European Neuropsychopharmacology 96 (2025) 71–83

Contents lists available at ScienceDirect

European Neuropsychopharmacology

journal homepage: www.sciencedirect.com/journal/european-neuropsychopharmacology

Pharmacological profiling of major depressive disorder-related multimorbidity clusters

Tamas Nagy [a,b,c,d], Xenia Gonda [a,b,c,e,f,*], Andras Gezsi [d], Nora Eszlari [a,b,c], Gabor Hullam [c,d], Rubèn González-Colom [g], Hannu Mäkinen [h], Teemu Paajanen [h], Dora Torok [a,b,c], Zsofia Gal [a,b,c], Peter Petschner [a,b,c], Isaac Cano [g,n], Mikko Kuokkanen [h,i,j], Carsten O. Schmidt [k], Sandra Van der Auwera [l,m], Josep Roca [g], Peter Antal [d], Gabriella Juhasz [a,b,c]

[a] Department of Pharmacodynamics, Faculty of Pharmaceutical Sciences, Semmelweis University, Nagyvárad tér 4., H-1089 Budapest, Hungary
[b] Center of Pharmacology and Drug Research & Development, Semmelweis University, Budapest, Hungary
[c] NAP3.0-SE Neuropsychopharmacology Research Group, Hungarian Brain Research Program, Semmelweis University, Üllői út 26., H-1085 Budapest, Hungary
[d] Department of Artificial Intelligence and Systems Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[e] Department of Psychiatry and Psychotherapy, Semmelweis University, Gyulai Pál utca 2., H-1085 Budapest, Hungary
[f] Department of Clinical Psychology, Semmelweis University, Üllői út 25, H-1092 Budapest, Hungary
[g] Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi i Sunyer (FRCB-IDIBAPS), Barcelona, Spain
[h] Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Mannerheimintie 166, 00300 Helsinki, Finland
[i] Department of Human Genetics and South Texas Diabetes and Obesity Institute, School of Medicine at University of Texas Rio Grande Valley, Brownsville, TX, United States
[j] Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Finland
[k] Institute for Community Medicine, University Medicine Greifswald, 17475 Greifswald, Germany
[l] Department of Psychiatry and Psychotherapy, University Medicine Greifswald, 17475 Greifswald, Germany
[m] German Centre for Neurodegenerative Diseases (DZNE), Site Rostock/Greifswald, 17475 Greifswald, Germany
[n] Universitat de Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Multimorbidity
Pharmacology
Major depressive disorder
Antidepressants

ABSTRACT

We previously identified seven distinct multimorbidity clusters associated with major depressive disorder through a comprehensive analysis of 1.2 million individuals of multiple cohorts. These clusters, characterized by unique clinical, genetic, and psychiatric and somatic illness risk profiles, implicate divergent treatment pathways and disease management strategies. This study aims to deepen the understanding of these clusters by analyzing drug prescriptions, evaluating the effectiveness of antidepressant treatment strategies, and identifying potential markers for personalized medicine.

Utilizing drug prescription data in the format of ATC codes, we performed epidemiological assessments, including multimorbidity (number of diseases), polypharmacy (number of chemical substances), and drug burden (number of prescriptions) analyses across the clusters. We applied linear regression models to assess strength and predictive capability of cluster membership on various metrics, and logistic regression to explore associations with treatment-resistant depression. We also quantified and visualized common antidepressant treatment sequences within each cluster.

Our findings indicate significant variations in polypharmacy and drug burden across clusters, with distinct patterns emerging that correlate with the clusters' profiles. Clusters liable to multimorbidity have higher drug burden, even after correction for number of diseases. Furthermore, the three clusters with higher risk for MDD showed different antidepressant treatment profiles; two required significantly more antidepressant prescriptions and had a higher risk for TRD.

The detailed pharmacological profiling presented in this study not only corroborates the initial cluster definitions but also enhances our predictive capabilities for treatment outcomes in MDD. By linking pharmacological data with comorbidity profiles, we pave the way for targeted therapeutic interventions.

* Correspondence author at: Gyulai Pal utca 2., Budapest, H-1085 Hungary.
E-mail address: gonda.xenia@semmelweis.hu (X. Gonda).

https://doi.org/10.1016/j.euroneuro.2025.05.007
Received 15 February 2025; Received in revised form 2 May 2025; Accepted 8 May 2025
Available online 8 June 2025

1. Introduction

Major Depressive Disorder (MDD) is a prevalent and debilitating psychiatric condition with a complex etiology, influenced by a combination of genetic, environmental, and personal risk factors (Malhi and Mann, 2018). This complexity results in heterogeneous clinical manifestations and wide variations in response to treatment, with one-third of patients showing treatment resistance to classical antidepressant medications (Perlis, 2016). Thus, one of the key challenges in MDD management is identifying the optimal pharmaceutical intervention considering the significant inter-individual variation in drug response (Barlati et al., 2023; Kishi et al., 2023), which is likely underlined by a divergent neurobiological background, in part reflected in distinct patterns of psychiatric and somatic comorbidities.

To optimise treatment selection one promising approach is to unravel the complexity of MDD by applying advanced analytical methods to large-scale health data in order to identify disease subgroups and their associated characteristics (Dwyer et al., 2018) on the clinical, phenotypic, and genomic levels. In the TRAJECTOME project (Hullam et al., 2023) we took significant steps toward understanding the heterogeneity of MDD by conducting a comprehensive analysis of longitudinal disease trajectories from 1.2 million individuals across three European population cohorts, including both MDD and nonMDD cases. Using only MDD-related multimorbidity trajectories for cluster identification, we identified seven clusters and demonstrated that each has a unique clinical, genetic, and non-genetic risk profile (Gezsi et al., 2024) (Table 1): Clusters 1–2 exhibit late MDD onset, low overall disease burden, and decreased prevalence of most diseases, indicating that individuals in these clusters tend to remain relatively healthy until old age; Clusters 3–4, despite a late onset of MDD and relatively low disease burden, show a propensity for cerebrovascular and metabolic disorders, respectively; Cluster 5 is characterized by early MDD onset, high disease burden, and is strongly associated with schizophrenia and musculoskeletal disorders that lead to pain, highlighting a higher risk profile; Cluster 6 also shows a high disease burden, but with a later onset of MDD, featuring prevalent stress reactions, somatoform disorders, and respiratory infections, likely tied to severe stress exposure; Cluster 7, which has an early MDD onset, primarily involves allergic and respiratory inflammatory diseases, along with migraines and dermatitis, while the disease burden for other diseases remain low.

Additionally, the utility of the clusters defined in our previous work is further demonstrated by two of our recently published studies. The first study demonstrated that several previous candidate gene findings with a focus on the impact of childhood trauma burden in the development of depression, can be replicated using the cluster membership probabilities instead of MDD diagnosis (Bonk et al., 2024), further supporting biological heterogeneity of depression. The second study introduced the Multimorbidity-Adjusted Disability Score (MADS), a novel tool that leverages the comorbidity network of MDD, that successfully enhanced health risk stratification, predicted disease progression, and pharmacological and nonpharmacological expenditures, particularly in the context of MDD (González-Colom et al., 2024).

Although the currently used treatment guidelines (Kendrick et al., 2022; Lam et al., 2024; Malhi et al., 2021) underwent a significant evolution over the past decades, they generally fail to address subtypes of depression and provide minimal guidance on how to address different within-category clinical differences with distinct pharmacological approaches. The improvements of the guidelines in part reflect changes in the official diagnostic systems including the International Statistical Classification of Diseases and Related Health Problems (ICD) (World Health Organization, 2004) or Diagnostic and Statistical Manual of Mental Disorders (DSM) (First, 2013), and in part correspond to novel and emerging high-quality evidence reflecting our changing knowledge concerning the efficacy of different treatment and pharmacological approaches in specific disease categories. There has also been some progress to include several subtypes or qualifiers/specifiers for MDD in the

Table 1
Major depressive disorder-related multimorbidity clusters based on the TRAJECTOME project (Gezsi et al., 2024).

Cluster ID	Level of Disease Burden	Prevalent Diseases	MDD risk	Time of MDD Onset	Genetic correlation with MDD/BD
1	Low	Few; overall decreased prevalence of most diseases	low	Late	-/-
2	Low	Few; overall decreased prevalence of most diseases	low	Late	-/-
3	Low	Cerebrovascular disease, Kidney diseases, Hypertension	average	Late	0
4	Low	Lipid metabolic disorders, Hypothyroidism	below average	Late	0
5	High	Schizophrenia, Musculoskeletal system diseases leading to pain disorders	high	Early	+/0
6	High	Reaction to severe stress, Somatoform disorders, Respiratory tract infections	high	Late	+/0
7	Variable (mostly low)	Allergic and respiratory system inflammatory diseases, Migraine, Dermatitis	high	Early	0/+

Each cluster is described based on its level of disease burden (number of comorbidities), prevalent diseases associated with the cluster, significant genetic correlation with MDD and BD (based on Psychiatric Genomic Consortium data), the risk (hazard ratio) of having MDD diagnosis (F32 or F33), and the time of MDD onset (members develop MDD early or late in their lifetime), -: significant negative genetic correlation, 0: no significant genetic correlation, +: significant positive genetic correlation.

official classification systems. However, these generally correspond to either the time or circumstances of the onset of the disorder, or some well-observable characteristics in manifestation. Thus, a more sophisticated subtyping corresponding to the putative genetic and neurobiological background is still lacking. In addition, updates to treatment guidelines are mostly confined to the inclusion of newly approved pharmacological treatments and other interventions, and incorporate only a few features to account for different clinical characteristics or associated conditions (Kendrick et al., 2022).

Nevertheless, based on the specific genetic and disease profile of the seven MDD-related multimorbidity clusters (Gezsi et al., 2024) we would expect that approaches to effective MDD treatment would also be distinct in the above clusters, providing an important tool and markers for more efficient and personalized treatment decisions in depression. Therefore, building on our previous work (Gezsi et al., 2024), this study aims to conduct a pharmacological profiling of MDD-related multimorbidity clusters using the UK Biobank (UKB) data, focusing on a narrower sub-cohort of participants born between 1946–1948 as the primary cohort. In addition, we compare the primary findings to the full dataset of UKB, which allows us to test the effect of evolving treatment guidelines in a broader age range. Furthermore, data from other countries, from Catalan Health Surveillance System (CHSS) cohort (Catalonia, Spain), and the Finnish Institute for Health and Welfare (THL) cohort (Finland) are used to identify similarities and differences between geographically different healthcare systems. Pharmaceutical profiling can elucidate trends in polypharmacy (the concurrent use of multiple medications) (Masnoon et al., 2017) and drug burden (the number of prescriptions), potentially highlighting variations in treatment patterns among the seven different MDD-related multimorbidity clusters in general, and specifically focusing on antidepressant treatment.

