the adjustment to the impact of specific disease conditions within distinct regions and enhances the overall applicability and adaptability of the tool. In this regard, this study has offered promising insights aligned with the developers' envisioned future features for integration into the AMG system. Serving as a proof of concept, it highlighted the potential improvements achievable within AMG by leveraging disease-disease associations, thereby shaping the roadmap for further AMG development.

MADS integration in precision medicine: advancing towards patient-centric strategies.

By assessing whether MADS is appropriate for the stratification of depression-related multimorbidity, we attempted to confirm its potential for contributing to precision medicine[49]. In the clinical arena, identifying individuals at elevated risk and customizing interventions enable healthcare providers to intervene proactively, potentially preventing, or lessening disease progression

and enhancing patient outcomes. These strategies not only yield immediate value in terms of improved patient care but also lay the foundation for the broader adoption of integrated care and precision medicine, particularly in the management of chronic conditions[50].

Incorporating systems medicine[51] methodologies and information technologies has prompted significant shifts in clinical research and practice, paving the way for holistic approaches, computational modelling, and predictive tools in clinical medicine. These advancements are driving the adoption of Clinical Decision Support Systems (CDSS), which use patient-specific data to generate assessments or recommendations, aiding clinicians in making informed decisions. It is well established that to improve predictive precision and aid clinical decision-making, implementing comprehensive methodologies that consider various influencing factors from multiple sources in patient health could enhance individual prognosis estimations[52].

This integration might facilitate predictive modelling methodologies for personalized risk prediction and intervention planning. This approach, known as Multisource Clinical Predictive Modelling (MCPM)[53,54] enables the integration of i) healthcare data and health determinants from other domains including: ii) Population health registry data; iii) Informal care data, including patients' self-tracking data, lifestyles, environmental, behavioral aspects, and sensors; and ideally iv) Biomedical research omics data. In this paradigm, it is crucial to acknowledge the pivotal role that multimorbidity groupers play in capturing the clinical complexity of individuals. Prior research[53,54] highlighted the synergy between patient clinical complexity (e.g., AMGs) and acute episode severity, correlating with higher risks of adverse health events. This opens avenues for further research, exploring how adjusted morbidity indicators like MADS can significantly contribute to predictive modelling, aiming at supporting the implementation of cost-effective, patient-centered preventive measures to manage chronic patients and potentially delay or prevent their progression to the highest risk levels in the stratification pyramid[55].

CONCLUSIONS

MADS showed to be a promising approach to estimate multimorbidity-adjusted risk of disease progression and measure MDD's impact on individuals and healthcare systems, which could be tested in other diseases. The novelty of the MADS approach lies in its unique capability to incorporate disease trajectories, providing a comprehensive understanding of depression-related

morbidity burden. In this regard, the BDMM method played a crucial role in isolating and identifying true direct disease associations. The results of the current study pave the way for the development of innovative digital tools to support advanced health risk assessment strategies. Nevertheless, clinical validation is imperative before considering the widespread adoption of MADS.

ACKNOWLEDGMENTS

The initiative was supported by ERA PerMed program (TRAJECTOME project, ERAPERMED2019-108). Locally, this study was supported by the Academy of Finland under the frame of ERA PerMed and the Hungarian National Research, Development, and Innovation Office 2019-2.1.7-ERA-NET-2020-00005 under the frame of ERA PerMed the Hungarian National Research, Development, and Innovation Office (K 143391, K 139330 and PD 134449 grants); the Hungarian Brain Research Program 3.0 (NAP2022-I-4/2022); and the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, under the TKP2021-EGA funding scheme (TKP2021-EGA-25 and TKP2021-EGA-02). Supported by the European Union project RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory.

We want to acknowledge, the earnest collaboration of the Digitalization for the Sustainability of the Healthcare (DS3) - IDIBELL group, for their support in the preparation of the Catalan cohort; that was extracted from the Catalan Health Surveillance System database, owned, and managed by the Catalan Health Service. Also, we want to acknowledge the participants and investigators of the FinnGen study and CSC – IT Center for Science, Finland, for computational resources.