2. Methods

2.1. Cohorts and participants

The primary cohort of our analyses is a sub-cohort of the UK Biobank (UKB, Application No: 1602) dataset, born between 1946 and 1948, inclusive, whose cluster membership probabilities were successfully determined, and who have drug prescription data (UKB $sample_{3\text{-}year}$). This decision was motivated by several factors. First, by limiting the focus to just three years, we aimed to create a more homogeneous group regarding life experiences and interactions with the healthcare system. This approach helps to reduce potential confounding in our data that could arise from shifts in healthcare policies or medical practices over a more extended period. Second, this three-year range represents the most densely populated age group within the UKB dataset (SFigure 2), covering $n = 25{,}063$ participants (16.3 %), and thereby increasing statistical power for our analyses. Finally, selecting participants based on birth year allows a more consistent comparison between different cohorts.

Our primary cohort results were compared to the results from the UK Biobank full cohort (UKB $sample_{full}$, without restriction for birth year, $n = 157{,}450$), and to sub-cohorts born between 1946 and 1948 ($sample_{3\text{-}year}$), inclusive, from the Catalan Health Surveillance System (CHSS, $n = 14{,}444$) cohort (Catalonia, Spain), and from the Finnish Institute for Health and Welfare (THL, $n = 2514$) cohort (Finland). For description of the comparison cohorts and for the ethical information regarding the investigated cohorts see Supplementary Methods.

2.2. Treatment-resistant depression cases

Diagnosis of MDD in individuals from the UK Biobank (UKB), CHSS, and THL cohorts was based on lifetime ICD-10 codes (F32 or F33) extracted from electronic health records (EHR), aligning with standard biobank methodologies. In case of UK Biobank, we used the "first occurrence" data fields (Category 1712), which also holds information from primary care, death register, hospital inpatient data and self-report. To more accurately characterize depressive profiles within each cluster, we defined treatment-resistant depression (TRD) for eligible participants, following the operationalisation of the general definition of inadequate response to a minimum of two antidepressants used in an appropriate dose for an appropriate time according to the criteria proposed by Fabbri et al. (Fabbri et al., 2021). This process is described in more detail in the Supplementary Methods, highlighting some minor deviations from the Fabbri et al. methodology (see also SFigure 5). In our study TRD-status refers to a sub-sample of participants, where it can be determined with certainty if they have TRD or not. In this context, TRD refers to individuals who have treatment-resistant depression, and nonTRD refers to MDD patients without treatment resistance (see Fig. 1).

2.3. Investigated phenotypes

2.3.1. Posterior probability of MDD-related multimorbidity cluster membership

Detailed description to derive the posterior probability of MDD-related multimorbidity cluster membership can be found in our previous paper (Gezsi et al., 2024). In short:

1. A Bayesian network-based Markov Chain Monte Carlo (BN-MCMC) (Marx et al., 2017) method was used to select comorbidities (minimum prevalence of 1 %) that are directly related to MDD, i.e. no other diseases mediated the relationship between MDD and the given comorbidity. The analysis was carried out in 4 time intervals (aged [0–20], [0–40], [0–60], and [0–70]) throughout the lifespan, and in 3 discovery cohorts (full samples without exclusions of UKB, THL, CHSS) resulting in 86 cross-cohort diseases (directly related to MDD for at least one time interval and for at least one cohort, and can be investigated in all cohorts).

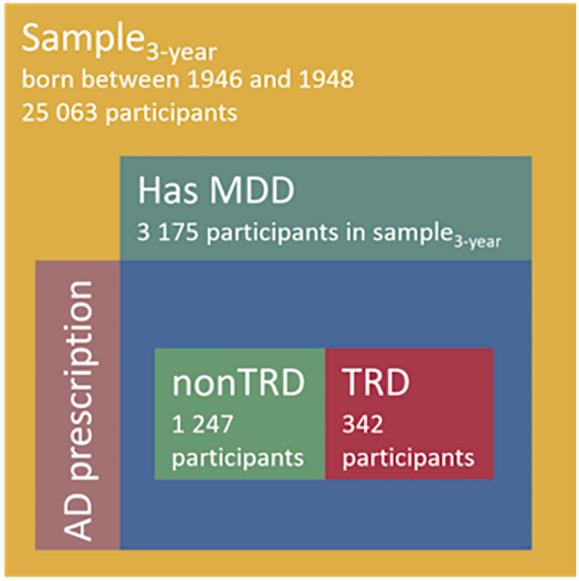

Fig. 1. Distribution of the UKB $sample_{3\text{-}year}$ cohort born between 1946 and 1948. Yellow: participants whose cluster membership probabilities were successfully determined, and who have drug prescription data (UKB $sample_{3\text{-}year}$); Purple: participants with at least one antidepressant (AD) prescription; Blue: participants with diagnosed major depression (MDD; based on the presence of F32 or F33 ICD-10 diagnosis); Red: participants who developed treatment resistant depression (TRD) based on their drug treatment profile; Green: participants who's AD prescription profile indicate a successful treatment.

2. Next, to determine the MDD-related multimorbidity burden throughout the participants' lifespan, MDD-related multimorbidity scores were computed for each participant and for each of the 4 time intervals based on the cross-cohort relevance scores of the selected 86 diseases.
3. The MDD-related multimorbidity scores were then used to identify the seven MDD-related multimorbidity clusters that have distinct temporal MDD-related multimorbidity burden profiles.
4. Based on the clustering, we were able to derive posterior probability of cluster membership variables for each participant and for each MDD-related multimorbidity cluster, which was used as a continuous weighting factor throughout the current study. The rationale behind this is that over the seven clusters, posterior probability adds up to 1 per person, thus each participant contributes the same amount to the analysis overall, but they have more influence over the statistics of the clusters they are more likely associated with.

2.3.2. UKB drug prescription data from primary care

At the time of our study, primary care data had been obtained for approximately 45 % of the UK Biobank cohort (about 230,000 participants). Because of the available data's completeness and level of detail varied between systems and suppliers (UK Biobank, 2019), we specifically used data from the SystemOne practice management system, provided by The Phoenix Partnership (TPP, https://tpp-uk.com/) for England, as this is the largest homogenous provider, encompassing about 165,000 participants. Because the boundaries of the data availability are not defined in an exact way for this cohort, we conservatively used the hard limits of January 1990 to January 2016 as it ensures complete time coverage of the TPP cohort. For our analysis, we leveraged the British National Formulary (BNF) drug codes (Dimond, 2003) and mapped these to the Anatomical Therapeutic Chemical (ATC) Classification System via the Systematized Nomenclature of Medicine (SNOMED) (El-Sappagh et al., 2018) to ensure comparability to other (CHSS, THL) cohorts. The mapping files can be found in the Supplementary Data. It is essential to note that these data were collected for patient care administration, which may implicate data quality issues, including change of general practitioners and potential gaps in data coverage that cannot be corrected in this study.

For drug dispensation data of CHSS and THL see Supplementary Methods section.

2.4. The ATC hierarchy

In our study, we aggregate and analyse data across different levels of the ATC classification system for drugs and medical products. The hierarchical structure of the ATC system allows for a multilevel analysis of medication data, ranging from broad anatomical categories to specific chemical compounds.

2.4.1. Antidepressant profiling based on synaptic targets

Because the statistical power was diminished at lower levels of the ATC hierarchy, an alternative categorization was also used for antidepressant drugs, taking into account the antidepressant actions at the synapse, based on (Malhi and Mann, 2018). This enabled us to aggregate drugs based on transporters and receptors they target, forming a one-to-many drug-to-target mapping. The exact mapping used can be seen in STable 3.

2.5. Calculation of multimorbidity, polypharmacy, and drug burden metrics

We defined three key epidemiological metrics based on disease onset and medication data: multimorbidity (the number of diseases), polypharmacy (number of consumed drug-types), and drug burden (number of prescriptions). Details about the derivation and usage of these variables can be found in the Supplementary Methods.

2.6. Statistics

2.6.1. Demonstrating the temporal characteristics of multimorbidity, polypharmacy and drug burden

A cumulative multimorbidity metric (number of disease onsets at or before the age) was calculated for every year of participant age. For polypharmacy (number of chemical substances) and drug burden (number of prescriptions), the values were counted at a given year of age (non-cumulative). To compare the temporal change in each cluster, cluster membership probability weighted average multimorbidity (top), polypharmacy (middle), and drug burden (bottom) in each cluster were plotted for each cohort.

2.6.2. Assessing the impact of cluster membership on cumulative metrics with regression analyses

To better understand the influence of the clusters on various metrics, we conducted regression analyses. The rationale behind this included the fact that statistical testing for differences between the clusters was not feasible with classical methods, as each participant was a member of each cluster, weighted by the posterior probability of MDD-related multimorbidity cluster membership. Thus, a significant association was determined based on the p-value of the corresponding cluster membership coefficients in the regression model. To compensate for the number of tests, we used Bonferroni correction. All analyses were performed using the statsmodels (v0.14.0) Python module. The predictor variables were standardized before model fitting. Further details of the regression analysis are available in the Supplementary methods.

Positive values of the regression coefficients for cluster probability log-odds signify that a higher likelihood of belonging to a cluster is associated with an increase in the outcome variable. Negative values indicate the opposite effect, linking higher cluster membership to lower outcome values.

2.6.3. Word tree analysis of antidepressant sequences

To identify common sequences of antidepressant treatment among different clusters, we employed a word tree approach, a concept borrowed from text analysis (Crichton, 2020/W. 2023). Here, each participant's drug prescription history was treated as a 'sentence,' where 'No drug' was the first 'word,' followed by the first prescription of different antidepressants in chronological order.

To perform this analysis, we used the open-source Python implementation of the word tree algorithm available at https://github.com/willcrichton/wordtree. We limited this analysis to antidepressants (ATC code starting with N06A) to focus on the typical sequences of treatment for MDD. Each branch of the resulting word tree represents a unique sequence of prescriptions. By comparing the weighted word trees for each cluster, we aimed to discern patterns in the sequences of antidepressant treatment across clusters.

3. Results

3.1. Descriptive statistics

In the UKB $\text{sample}_{3\text{-year}}$ cohort majority of the participants have a higher posterior probability belonging to clusters 1–4, which showed a protective profile regarding MDD and most of its comorbidities in our original study (Gezsi et al., 2024). Belonging to clusters 5–7 have a lower posterior probability but in these clusters a higher percentage of participants have MDD (F32 or F33) diagnosis, which is highest for cluster 6, followed by cluster 5 and cluster 7 (STable 2). It is noteworthy that depression shows a relatively late onset, with 48.9 years in UKB $\text{sample}_{3\text{-year}}$, while cluster 5 shows the earliest mean onset age of 45.73 years.