This research has been conducted using the UK Biobank Resource under Application Number 1602.

Authors' Contributions

PA, GJ, and IC designed the study and directed the project. RGC, KM an IC led the design of MADS. RGC, KM, AG and TP executed the quantitative analysis; processed the experimental data, performed the statistical analysis and created the figures. EV generated the CHSS database and

provided statistical support. ZG, GH, HM, TN, MK, JPJ and JR provided insightful information to the study. The manuscript was first drafted by RGC, IC and JR and thoroughly revised by KM, EV, AG, TP, ZG, GH, HM, TN, MK, JPJ, PA, and GJ. All authors approved the final version of the manuscript and are accountable for all aspects of the work in ensuring its accuracy and integrity.

Competing interests

All authors declare no financial or non-financial competing interests.

Data availability

Data for this study are not publicly available due to patient privacy concerns.

Code availability

The scripts used to compute MADS are available from the corresponding author upon reasonable request.

ACRONYMS

AISBE- Health District of Barcelona-Esquerra
AMG – Adjusted Morbidity Groups
BDMM – Bayesian Direct Morbidity Maps
CHSS – Catalan Health Surveillance System
DW – Disability Weights
GBD - Global Burden of Disease
MADS – Multimorbidity Adjusted Disability Score
MCPM – Multisource Clinical Predictive Modelling
MDD – Major Depressive Disorder
PR – Probability of Relevance
THL - Finnish National Institute for Health and Welfare Biobank
UKB – United Kingdom Biobank

REFERENCES

1. VETRANO DL ET AL. An International Perspective on Chronic Multimorbidity: Approaching the Elephant in the Room. *J Gerontol A Biol Sci Med Sci*. 2018;73(10):1350-1356. doi:10.1093/gerona/glx178
2. SALIVE ME. Multimorbidity in Older Adults. *Epidemiol Rev*. 2013;35(1):75-83. doi:10.1093/epirev/mxs009
3. GARIN N ET AL. Impact of Multimorbidity on Disability and Quality of Life in the Spanish Older Population. *PLoS One*. 2014;9(11):e111498. doi:10.1371/JOURNAL.PONE.0111498
4. MONTERDE D ET AL. Multimorbidity as a predictor of health service utilization in primary care: a registry-based study of the Catalan population. 2020;21(1).
5. VALDERAS JM ET AL. Defining comorbidity: implications for understanding health and health services. *Ann Fam Med*. 2009;7(4):357-363. doi:10.1370/AFM.983
6. RAHMAN MH ET AL. Bioinformatics and system biology approaches to identify pathophysiological impact of COVID-19 to the progression and severity of neurological diseases. *Comput Biol Med*. 2021;138:104859. doi:10.1016/J.COMPBIOMED.2021.104859
7. TARASCHI A ET AL. Human Immune System Diseasome Networks and Female Oviductal Microenvironment: New Horizons to be Discovered. *Front Genet*. 2022;12:2752. doi:10.3389/FGENE.2021.795123/BIBTEX
8. CHAUHAN PK ET AL. Integrative network analysis interweaves the missing links in cardiomyopathy diseasome. *Sci Reports 2022 121*. 2022;12(1):1-11. doi:10.1038/s41598-022-24246-x
9. CALDERÓN-LARRAÑAGA A ET AL. Multimorbidity and functional impairment-bidirectional interplay, synergistic effects and common pathways. *J Intern Med*. 2019;285(3):255-271. doi:10.1111/JOIM.12843
10. ZHOU X ET AL. Human symptoms-disease network. *Nat Commun*. 2014;5. doi:10.1038/NCOMMS5212
11. MENCHE J ET AL. Uncovering disease-disease relationships through the incomplete human interactome. *Science*. 2015;347(6224):1257601. doi:10.1126/SCIENCE.1257601
12. GOH K IL ET AL. Exploring the human diseasome: the human disease network. *Brief Funct Genomics*. 2012;11(6):533-542. doi:10.1093/BFGP/ELS032
13. BARABÁSI AL. Network Medicine — From Obesity to the "Diseasome." *N Engl J Med*. 2007;357(4):404-407. doi:10.1056/NEJME078114

14. MURRAY SA ET AL. Illness trajectories and palliative care. *BMJ*. 2005;330(7498):1007-1011. doi:10.1136/BMJ.330.7498.1007
15. JENSEN AB ET AL. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nat Commun 2014 51*. 2014;5(1):1-10. doi:10.1038/NCOMMS5022
16. TINETTI ME ET AL. Designing Health Care for the Most Common Chronic Condition—Multimorbidity. *JAMA*. 2012;307(23):2493. doi:10.1001/JAMA.2012.5265
17. ELIXHAUSER A ET AL. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27. doi:10.1097/00005650-199801000-00004
18. PARKERSON GR ET AL. The Duke Severity of Illness Checklist (DUSOI) for measurement of severity and comorbidity. *J Clin Epidemiol*. 1993;46(4):379-393. doi:10.1016/0895-4356(93)90153-R
19. CHARLSON ME ET AL. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383. doi:10.1016/0021-9681(87)90171-8
20. STARFIELD B ET AL. Ambulatory care groups: a categorization of diagnoses for research and management. *Health Serv Res*. 1991;26(1):53.
21. MONTERDE D ET AL. Adjusted morbidity groups: A new multiple morbidity measurement of use in Primary Care. *Aten primaria*. 2016;48(10):674-682. doi:10.1016/J.APRIM.2016.06.003
22. STEFFEN A ET AL. Mental and somatic comorbidity of depression: A comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. *BMC Psychiatry*. 2020;20(1):1-15. doi:10.1186/S12888-020-02546-8/FIGURES/4
23. CEREZO-CEREZO J ET AL. *GOOD PRACTICE BRIEF - Population Stratification: A Fundamental Instrument Used for Population Health Management in Spain*.; 2018.
24. JOHNS HOPKINS ACG® SYSTEM. Accessed August 29, 2023. https://www.hopkinsacg.org/
25. 3M™ CLINICAL RISK GROUPS (CRGS) | 3M. Accessed August 29, 2023. https://www.3m.com/3M/en_US/health-information-systems-us/drive-value-based-care/patient-classification-methodologies/crgs/
26. UHLIG K ET AL. A framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med*. 2014;29(4):670-679. doi:10.1007/S11606-013-2659-Y/FIGURES/1
27. MARX P ET AL. Comorbidities in the diseasome are more apparent than real: What Bayesian filtering reveals about the comorbidities of depression. *PLoS Comput Biol*. 2017;13(6).

doi:10.1371/journal.pcbi.1005487

28. BOLGÁR B ET AL. VB-MK-LMF: Fusion of drugs, targets and interactions using variational Bayesian multiple kernel logistic matrix factorization. *BMC Bioinformatics*. 2017;18(1):1-18. doi:10.1186/s12859-017-1845-z
29. ICD - ICD-10-CM - INTERNATIONAL CLASSIFICATION OF DISEASES, TENTH REVISION, CLINICAL MODIFICATION. Accessed July 21, 2021. https://www.cdc.gov/nchs/icd/icd10cm.htm
30. JUHASZ G ET AL. Unique genetic and risk-factor profiles in multimorbidity clusters of depression-related disease trajectories from a study of 1.2 million subjects. *Nat Commun Prepr available Res Sq*. Published online August 2, 2023. doi:10.21203/RS.3.RS-3199113/V1
31. TRAJECTOME (2020-2023). Temporal disease map-based stratification of depression-related multimorbidities: towards quantitative investigations of patient trajectories and predictions of multi-target drug candidates. Published 2020. https://semmelweis.hu/trajectome/en/
32. ABBAFATI C ET AL. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9
33. FARRÉ N ET AL. Medical resource use and expenditure in patients with chronic heart failure: a population-based analysis of 88 195 patients. *Eur J Heart Fail*. 2016;18(9):1132-1140. doi:10.1002/ejhf.549
34. SUDLOW C ET AL. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med*. 2015;12(3):1001779. doi:10.1371/journal.pmed.1001779
35. THL BIOBANK - THL. Accessed April 19, 2023. https://thl.fi/en/web/thl-biobank
36. BORODULIN K ET AL. Cohort profile: The national finRiSK study. *Int J Epidemiol*. 2018;47(3):696-696I. doi:10.1093/ije/dyx239
37. VALSTA L ET AL. *FinHealth 2017 Study : Methods*. THL; 2019.
38. HEISTARO S. *Methodology Report : Health 2000 Survey*. Kansanterveyslaitos; 2008.
39. HILDERINK HBM ET AL. Accounting for multimorbidity can affect the estimation of the Burden of Disease: a comparison of approaches. Published online 2016. doi:10.1186/s13690-016-0147-7
40. HAAGSMA JA ET AL. *The Effect of Comorbidity on Health-Related Quality of Life for Injury Patients in the First Year Following Injury: Comparison of Three Comorbidity Adjustment Approaches*.; 2011. doi:10.1186/1478-7954-9-10
41. WORLD HEALTH ORGANIZATION(WHO). Guidelines for ATC classification and DDD