Descriptive statistics across replication cohorts (UKB $\text{sample}_{\text{full}}$, the CHSS $\text{sample}_{3\text{-year}}$ and THL $\text{sample}_{3\text{-year}}$) are available in the supplementary tables S2 for comparison, and the main characteristics were

similar to UKB sample$_{3\text{-year}}$. A notable difference is that in case of UKB sample$_{full}$ cluster 5 has a higher posterior probability share (16.75 %) and the highest percentage of participants with MDD diagnosis (19.46 %), although closely followed by cluster 6 (19.37 %). Another difference is that cluster 3 has the second highest percentage of MDD in the CHSS sample$_{3\text{-year}}$ (21.91 %), and the third in THL sample$_{3\text{-year}}$ (12.26 %). In cohorts UKB sample$_{full}$ and THL sample$_{3\text{-year}}$ cluster 5 shows the earliest mean onset age of MDD (41.03 and 51.64 years, respectively), similarly to UKB sample$_{3\text{-year}}$, while it was cluster 6 (60.60 years) in the CHSS sample$_{\mathbf{3\text{-year}}}$ cohort.

3.2. Temporal characteristics of polypharmacy and drug burden mainly follow multimorbidity

The patterns of multimorbidity differed notably across clusters (Fig. 2). Clusters 5 and 7 showed an early onset of multimorbidity. Cluster 6 reached above-average multimorbidity later in life, and remained the cluster with highest multimorbidity values followed by cluster 5 and 7. Interestingly, cluster 3 showed a sharp increase in multimorbidity after 65 years of age and reached the second most prevalent multimorbidity values by 70 years of age.

Polypharmacy and drug burden follow similar characteristics. However, taking into account that pharmacological data was available for the period of 1990–2016, compared to multimorbidity, the early onset of clusters 5 and 7 is not reflected in drug prescriptions. Also, while clusters 5–7 reached above-average multimorbidity values at the end of prescription data coverage, regarding polypharmacy and drug burden the same trend could be seen only for cluster 6 and cluster 5.

The corresponding figures for the UKB sample$_{full}$, CHSS sample$_{3\text{-year}}$ and THL sample$_{3\text{-year}}$ cohorts can be found in the Supplementary material (SFigures 6–8). Multimorbidity patterns replicated in all the three replication cohorts. In the UKB sample$_{full}$ polypharmacy and drug burden followed the pattern of multimorbidity throughout the ages with cluster 3 approaching cluster 6 and overtaking cluster 5 in polypharmacy and drug burden by 70 years of age. In the CHSS sample$_{3\text{-year}}$ cohort, the limited pharmacological data available only from 2011 resulted in a pattern of polypharmacy and drug burden for participants in their late sixties where cluster 6, cluster 3 and cluster 5 showed the highest values, respectively, followed by above average polypharmacy and drug burden values in cluster 7. In the THL sample$_{3\text{-year}}$ cohort for

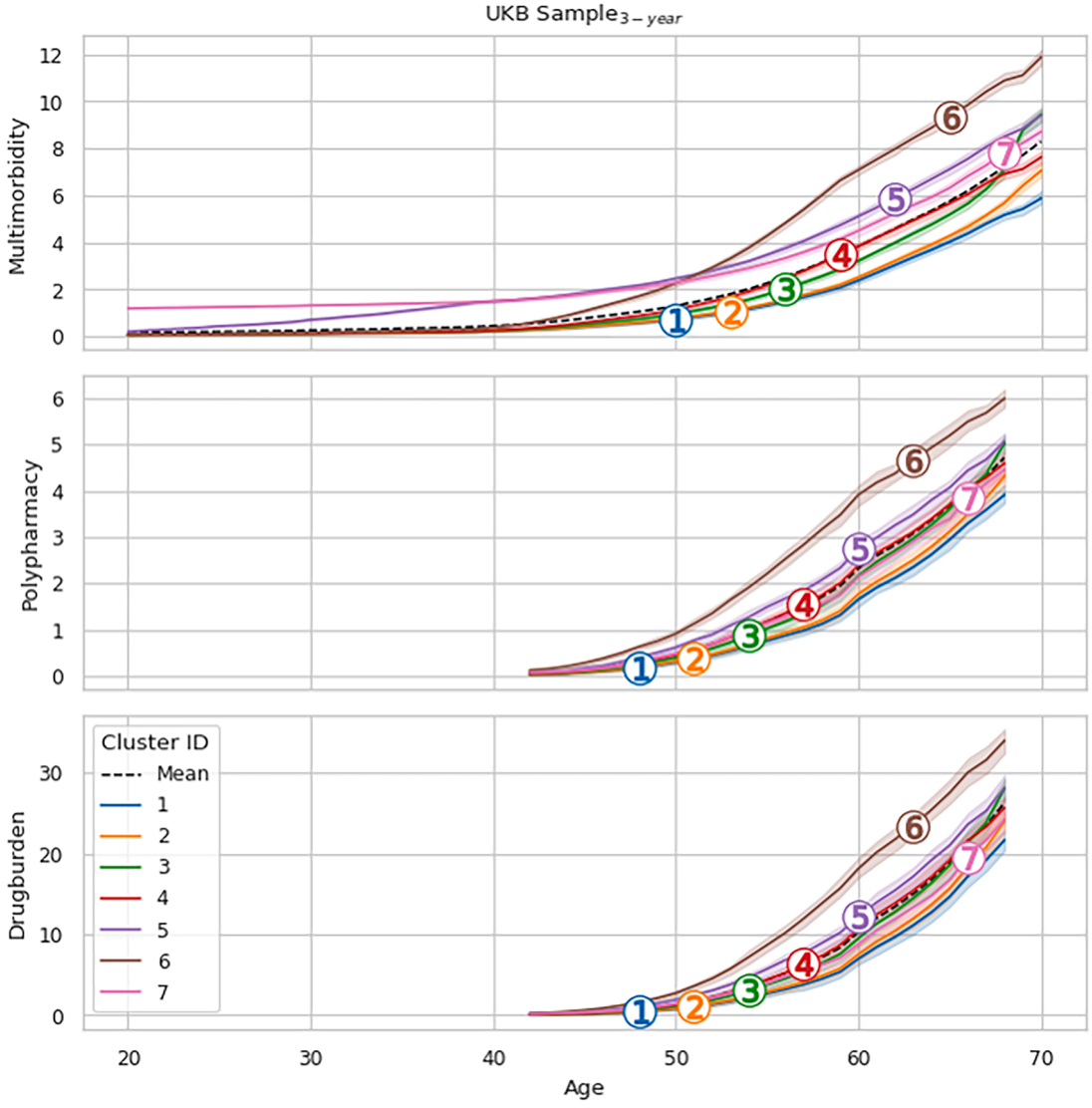

Fig. 2. Temporal characteristics of multimorbidity, polypharmacy and drug burden in UK Biobank sample$_{3\text{-year}}$. Temporal change of cluster membership probability weighted average multimorbidity (top), polypharmacy (middle), and drug burden (bottom) in each cluster for participants of the UKB sample$_{3\text{-year}}$. A cumulative multimorbidity metric (number of disease onsets at or before the age) was calculated for every year of participant age. For polypharmacy (number of chemical substances) and drug burden (number of prescriptions), the values were counted at a given year of age (non-cumulative).

polypharmacy the pattern follows multimorbidity data, while for drug burden cluster 5 has the highest value by 70 years of age.

3.2.1. Significant cluster differences in cumulative multimorbidity, polypharmacy, and drug burden metrics

The regression analysis of cumulative multimorbidity, polypharmacy and drug burden data in the UKB sample$_{3\text{-year}}$ showed similar results to each other, with cluster 6 having the highest regression coefficient, followed by cluster 5, and with the lowest values for clusters 1 and 2 (Fig. 3A). Adjusting the regression analysis for multimorbidity by adding it as a covariate in the equations we can see that general trends are retained for clusters 1, 2, 4 and 5–6. For cluster 3 the sign of the coefficients has changed, indicating a higher polypharmacy and drug burden independently of the number of multimorbidities, though this is only nominally significant. For cluster 5, the coefficients are reduced, while for cluster 6, a strong association is retained. However, the largest coefficients for cluster 6 decrease to about one-fifth of their previous value, indicating that multimorbidity accounts for a substantial portion of the variance in the model.

Regarding the regression analysis of multimorbidity, polypharmacy and drug burden, with age and sex as covariates in the model, in the UKB sample$_{full}$ (Fig. 3B), CHSS sample$_{3\text{-year}}$ (Fig. 3C) and THL sample$_{3\text{-year}}$ (Fig. 3D) cohorts we replicated that cluster 6 has the highest regression coefficients, followed by cluster 5, and clusters 1–2 have the lowest values. While cluster 3 has lower than average values for multimorbidity, polypharmacy and drug burden in the UKB sample$_{full}$ cohort (similarly to UKB sample$_{3\text{-year}}$), these were significantly higher in the CHSS sample$_{3\text{-year}}$ and THL sample$_{3\text{-year}}$ cohorts.

After correcting for multimorbidity, the polypharmacy and drug burden pattern became more divergent between cohorts, likely due to the different timeframes covered by the multimorbidity and medication data. Regarding polypharmacy, replicated findings are the highest regression coefficients corresponding to cluster 6 both in the UKB sample$_{full}$ and the THL sample$_{3\text{-year}}$ cohorts, and the lowest regression coefficient corresponding to cluster 1 in all cohorts. Regarding drug burden, the highest regression coefficients are related to cluster 6 in all cohorts, although becoming non-significant in the CHSS sample$_{3\text{-year}}$ and THL sample$_{3\text{-year}}$ cohorts. The lowest drug burden regression coefficients replicated for clusters 1 and 2 in all cohorts, but with a diminished significance in the CHSS sample$_{3\text{-year}}$ and THL sample$_{3\text{-year}}$ cohorts.

More detailed results regarding how cluster membership affects drug burden in terms of the main ATC categories in our cohorts can be seen in SFigure 9.