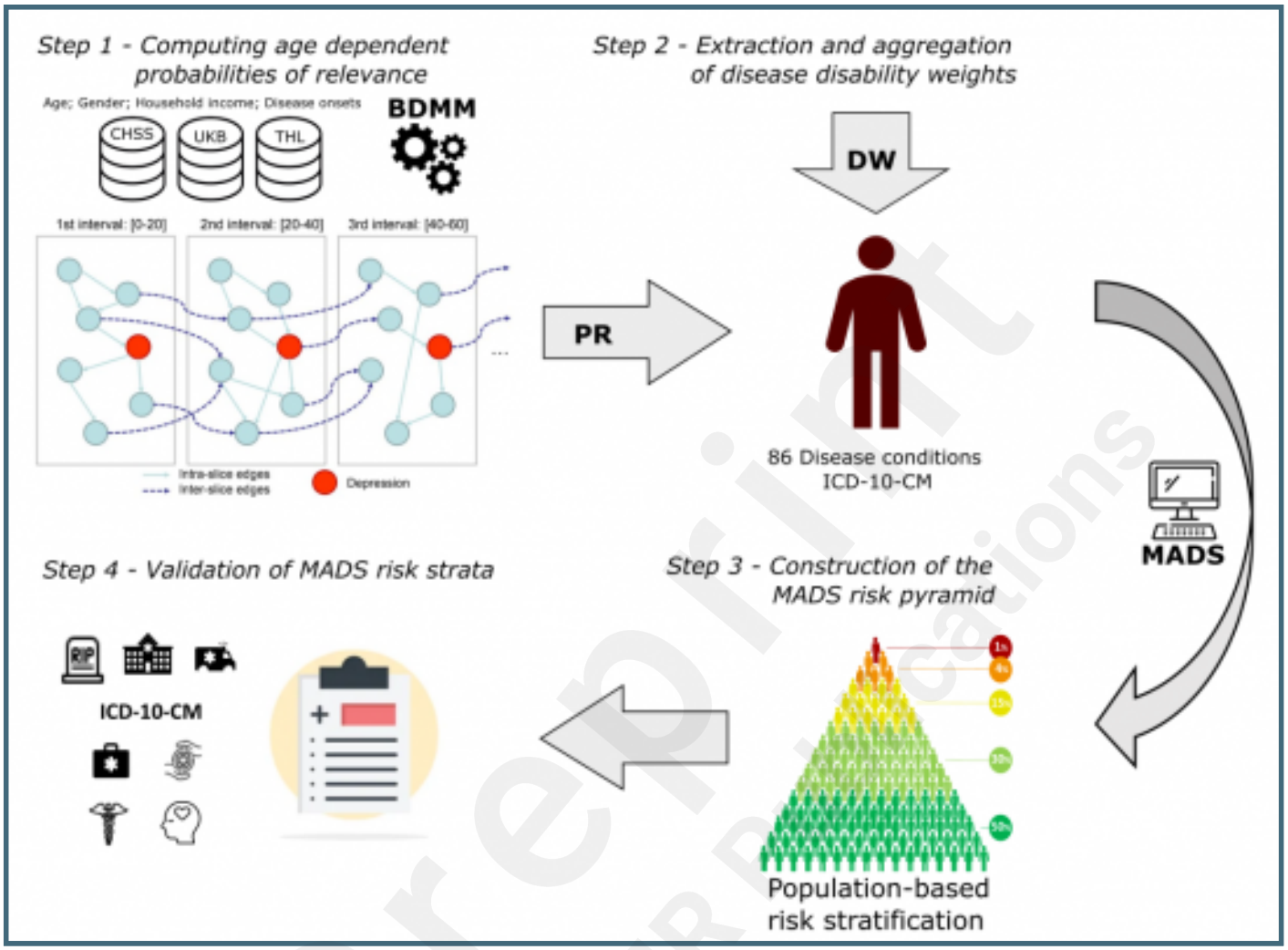
assignment. *Cent Drug Stat Methodol*. Published online 2021.