3.2.2. Cluster membership significantly affects drug burden of antidepressants in MDD patients

Next, we focused on MDD patients (ICD-10 F32 or F33 diagnoses) and their antidepressant treatment. Comparison of chemical subgroups within the antidepressant categories (ATC code N06A) showed significant differences between clusters (Fig. 4). Namely, in the UKB sample$_{3\text{-year}}$ cohort clusters 1–2 are both negatively associated with categories

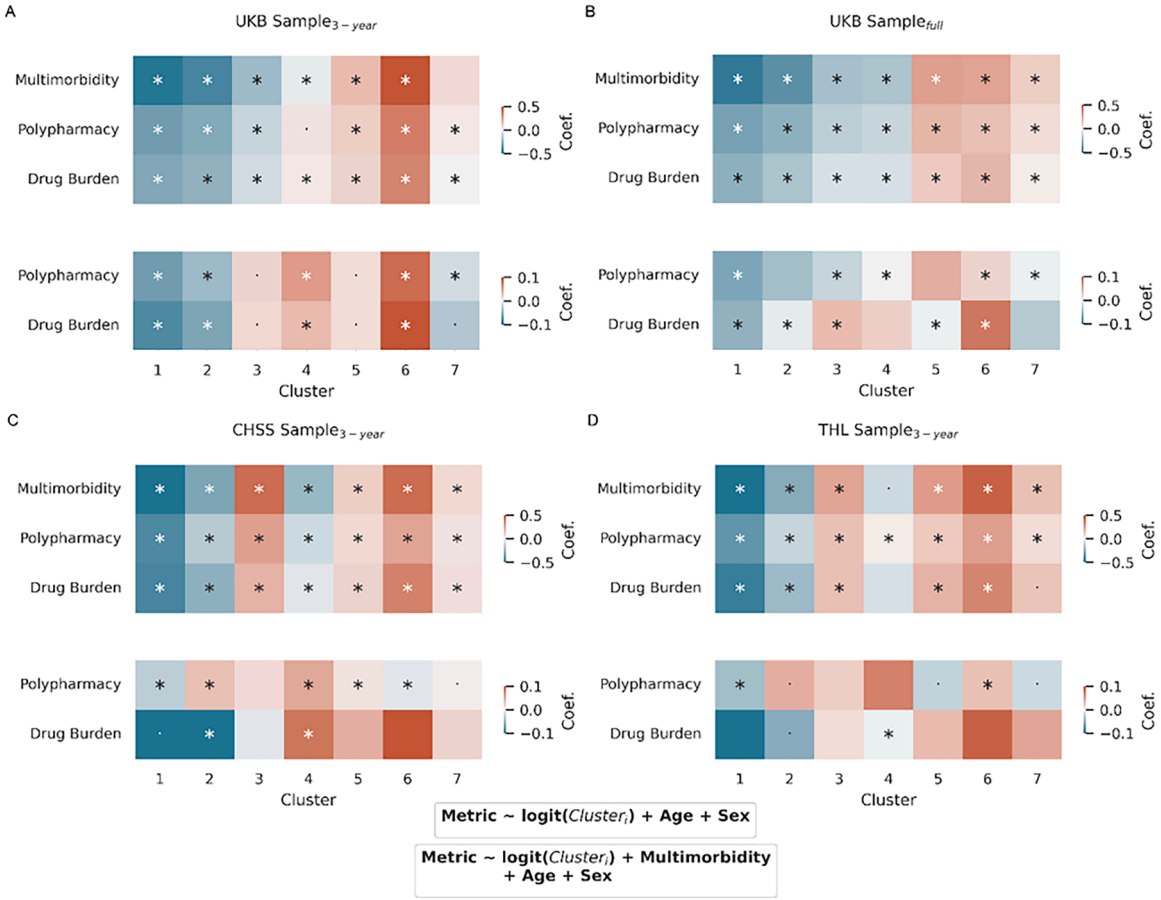

Fig. 3. Effects of cluster-membership on the cumulative multimorbidity, polypharmacy and drug burden metrics in the different cohorts. Coefficients for the cluster memberships log-odds of each cluster in the linear regression model with main metrics as outcome variables, performed on A - UKB sample$_{3\text{-year}}$; B - UKB sample$_{full}$; C - CHSS sample$_{3\text{-year}}$; D - THL sample$_{3\text{-year}}$. Top- results within each panel include multimorbidity as an outcome; age and sex as covariates. Bottom panels showcase the same regression analysis, but multimorbidity was included as a covariate (rather than outcome) in each model. Asterisk (*) marks significant results after Bonferroni correction; point (·) marks nominally significant results ($p < 0.05$). In case of THL, the polypharmacy metric was calculated based on the uniqueness of level 3 ATC codes, as the full length was not available.

N06AA ("Non-selective monoamine reuptake inhibitors," incorporating TCAs) and N06AB (SSRIs). On the other hand, cluster 6 membership was positively associated with increased drug burden in these two categories. Interestingly, clusters 5 and 7, which are other MDD risk clusters, show no significant associations with antidepressant chemical subgroups, except for a nominally significant positive association for cluster 5 with the N06AA ("Non-selective monoamine reuptake inhibitors," incorporating TCAs) ATC category (Fig. 4).

As data for ATC level 4 and level 5 are not available in the THL sample$_{3\text{-year}}$ this cohort was not used for comparisons in the remaining analyses.

Regarding the UKB sample$_{full}$ the overall pattern for chemical subgroups within the antidepressant categories (ATC: N06A) replicated. In addition, the regression coefficients showed a positive significant association for N06AA ("Non-selective monoamine reuptake inhibitors," incorporating TCAs), N06AB (SSRIs), and N06AX (Other antidepressants) both in cluster 5 and 6, while clusters 1–4 were significantly negatively associated with the drug burden of these chemical subgroups, with no significant results emerging for cluster 7.

In the CHSS sample$_{3\text{-year}}$ cohort, despite the similar overall pattern, only the negative significant association of the drug burden regarding N06AB (SSRIs) replicated for cluster 1.

Further details regarding cluster membership effects on drug burden of antidepressants at the ATC 5th level (chemical substances) in MDD patients can be seen in SFigure 10.

3.2.3. *Cluster membership significantly affects synaptic target profile of antidepressants in MDD patients but in a cohort-dependent manner*

To investigate whether specific classes and synaptic targets of the prescribed antidepressants show differences according to the MDD-related multimorbidity clusters we re-categorized the ATC-coded antidepressants according to their physiological targets based on (Malhi and Mann, 2018) (STable 3), and calculated a normalized value of drug burden: percentage of prescriptions of the drug target within antidepressants, e.g. number of prescriptions of SSRIs divided by the number of prescriptions of all AD (see Supplementary methods).

In the UKB sample$_{3\text{-year}}$ cohort using this method we can see that the percentage of prescribed SSRIs is nominally increased for clusters 2–3, and the percentage of prescribed TCAs is nominally decreased in clusters 2 and 4, and there is a significant negative association with normalized drug burden for TCAs in cluster 3. For cluster 5 the opposite pattern emerged: nominally significant associations for higher percentage of prescribed TCAs and lower percentage of prescribed SSRIs. In line with the AD class data, the synaptic drug targets in cluster 5 showed significant positive associations with postsynaptic M1 muscarinic acetylcholine receptor and NA transporter, and nominally significant positive associations with postsynaptic serotonin 5-HT2 receptor, histamine H1 receptor, and noradrenalin alfa-1alfa-2 receptor. For clusters 1–4, although several AD targets were nominally significantly negatively associated, the only significant finding was a significant negative association of postsynaptic M1 muscarinic acetylcholine receptor with cluster 3 (Fig. 5).

In the UKB sample$_{full}$ cohort very similar pattern emerged but with more significant findings.

However, in the CHSS sample$_{3\text{-year}}$ cohort the pattern does not resemble the UKB results, suggesting a different treatment selection strategy for MDD patients, probably affected by the more recent and shorter period medication data, and potential healthcare differences. In the CHSS sample$_{3\text{-year}}$ cohort only nominally significant findings were observed, mainly related to cluster 3 and postsynaptic histamine H1 receptor, M1 muscarinic acetylcholine receptor, noradrenalin alfa1alfa2 receptor, and noradrenalin transporter as physiological drug targets.

3.3. *Cluster membership effect on treatment-resistance in MDD patients*

For participants where TRD-status was determinable, we defined the baseline as cluster 1 and calculated odds ratios compared to other clusters (Table 2). Based on this, in the UKB sample$_{3\text{-year}}$ cohort cluster 5

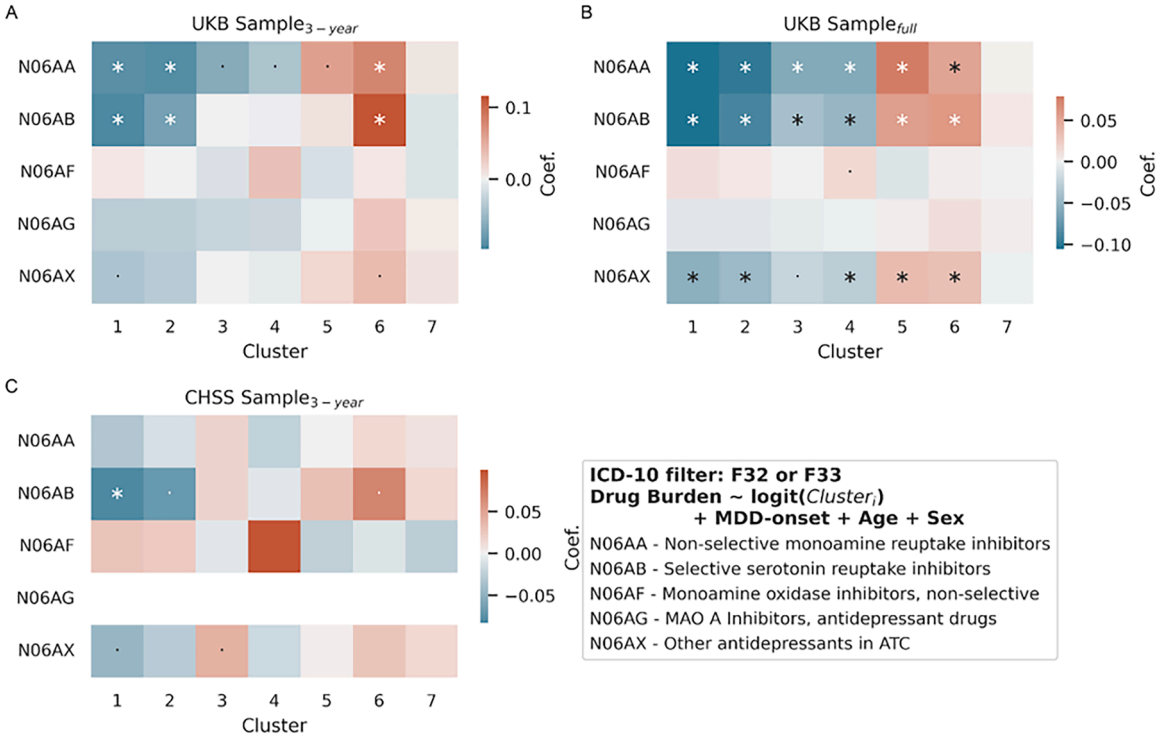

Fig. 4. Effects of cluster membership in MDD patients on the drug burden of the antidepressant chemical subgroups ATC categories (4th level). Coefficients for the cluster memberships log-odds of each cluster in the linear regression model with drug burdens in chemical subgroups of the ATC antidepressant (N06A) category as outcome variables, performed on MDD patients of A - UKB sample$_{3\text{-year}}$; B - UKB sample$_{full}$; C - CHSS sample$_{3\text{-year}}$. Age, sex, and depression onset age were included as covariates in the models. Asterisk (*) marks significant results after Bonferroni correction; point (·) marks nominally significant results ($p < 0.05$).