42. CATALAN HEALTH SERVICE. Accessed January 21, 2023. https://catsalut.gencat.cat/ca/inici/
43. R CORE TEAM. R: A language and environment for statistical computing. Published online 2021.
44. HAAGSMA JA ET AL. Review of disability weight studies: Comparison of methodological choices and values. *Popul Health Metr*. 2014;12(1). doi:10.1186/s12963-014-0020-2
45. MAKOVSKI TT ET AL. Multimorbidity and quality of life: Systematic literature review and meta-analysis. *Ageing Res Rev*. 2019;53. doi:10.1016/J.ARR.2019.04.005
46. YAMANASHI H ET AL. The role of mental disease on the association between multimorbidity and medical expenditure. *Fam Pract*. 2020;37(4):453-458. doi:10.1093/FAMPRA/CMAA015
47. DRAGIOTI E ET AL. Impact of mental disorders on clinical outcomes of physical diseases: an umbrella review assessing population attributable fraction and generalized impact fraction. *World Psychiatry*. 2023;22(1):86-104. doi:10.1002/WPS.21068
48. JADECARE (2020-2023). Joint Action on implementation of digitally enabled integrated person-centred care. Published 2020. https://www.jadecare.eu/
49. MUSKER M ET AL. Treating Depression in the Era of Precision Medicine: Challenges and Perspectives. *Neurobiol Depress Road to Nov Ther*. Published online January 1, 2019:265-275. doi:10.1016/B978-0-12-813333-0.00023-8
50. DUEÑAS-ESPIN I ET AL. Proposals for enhanced health risk assessment and stratification in an integrated care scenario. *BMJ Open*. 2016;6(4):e010301.
51. FEDEROFF HJ ET AL. Evolving from reductionism to holism: Is there a future for systems medicine? *JAMA - J Am Med Assoc*. 2009;302(9):994-996. doi:10.1001/jama.2009.1264
52. CANO I ET AL. Perspectives on Big Data applications of health information. *Curr Opin Syst Biol*. 2017;3:36-42. doi:10.1016/j.coisb.2017.04.012
53. CALVO M ET AL. Health outcomes from home hospitalization: Multisource predictive modeling. *J Med Internet Res*. 2020;22(10):e21367. doi:10.2196/21367
54. GONZÁLEZ-COLOM R ET AL. Computational modelling for prevention of unplanned hospital admissions in multimorbid patients. *J Med Internet Res*. Published online 2023. doi:10.2196/40846
55. BALTAXE E ET AL. Population-based analysis of COPD patients in Catalonia: implications for case management. *Eur Respir J*. 2017;50(suppl 61):PA4956. doi:10.1183/1393003.CONGRESS-2017.PA4956