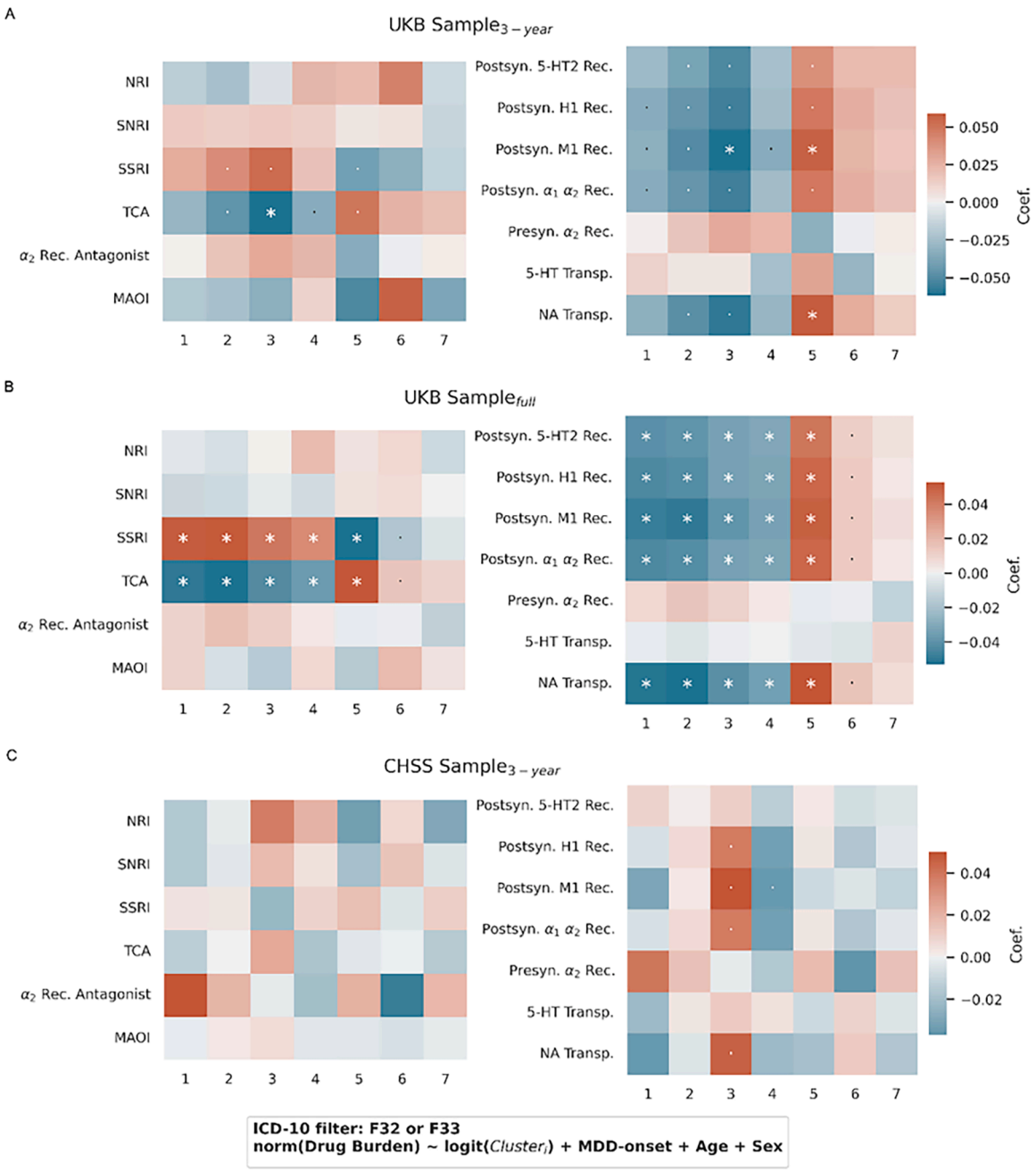

Fig. 5. Effects of cluster membership on the normalized drug burden of the alternative antidepressant categories and physiological targets. Coefficients for the cluster memberships log-odds of each cluster in the linear regression model with normalized AD drug class burden (left sub-panels) and normalized physiological target burden (right sub-panels), performed on MDD patients of: A - UKB sample$_{3\text{-year}}$; B - UKB sample$_{full}$; C - CHSS sample$_{3\text{-year}}$. In this case, the AD consumption was summarized and each target variable (e.g. drug class or physiological targets) represents a percentage share taken from the total drug-burden of AD treatment. Age, sex, and depression onset age were included as covariates in the models. Asterisk (*) marks significant results after Bonferroni correction; point (·) marks nominally significant results ($p < 0.05$). Rec=receptor. Transp=transporter. 5-HT=serotonin. NA=noradrenaline. *H*=histamine. *M*=muscarine. DA=dopamine. MAO=monoamine oxidase. TCAs=tricyclic antidepressants. NDRIs=noradrenaline dopamine reuptake inhibitors. SSRIs=selective serotonin reuptake inhibitors. SNRIs=serotonin-noradrenaline reuptake inhibitors. NRIs=noradrenaline reuptake inhibitors. Presyn= presynaptic targets. Postsyn=postsynaptic targets.

showed the highest OR with 1.80, followed by cluster 6 with 1.52, the other clusters' ratio was closer to 1.

By performing logistic regression for the TRD-status as the outcome variable, we were able to compensate for the effects of different age and sex distributions in the clusters. The results show a significant positive association between TRD and cluster 5 and 6 membership, which also have the highest percentage of participants with TRD. All the other clusters were significantly negatively associated, and the results for cluster 7 were not significant.

Results for the UKB sample$_{full}$ (STable 6) and for CHSS sample$_{3\text{-year}}$ (STable 7) are available in the Supplementary material. Notably, the results of the UKB sample$_{3\text{-year}}$ and the UKB sample$_{full}$ are almost identical with an exemption that cluster 6 has an equal TRD risk as cluster 5. Regarding the CHSS sample$_{3\text{-year}}$ cohort, the shorter drug data coverage made it more challenging to determine TRD status, increasing the prevalence of false-negative non-TRD records. This might be responsible for the generally lower TRD prevalence and different TRD profile, where cluster 7 and cluster 3 showed nominally significant increased risk for

Table 2
Weighted prevalence of treatment-resistant depression (TRD) and association with cluster membership in UK Biobank sample$_{3\text{-year}}$.

Cluster ID	TRD prevalence	$OR_{ci/c1}$	Coef.	p-value	Signif.
1	12 %	1.00	−0.2823	6.54E-05	*
2	12 %	1.01	−0.2577	1.43E-04	*
3	13 %	1.05	−0.2017	2.15E-03	*
4	13 %	1.05	−0.1927	3.84E-03	*
5	22 %	1.80	0.1079	8.57E-04	*
6	19 %	1.52	0.1012	2.94E-03	*
7	14 %	1.14	0.0035	9.17E-01	

Distribution of treatment-resistant depression (TRD) in each cluster, as cluster probability weighted percentage of participants with known TRD-status (TRD or nonTRD) and results of the logistic regression to characterize the influence of cluster membership on TRD-status (covariates included age and sex). Odds-ratios (OR) are compared to 100 % certainty of cluster 1 membership.

TRD.

3.4. Cluster membership effect on antidepressant treatment selection strategies in MDD patients represented by drug trees

To understand the sequence of antidepressant treatments among participants with MDD diagnosis within each cluster, we conducted a drug tree analysis, akin to a word tree, which mapped the most common prescription paths initiated by the first-time prescription of a new drug for a given individual (Fig. 6). Irrespective of the cluster memberships, we calculated the distribution of individuals on these treatment trajectories, and it is shown as the percentage of individuals who reached at least the given node in the tree. After that we calculated the same percentages weighted with cluster memberships and calculated a ratio with the overall population, to see if a given trajectory is characteristic for a cluster.

Overall, the most frequent first-line AD was amitriptyline (N06AA09) with 40.54 %, which was especially overrepresented in clusters 5 and 6. Three SSRIs, fluoxetine (N06AB03, 16.07 %), sertraline (N06AB06, 12.74 %) and citalopram (N06AB04, 10.61 %) were the second, third, and fourth most popular first choice ADs. They were especially characteristic for clusters 1, 2, 3 and 4. Cluster 7′s characteristics trajectories also included citalopram and the fifth most common first-line AD, mirtazapine (N06AX11, 8.68 %). Further down the tree, the general tendency was a reduced representation of clusters 1–4; an increased representation of clusters 5–6; and varying tendencies for cluster 7.

In the UKB sample$_{full}$, a similar trend could be observed (SFigure 11), except that citalopram (25.27 %) overtook other SSRIs; but not amitriptyline, which remained the most popular first choice with 39.37 %). In case of the CHSS sample$_{3\text{-year}}$ (SFigure 12), possibly because of the short and recent drug coverage period, the top 3 first-line ADs (paroxetine 19.34 %, citalopram 14.23 % and sertraline 12.41 %) were from the SSRI class.

4. Discussion

Our results support that MDD-related multimorbidity clusters, identified in the TRAJECTOME project (Gezsi et al., 2024), not only showed distinct multimorbidity profiles but also specific and different polypharmacy and drug burden characteristics. These differences suggest that the neurobiological and genetic underpinnings of depressive syndromes influence treatment choices in large populations. The present results support this, demonstrating that the clusters identified based on MDD-related multimorbidity trajectories also showed significant differences in polypharmacy and drug burden, even after appropriate statistical corrections. (Gezsi et al., 2024). Thus, significant variations in drug usage patterns across different clusters of MDD patients have important implications for clinical practice and health policy, emphasizing the need for personalized treatment strategies based on individual comorbidity profiles.

Accumulating evidence suggest that depression is not only frequently comorbid with other somatic disorders (Gold et al., 2020; Read et al.,

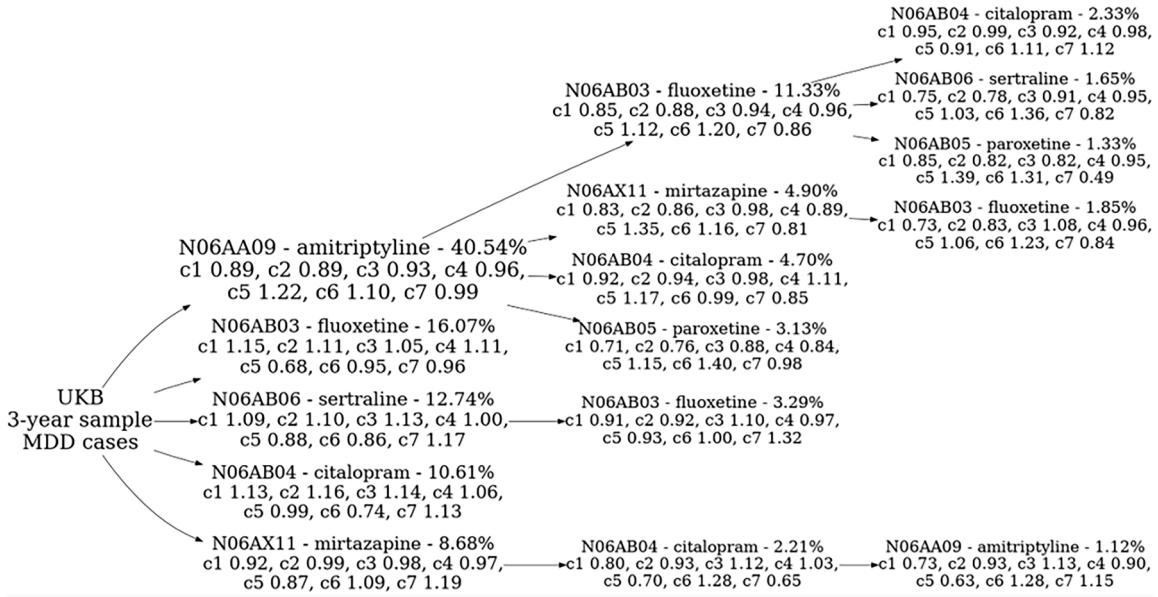

Fig. 6. "Drug tree" representation of the most common antidepressant treatment courses of MDD cases in UK Biobank sample$_{3\text{-year}}$. This figure illustrates the antidepressant treatment paths for MDD cases in the UK Biobank sample$_{3\text{-year}}$. Each percentage shown at a given point of the branch represents the proportion of the entire sample that was prescribed the corresponding drug at that stage in the treatment path. The cluster values (c1 to c7) show the proportion in the cluster-weighted case compared to the full population without weighting. This ratio represents how each MDD cluster's use of the treatment compares to the overall sample. Values above 1 indicate that the drug was more characteristic in the cluster compared to sample average and values below 1 indicate that they were less characteristic. De Prisco, M., Vieta, E., 2023. The never-ending problem: Sample size matters. European Neuropsychopharmacology: the Journal of the European College of Neuropsychopharmacology 79, 17–18. Oliva, V., Vieta, E., 2025. Predicting the past: The risks and rewards of post-hoc findings. European neuropsychopharmacology: the journal of the European College of Neuropsychopharmacology 92, 21–22.

2017), especially with cardiovascular (e.g. hypertension), pain (e.g. dorsopathies, migraine), and metabolic diseases (e.g. hypothyreosis, diabetes mellitus, hypercholesterinaemia, obesity) (Arnaud et al., 2022; Lo et al., 2024; Steffen et al., 2020), but also a major hub in the diseasome network, underlined by common molecular-level mechanisms (Marx et al., 2017; Menche et al., 2015), and other potential factors, such as life-style (Hullam et al., 2019) and treatment-related effects (Gold et al., 2020). However, the actual shared etiology or the interaction between such comorbid conditions and depression is still poorly understood, and while complex and potentially multidirectional, it is not sufficiently reflected in deliberate treatment choices focusing on the potential etiological link. Therefore, it is noteworthy that our method identified important differences among the clusters with varying MDD burden.

Namely, cluster 6 with high risk of MDD, characterised by mainly stress-related disorders, respiratory tract infections, and an overall high disease burden, pinpointing to mechanisms related to hypothalamo-pituitary-axis (HPA) dysregulation (Iob et al., 2020) and inflammation (Gezsi et al., 2024), showed the highest cumulative drug burden and drug burden of nervous system drugs, across all cohorts even after correcting for the number of multimorbidities. Based on the UK Biobank cohorts, where we had enough power to further investigate participants with MDD diagnosis, this cluster also showed significantly increased prevalence of TRD, and more frequent switches in the treatment path as demonstrated by the drug-tree analysis. Our findings are in line with previous observations that TRD patients more frequently suffer from other psychiatric and somatic comorbidities, and also in turn TRD increases the risk of developing physical disorders (McIntyre et al., 2023). Despite this fact, pharmacogenomic studies aiming to support treatment selection so far mostly excluded subjects with multimorbidity, and even if included them, have not investigated life stressors or comorbid physical conditions as factors for improved prediction models (Minelli et al., 2022). Consequently, our identified clusters, which can be considered more homogenous subtypes of depression, may complement such observations by providing a deeper understanding of shared genetic and etiological factors underlying such comorbid and multimorbid conditions and help pave the way towards a more personalised medication choice to exploit the potentials of treatment to treat both depression and its comorbidities more effectively.

Indeed, we dissected high MDD-risk patients into three clusters. Cluster 5 showed similarities to cluster 6 in terms of high multimorbidity, polypharmacy, drug burden, high TRD prevalence, and frequent antidepressant changes in the treatment path. However, cluster 5 was distinguished from cluster 6 by earlier MDD onset and increased prevalence of schizophrenia and pain disorders (dorsalgia, intervertebral disc disorders, female genital organ-related pain). In addition, this cluster showed the highest TCA drug burden along with broad synaptic treatment targets. Previous studies demonstrated that despite the extremely high symptomatic heterogeneity of depression (Fried and Nesse, 2015) simply focusing on symptom profiles are not enough for MDD subtyping without considering other psychiatric conditions, psychosocial features and age at onset (Nguyen et al., 2022; 2023; Wardenaar et al., 2014). And despite these subtypes showed different severities and genetic heritabilities their implementation in clinical practice remained challenging (Lynall and McIntosh, 2023). Thus, our dissection of cluster 5 based on MDD-related multimorbidity trajectories is noteworthy because of the distinct pharmacological profile. Although there is no clear treatment guideline for TCA use in MDD treatment across different countries and healthcare systems (Vos et al., 2021), our results suggested that cluster 5 (and to certain extent cluster 6) increase the prevalence of starting on or switching to TCA during the treatment path. While most guidelines focus on when not to use TCA (e.g. due to cardiotoxicity), we could not exclude that our results partially reflected the fact that more severe MDD patients are treated in psychiatric settings, where greater variety of antidepressants are used compared to primary care or non-psychiatric clinicians (Rathnam et al., 2024).

Our third cluster of increased MDD prevalence, cluster 7, included higher risk for inflammatory disorders, where allergens (e.g. asthma, allergic rhinitis, dermatitis), infections (e.g. acute tonsillitis), or nervous sensitisation (e.g. migraine) play etiopathophysiological role. Most notably, this cluster had the strongest inflammation-related genetic predisposition (Gezsi et al., 2024) that may contribute to the development of early-onset MDD (Tylee et al., 2022). In line with this, cluster 7 showed increased cumulative polypharmacy and drug burden across cohorts, although in a smaller degree than clusters 6 and 5, but specific antidepressant usage pattern could not be demonstrated probably due to lower power.

Regarding the relatively healthier clusters, cluster 3 has a unique profile with sharp increase in multimorbidity in later life resulting in an increased drug burden either before (CHSS, THL) or after (UKB $sample_{3\text{-}year}$ and UKB $sample_{full}$) correction for multimorbidities, depending on the available treatment data. This cluster was associated with increased risk of kidney diseases, hypertension, and stroke, but with an average-risk MDD that showed later onset (Gezsi et al., 2024). Interestingly, in cluster 4, where besides hypertension we can see increased risk for hypothyroidism and hyperlipidaemias (Gezsi et al., 2024), late life multimorbidity and MDD risk remained average with more divergent drug burden profile across cohorts. Furthermore, focusing on the UK Biobank cohorts antidepressant drug burden pattern of clusters 3–4 was similar to clusters 1–2, supporting previous findings where first choice of antidepressant treatment of MDD patients with or without hypertension was not significantly different (Fugger et al., 2019). More precisely, clusters 1–2 showed an increased prevalence to be treated by SSRIs, the most frequently used first choice antidepressants, and required less switches to other antidepressants throughout the treatment path. This is in line with previous observations where late-onset patients more frequently experienced single major depressive episodes, and lower current suicidal risk (Bartova et al., 2024). A potential explanation for these observations is that clusters 1–2 were characterised by protective genetic profile against diseases caused by immune-related biological processes in our previous study (Gezsi et al., 2024), and thus showed the lowest level of multimorbidity, polypharmacy, and drug burden in the present study, with the smallest prevalence of TRD.

To summarise our main findings, we replicated the importance of immunometabolic depression (Milaneschi et al., 2020) with more detailed differences between clusters that showed important distinctions regarding treatment choices. All three clusters with increased MDD risk showed association, to some extent, with inflammatory processes but with different ethiopathologic backgrounds consisting of features such as infections (clusters 6–7), behavioural and lifestyle factors (clusters 5–6), and genetic predisposition (cluster 5 and 7). Thus, even within inflammation-related depression important subtypes exist (Ioannou et al., 2021), and as we demonstrated, with different therapeutic implications and outcomes that our method was able to entangle. This could also explain that while antidepressant treatment decreased inflammatory markers in moderate-severe depression, these markers alone were not able to predict long-term disease course (Kofod et al., 2022). Our results also supported that metabolic alterations and cardiovascular diseases not necessarily increase MDD risk substantially, as we have seen in clusters 3–4, reflecting the results of genetic studies, which showed moderate genetic overlap between MDD and cardiovascular diseases but causal genetic liability only from MDD to cardiovascular diseases but not vica versa (Bergstedt et al., 2024). The different genetic vulnerability according to clusters, and the variability of the mediating metabolic and lifestyle factors might explain mixed results of RCTs investigating how depression treatment impacts cardiovascular comorbidities (Arnaud et al., 2023).