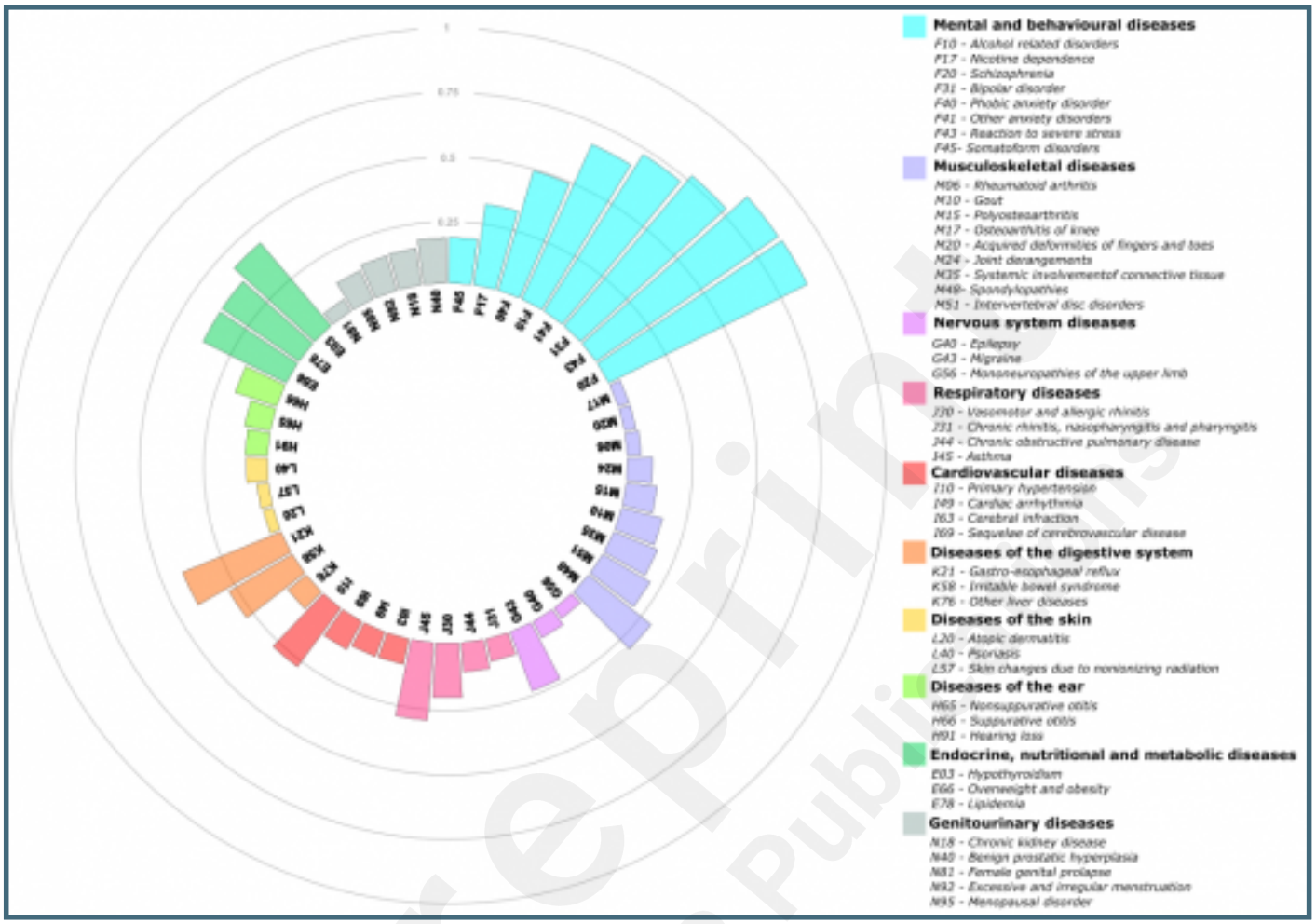
Supplementary Files

Figures

Workflow for building and validation of the MADS. BDMM stands for Bayesian Direct Multimorbidity Maps, PR for Probabilities of Relevance, and DW for Disability Weights.



Average probabilities of relevance between Major Depressive Disorder and 45 chronic conditions utilized to compute MADS.



Multimedia Appendixes

Supplementary material.
URL: http://asset.jmir.pub/assets/b17e0674d26e5d6682abe2304c7c8043.docx