Considering the state of precision approaches to treatment, psychiatry significantly lags behind other disciplines in medicine. Although, some healthcare guidelines (Lam et al., 2016) already started to embrace understanding of comorbidities on the clinical level, most of them approach this issue from the direction of the somatic illness rather than

depression. Despite that currently used antidepressants have low or moderate effect size compared to placebo (Gold et al., 2020; Simon et al., 2024) even the new guideline by the European Medicines Agency (Butlen-Ducuing et al., 2023) for research and drug development for the treatment of depression does not go beyond emphasising targeting symptom clusters rather than diagnostic categories. As a novel way, the multimorbidity clusters we used in this study have the potential clinical utility in biomarker development as they reflect different etiological pathways involved in the constellation of parallel conditions. For instance, clusters 5 and 6—characterized by elevated risk for treatment-resistant depression (TRD)—exhibit distinct antidepressant prescription patterns and increased polypharmacy. These observations suggest that such clusters could be used to predict differential responses to specific antidepressants, ultimately guiding pharmacological choices. Thus, our study shows the potential to prioritize novel antidepressants based on distinct MDD-related multimorbidity clusters as each cluster provides insights into biological pathways for targeted interventions. In addition, if validated in future prospective studies, this method might contribute to the development of clinical decision-support systems enhancing personalised care and potentially improving patient outcomes in depression.

5. Limitations

Certain limitations of our study should be acknowledged. The availability of drug coverage varies across cohorts, with the UK Biobank (UKB) spanning from 1990 to 2016, resulting in incomplete data for earlier periods and potentially hindering our ability to fully capture long-term trends. Similarly, CHSS and THL cohorts offer shorter and more recent timespans, introducing variability in the temporal resolution of medication data.

Differences in healthcare systems further complicate direct cross-cohort comparisons. Changes in healthcare policies, diagnostic criteria, and prescription practices over time can significantly impact our findings. Moreover, the reliance on prescription and dispensation records presents inherent challenges, as they do not necessarily reflect actual medication adherence or usage, introducing a potential gap between prescribed treatments and real-world practices. The THL dataset, in particular, is constrained to ATC level 3 codes, limiting the granularity of drug analyses.

Our analyses lacked direct clinical measures of depression severity, limiting our ability to fully capture clinical nuances and the potential heterogeneity within treatment groups. Although some questionaire based depression severity metrics were available in the UK Biobank, their coverage was incomplete, not of clinical quality, separate in time from the drug data, and not applicable to other cohorts, further constraining the interpretability and comparability of our findings across datasets.

Furthermore, our definition of TRD although aligns with current research standards constrained by data availability in biobanks, alternative TRD definitions representing other clinical approaches, such as augmentation strategies, could potentially modify the magnitude of risk of TRD in different MDD-related multimorbidity clusters. Nevertheless, we would not expect a different pattern as cluster 5 and 6 have the highest MDD risk, and showed high cumulative drug burden and drug burden of nervous system drugs.

Additionally, while our analyses were planned prospectively and leveraged multiple independent datasets across different cohorts and timeframes, and proper statistical corrections for multiple hypothesis testing, we acknowledge the need of registration policies, such as proposed in (Oliva and Vieta, 2025) and future prospective studies explicitly designed for validating these multimorbidity clusters and their associated treatment patterns.

Lastly, statistical power in certain clusters and subsamples, such as Cluster 7 or the TRD samples, may be limited due to the smaller number of participants with high posterior probabilities. This limitation could impair our ability to detect subtle treatment patterns. Moreover, the asymmetrical distribution of cluster posterior probabilities may lead to uneven statistical power across clusters, potentially influencing the distribution of significant results (De Prisco and Vieta, 2023).

Contributors

Tamas Nagy contributed to the conceptualization of the study, performed data analysis, and drafted the manuscript. Gabriella Juhasz took major part in conceptualization, supervision, and writing. Andras Gezsi, Dóra Török, Zsofia Gal, Isaac Cano, Mikko Kuokkanen, Carsten O. Schmidt, Sandra Van der Auwera, Josep Roca, Peter Petschner, and Peter Antal contributed to the conceptualization of the study. Xenia Gonda, Nora Eszlari, Gabor Hullam, Peter Petschner, Peter Antal, and Gabriella Juhasz were involved in the writing and revision of the manuscript. Rubèn González-Colom, Hannu Mäkinen, and Teemu Paajanen conducted auxilary calculations for the study.

All authors contributed to and have approved the final manuscript.

Role of funding source

This study was supported by the Hungarian National Research, Development, and Innovation Office 2019–2.1.7-ERA-NET-2020–00005 under the frame of ERA PerMed (ERAPERMED2019–108); by the Hungarian National Research, Development, and Innovation Office OTKA K 143391, K 139330, PD 146014, and PD 134449 grants; by the Hungarian Brain Research Program 3.0 (NAP2022-I-4/2022); and by TKP2021-EGA-25 and TKP2021-EGA-02, supported by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, under the TKP2021-EGA funding scheme; and the European Union project [RRF-2.3.1–21–2022–00004] within the framework of the Artificial Intelligence National Laboratory. Dora Torok is supported by EKÖP-2024–68. Nora Eszlari is supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

The study was funded by the Academy of Finland under the frame of ERA PerMed (TRAJECTOME project, ERAPERMED2019–108). The Catalan cohort was extracted from the Catalan Health Surveillance System database, owned and managed by the Catalan Health Service, with the earnest collaboration of the Digitalization for the Sustainability of the Healthcare (DS3) - IDIBELL group. The study was supported by the Catalan Department of Health (SLD002/19/000002) under the frame of ERA PerMed (ERAPERMED2019–108).

The funders had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Declaration of competing interest

The authors declare no conflict of interest

Acknowledgement

The authors wish to acknowledge the participants and investigators of the Finrisk (1992, 1997, 2002, 2007, 2012), Health 2000/2011 and FinHealth 2017 studies. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.euroneuro.2025.05.007.

References

Arnaud, A.M., Brister, T.S., Duckworth, K., Foxworth, P., Fulwider, T., Suthoff, E.D., Werneburg, B., Aleksanderek, I., Reinhart, M.L., 2022. Impact of major depressive disorder on comorbidities: a systematic literature review. J. Clin. Psychiatry 83 (6), 21r14328.

Arnaud, A.M., Brister, T.S., Duckworth, K., Foxworth, P., Fulwider, T., Suthoff, E.D., Werneburg, B., Aleksanderek, I., Reinhart, M.L., 2023. Impact of treating depression

on associated comorbidities: a systematic literature review. Prim, Care Companion CNS Disord 25 (1), 44925.
Barlati, S., Minelli, A., Nibbio, G., Bertoni, L., Necchini, N., Paolini, S., Muscarella, A., Ubertino, U.B., Calzavara-Pinton, I., Vita, A., Gennarelli, M., 2023. The role of pharmacogenetics in the treatment of major depressive disorder: a critical review. Front Psychiatry 14, 1307473.
Bartova, L., Fugger, G., Dold, M., Kautzky, A., Bairhuber, I., Kloimstein, P., Fanelli, G., Zanardi, R., Weidenauer, A., Rujescu, D., Souery, D., Mendlewicz, J., Zohar, J., Montgomery, S., Fabbri, C., Serretti, A., Kasper, S., 2024. The clinical perspective on late-onset depression in European real-world treatment settings. Eur. Neuropsychopharmacol. 84, 59–68.
Bergstedt, J., Pasman, J.A., Ma, Z., Harder, A., Yao, S., Parker, N., Treur, J.L., Smit, D.J. A., Frei, O., Shadrin, A.A., Meijsen, J.J., Shen, Q., Hägg, S., Tornvall, P., Buil, A., Werge, T., Hjerling-Leffler, J., Als, T.D., Børglum, A.D., Fang, F., 2024. Distinct biological signature and modifiable risk factors underlie the comorbidity between major depressive disorder and cardiovascular disease. Nature Cardiovasc. Res. 3 (6), 754–769.
Bonk, S., Eszlari, N., Kirchner, K., Gezsi, A., Garvert, L., Kuokkanen, M., Cano, I., Grabe, H.J., Antal, P., Juhasz, G., Van der Auwera, S., 2024. Impact of gene-by-trauma interaction in MDD-related multimorbidity clusters. J. Affect. Disord. 359, 382–391.
Butlen-Ducuing, F., Haberkamp, M., Aislaitner, G., Bałkowiec-Iskra, E., Mattila, T., Doucet, M., Kollb-Sielecka, M., Balabanov, P., Leuchs, A.-K., Elferink, A., 2023. The new European Medicines Agency guideline on antidepressants: a guide for researchers and drug developers. Eur. Psychiatry 67 (1), e2.
Crichton, W. (2023). *Wordtree* [Python]. https://github.com/willcrichton/wordtree (Original work published 2020).
De Prisco, M., Vieta, E., 2023. The never-ending problem: sample size matters. Eur. Neuropsychopharmacol. 79, 17–18.
Dimond, B., 2003. The British National Formulary and other medicine indexes. Brit. J. Nursing 12 (10), 620–622.
Dwyer, D.B., Falkai, P., Koutsouleris, N., 2018. Machine learning approaches for clinical psychology and psychiatry. Ann. Rev. Clin. Psychology 14, 91–118.
El-Sappagh, S., Franda, F., Ali, F., Kwak, K.S., 2018. SNOMED CT standard ontology based on the ontology for general medical science. BMC medical informatics and decision making 18 (1), 76. https://doi.org/10.1186/s12911-018-0651-5.
Fabbri, C., Hagenaars, S.P., John, C., Williams, A.T., Shrine, N., Moles, L., Hanscombe, K. B., Serretti, A., Shepherd, D.J., Free, R.C., Wain, L.V., Tobin, M.D., Lewis, C.M., 2021. Genetic and clinical characteristics of treatment-resistant depression using primary care records in two UK cohorts. Mol. Psychiatry 26 (7), 3363–3373.
First, M.B., 2013. Diagnostic and statistical manual of mental disorders, 5th edition, and clinical utility. J. Nerv. Ment. Dis. 201 (9), 727–729.
Fried, E.I., Nesse, R.M., 2015. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. J. Affect. Disord. 172, 96–102.
Fugger, G., Dold, M., Bartova, L., Kautzky, A., Souery, D., Mendlewicz, J., Serretti, A., Zohar, J., Montgomery, S., Frey, R., Kasper, S., 2019. Comorbid hypertension in patients with major depressive disorder—Results from a European multicenter study. Eur. Neuropsychopharmacol. 29 (6), 777–785.
Gezsi, A., Van der Auwera, S., Mäkinen, H., Eszlari, N., Hullam, G., Nagy, T., Bonk, S., González-Colom, R., Gonda, X., Garvert, L., Paajanen, T., Gal, Z., Kirchner, K., Millinghoffer, A., Schmidt, C.O., Bolgar, B., Roca, J., Cano, I., Kuokkanen, M., Juhasz, G., 2024. Unique genetic and risk-factor profiles in clusters of major depressive disorder-related multimorbidity trajectories. Nat. Comm. 15 (1), 7190.
Gold, S.M., Köhler-Forsberg, O., Moss-Morris, R., Mehnert, A., Miranda, J.J., Bullinger, M., Steptoe, A., Whooley, M.A., Otte, C., 2020. Comorbid depression in medical diseases. Nat. Rev. Dis. Primers 6 (1), 69.
González-Colom, R., Mitra, K., Vela, E., Gezsi, A., Paajanen, T., Gál, Z., Hullam, G., Mäkinen, H., Nagy, T., Kuokkanen, M., Piera-Jiménez, J., Roca, J., Antal, P., Juhasz, G., Cano, I., 2024. Multicentric assessment of a multimorbidity-adjusted disability score to stratify depression-related risks using temporal disease maps: instrument validation study. J. Med. Internet Res. 26 (1), e53162.
Hullam, G., Antal, P., Petschner, P., Gonda, X., Bagdy, G., Deakin, B., Juhasz, G., 2019. The UKB envirome of depression: from interactions to synergistic effects. Sci. Rep. 9 (1), 9723.
Hullam, G., Gal, Z., Gonda, X., Nagy, T., Gezsi, A., Cano, I., Van der Auwera, S., Koukkanen, M., Antal, P., Juhasz, G., 2023. A sound mind in a sound body: a novel concept unravelling heterogeneity of depression. Neuropsychpharm. Hung. 25 (4), 183–193.
Ioannou, M., Foiselle, M., Mallet, J., Stam, E.L., Godin, O., Dubertret, C., Terro, E., Sommer, I.E.C., Haarman, B.C.M., Leboyer, M., Schoevers, R.A., 2021. Towards precision medicine: what are the stratification hypotheses to identify homogeneous inflammatory subgroups. Eur Neuropsychopharmacology 45, 108–121.
Iob, E., Kirschbaum, C., Steptoe, A., 2020. Persistent depressive symptoms, HPA-axis hyperactivity, and inflammation: the role of cognitive-affective and somatic symptoms. Mol. Psychiatry 25 (5), 1130–1140.
Kendrick, T., Pilling, S., Mavranezouli, I., Megnin-Viggars, O., Ruane, C., Eadon, H., Kapur, N., Committee, Guideline, 2022. Management of depression in adults: summary of updated NICE guidance. BMJ 378, o1557.
Kishi, T., Ikuta, T., Sakuma, K., Okuya, M., Hatano, M., Matsuda, Y., Iwata, N., 2023. Antidepressants for the treatment of adults with major depressive disorder in the maintenance phase: a systematic review and network meta-analysis. Mol. Psychiatry 28 (1), 402–409.
Kofod, J., Elfving, B., Nielsen, E.H., Mors, O., Köhler-Forsberg, O., 2022. Depression and inflammation: correlation between changes in inflammatory markers with antidepressant response and long-term prognosis. Eur Neuropsychopharmacology 54, 116–125.
Lam, R.W., Kennedy, S.H., Adams, C., Bahji, A., Beaulieu, S., Bhat, V., Blier, P., Blumberger, D.M., Brietzke, E., Chakrabarty, T., Do, A., Frey, B.N., Giacobbe, P., Gratzer, D., Grigoriadis, S., Habert, J., Ishrat Husain, M., Ismail, Z., McGirr, A., Milev, R.V., 2024. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2023 update on Clinical Guidelines for management of major depressive disorder in adults: réseau canadien pour les traitements de l'humeur et de l'anxiété (CANMAT) 2023 : mise à jour des lignes directrices cliniques pour la prise en charge du trouble dépressif majeur chez les adultes. Can. J. Psychiatry 69 (9), 641–687.
Lam, R.W., McIntosh, D., Wang, J., Enns, M.W., Kolivakis, T., Michalak, E.E., Sareen, J., Song, W.Y., Kennedy, S.H., MacQueen, G.M., Milev, R.V., Parikh, S.V., Ravindran, A. V., CANMAT Depression Work Group, 2016. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 1. Disease Burden and Principles of Care. Canadian journal of psychiatry. Revue canadienne de psychiatrie 61 (9), 510–523. https://doi.org/10.1177/0706743716659416.
Lo, H.K.Y., Chan, J.K.N., Wong, C.S.M., Chung, K.F., Correll, C.U., Solmi, M., Baum, L.W., Thach, T.Q., Sham, P.C., Chang, W.C., 2024. Excess mortality and life-years lost in people diagnosed with depression: a 20-year population-based cohort study of 126,573 depressed individuals followed for 1139,073 persons-years. Eur. Neuropsychopharmacology 91, 1–8.
Lynall, M.-E., McIntosh, A.M., 2023. The heterogeneity of depression. Am. J. Psychiatry 180 (10), 703–704.
Malhi, G.S., Bell, E., Bassett, D., Boyce, P., Bryant, R., Hazell, P., Hopwood, M., Lyndon, B., Mulder, R., Porter, R., Singh, A.B., Murray, G., 2021. The 2020 Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Austr. NZ. J. Psychiatry 55 (1), 7–117.
Malhi, G.S., Mann, J.J., 2018. Depression. Lancet 392 (10161), 2299–2312.
Marx, P., Antal, P., Bolgar, B., Bagdy, G., Deakin, B., Juhasz, G., 2017. Comorbidities in the diseasome are more apparent than real: what Bayesian filtering reveals about the comorbidities of depression. PLoS Comput. Biol. 13 (6), e1005487.
Masnoon, N., Shakib, S., Kalisch-Ellett, L., Caughey, G.E., 2017. What is polypharmacy? A systematic review of definitions. BMC Geriatr 17 (1), 230.
McIntyre, R.S., Alsuwaidan, M., Baune, B.T., Berk, M., Demyttenaere, K., Goldberg, J.F., Gorwood, P., Ho, R., Kasper, S., Kennedy, S.H., Ly-Uson, J., Mansur, R.B., McAllister-Williams, R.H., Murrough, J.W., Nemeroff, C.B., Nierenberg, A.A., Rosenblat, J.D., Sanacora, G., Schatzberg, A.F., Maj, M., 2023. Treatment-resistant depression: definition, prevalence, detection, management, and investigational interventions. World Psychiatry 22 (3), 394–412.
Menche, J., Sharma, A., Kitsak, M., Ghiassian, S.D., Vidal, M., Loscalzo, J., Barabási, A.-L., 2015. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. Science 347 (6224), 1257601.
Milaneschi, Y., Lamers, F., Berk, M., Penninx, B.W.J.H., 2020. Depression heterogeneity and its biological underpinnings: toward immunometabolic Depression. Biol. Psychiatry 88 (5), 369–380.
Minelli, A., Barlati, S., Baune, B.T., European College of Neuropsychopharmacology (ECNP) Pharmacogenomics & Transcriptomics Network, 2022. Evaluating study designs and treatment outcomes of antidepressant pharmacogenetic clinical trials—Challenges and future perspectives. A critical review. Eur. Neuropsychopharmacol. 59, 68–81.
Nguyen, T.-D., Harder, A., Xiong, Y., Kowalec, K., Hägg, S., Cai, N., Kuja-Halkola, R., Dalman, C., Sullivan, P.F., Lu, Y., 2022. Genetic heterogeneity and subtypes of major depression. Mol. Psychiatry 27 (3), 1667–1675.
Nguyen, T.-D., Kowalec, K., Pasman, J., Larsson, H., Lichtenstein, P., Dalman, C., Sullivan, P.F., Kuja-Halkola, R., Lu, Y., 2023. Genetic contribution to the heterogeneity of major depressive disorder: evidence from a sibling-based design using Swedish national registers. Am. J. Psychiatry 180 (10), 714–722.
Oliva, V., Vieta, E., 2025. Predicting the past: the risks and rewards of post-hoc findings. Eur. Neuropsychopharmacology (92), 21–22.
Perlis, R.H., 2016. Abandoning personalization to get to precision in the pharmacotherapy of depression. World Psychiatry 15 (3), 228–235.
Rathnam, S., Hart, K.L., Sharma, A., Verhaak, P.F., McCoy, T.H., Doshi-Velez, F., Perlis, R.H., 2024. Heterogeneity in antidepressant treatment and major depressive disorder outcomes among clinicians. JAMA Psychiatry 81 (10), 1003–1009.
Read, J.R., Sharpe, L., Modini, M., Dear, B.F., 2017. Multimorbidity and depression: a systematic review and meta-analysis. J. Affect. Disord. 221, 36–46.
Simon, G.E., Moise, N., Mohr, D.C., 2024. Management of depression in adults: a review. JAMA 332 (2), 141–152.
Steffen, A., Nübel, J., Jacobi, F., Bätzing, J., Holstiege, J., 2020. Mental and somatic comorbidity of depression: a comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. BMC Psychiatry 20 (1), 142.
Tylee, D.S., Lee, Y.K., Wendt, F.R., Pathak, G.A., Levey, D.F., De Angelis, F., Gelernter, J., Polimanti, R., 2022. An atlas of genetic correlations and genetically informed associations linking psychiatric and immune-related phenotypes. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2022.0914.
UK Biobank. (2019, September). *Primary Care Linked Data*. https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/primary_care_data.pdf.
Vos, C.F., Aarnoutse, R.E., Op de Coul, M.J.M., Spijker, J., Groothedde-Kuyvenhoven, M. M., Mihaescu, R., Wessels-Basten, S.J.W., Rovers, J.J.E., ter Hark, S.E., Schene, A.H., Hulscher, M.E.J.L., & Janzing, J.G.E. (2021). Tricyclic antidepressants for major

depressive disorder: a comprehensive evaluation of current practice in The Netherlands. *BMC Psychiatry* 21.

Wardenaar, K.J., van Loo, H.M., Cai, T., Fava, M., Gruber, M.J., Li, J., de Jonge, P., Nierenberg, A.A., Petukhova, M.V., Rose, S., Sampson, N.A., Schoevers, R.A., Wilcox, M.A., Alonso, J., Bromet, E.J., Bunting, B., Florescu, S.E., Fukao, A., Gureje, O., Kessler, R.C., 2014. The effects of co-morbidity in defining major depression subtypes associated with long-term course and severity. Psychol. Med. 44 (15), 3289–3302. https://doi.org/10.1017/S0033291714000993.

World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems : tenth revision*. World Health Organization. https://iris.who.int/handle/10665/42980.