

Review Article

Cite this article: Davis, K. A. S., Mirza, L., Clark, S. R., Coleman, J. R. I., Kassam, A. S., Mills, N. T., Zadow, A., McIntosh, A. M., & Hotopf, M. (2025). Unlocking mental health insights with UK Biobank data: Past use and future opportunities. *Psychological Medicine*, **55**, e244, 1–8
<https://doi.org/10.1017/S0033291725101359>

Received: 02 July 2025

Revised: 02 July 2025

Accepted: 12 July 2025

Keywords:

cohort study; epidemiology; mental health; methodology; research; UK Biobank






Corresponding author:

Katrina A. S. Davis;

Email: katrina.davis@kcl.ac.uk

K.A.S.D. and L.M. are joint first authors.

Unlocking mental health insights with UK Biobank data: Past use and future opportunities

Katrina A. S. Davis^{1,4} , Luwaiza Mirza², Scott R. Clark³, Jonathan R. I. Coleman¹ , Aliyah S. Kassam⁴ , Natalie T. Mills³, Amy Zadow⁵, Andrew M. McIntosh⁶  and Matthew Hotopf^{1,4} 

¹King's College London Institute of Psychiatry, Psychology and Neuroscience, London, England, UK; ²West London NHS Trust, Mental Health Services, London, England, UK; ³University of Adelaide, Discipline of Psychiatry, South Australia, Australia; ⁴South London and Maudsley NHS Trust, Biomedical Research Centre, London, England, UK; ⁵University of Adelaide, Faculty of Health and Medical Sciences, South Australia, Australia and ⁶University of Edinburgh, Division of Psychiatry, Edinburgh, Scotland, UK

Abstract

UK Biobank (UKB) is a large-scale, prospective resource offering significant opportunities for mental health research. Data include genetic and biological data, healthcare linkage, and mental health enhancements. Challenges arise from incomplete linkage of some sources and the incomplete coverage for enhancements, which also occur at different times post-baseline. We searched for publications using UKB for mental health research from 2016 to 2023 to describe and inspire future use. Papers were classified by mental health topic, 'additional' aspects, and the data used to define the mental health topic. We identified 480 papers, with 338 focusing on mental health disorder topics (affective, anxiety, psychotic, multiple, and transdiagnostic). The most commonly studied disorder was depression (41%). The most common single method of ascertaining mental disorder status was the Mental Health Questionnaire (26%), with genetic risk, for example, using polygenic risk scores, also frequent (21%). Common additional aspects included brain imaging, gene–environment interaction, and the relationship with physical health. The review demonstrates the value of UKB to mental health research. We explore the strengths and weaknesses, producing resources informed by the review. A strength is the flexibility: conventional epidemiological studies are present, but also genomics, imaging, and other tools for understanding mental health. A major weakness is selection effects. UKB continues to hold potential, especially with additional data continuing to become available.

Background

Mental disorders are common, costly, and burdensome (GBD Mental Health Collaborators, 2022). Understanding the interplay between genetic predisposition, lifestyle and environmental factors in mental health disorders are crucial for developing effective preventative, diagnostic, and management approaches (Wykes et al., 2023). Such research requires large sample sizes, such as population-based cohorts. The UK Biobank (UKB) is a large-scale, prospective cohort with extensive data for a sample of over 500,000 individuals aged between 40 and 69 at recruitment (2006–2010), providing a valuable resource for researchers (Sudlow et al., 2015). It includes questionnaire data, biological samples, and genetic profiles, supplemented by regularly updated linkage to National Health Service (NHS) health records. This allows clinical diagnoses to be used as risk factors and outcomes in longitudinal studies (Allen et al., 2024).

UKB has undergone several 'enhancements', involving additional participation from volunteers since the baseline assessment. These include activity monitoring, multimodal imaging, and online questionnaires. Two mental health questionnaires, MHQ1 (primarily completed in 2016) and MHQ2 (primarily completed in 2022), were introduced to assess lifetime common mental disorders (Davis et al., 2020, 2025). Intended primarily to provide phenotypes of lifetime common mental disorders, these questionnaires include measures on depression, mania, generalized anxiety, post-traumatic stress disorder (PTSD), and alcohol use disorder, with panic disorder and eating disorders added in MHQ2. Additionally, transdiagnostic features were assessed, such as psychotic experiences, self-harm, quality of life, and resilience. The questionnaires also collected information on known mental health risk factors, including childhood and adult adverse events, cannabis use, loneliness, social isolation, and COVID-19 exposure in MHQ2. Figure 1 presents a timeline of UKB's resource accumulation before and after baseline, while Table 1 outlines available data sources for key mental disorder topics.

As demonstrated by Figure 1 and Table 1, data on mental health and mental disorders in UKB come from multiple sources, including record linkages, baseline data collection, and

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

UK Biobank Timeline

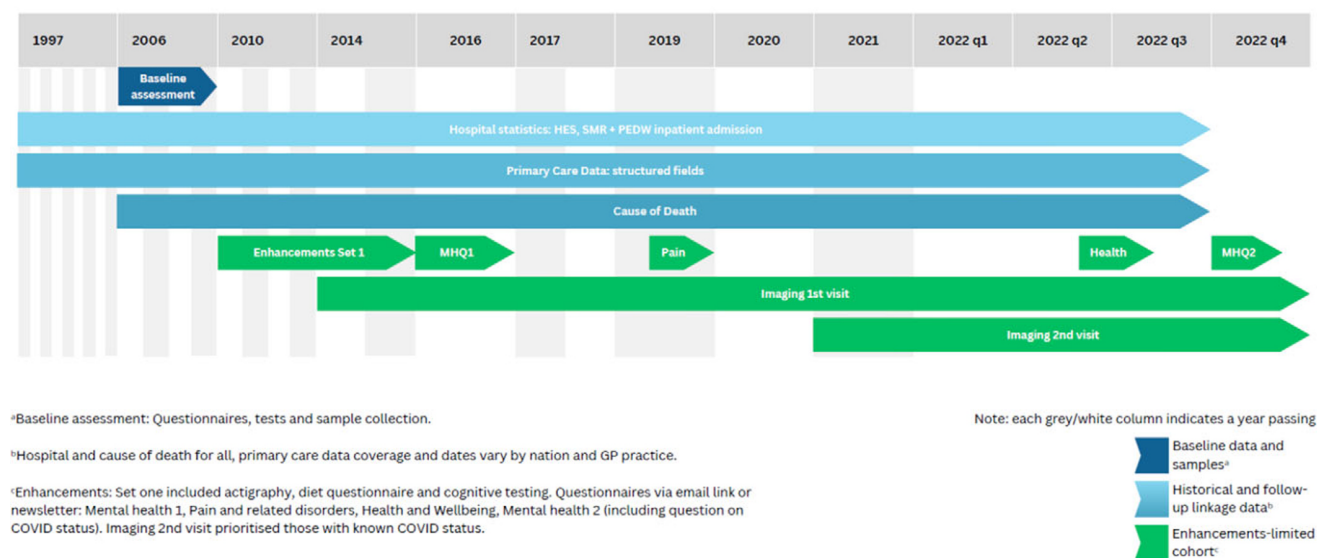


Figure 1. UK Biobank timeline showing timing (but not completeness) of enhancements.

enhancements requiring a repeat participant contact, like imaging and questionnaires. Navigating this rich yet complex data landscape is challenging, and no single resource comprehensively describes the breadth of available data. This article aimed to address this gap.

We aimed to explore UKB in a way that will be valuable for anyone planning to use UKB mental health data and those who wish to know more to evaluate UKB research. The next step is to provide an overview of the research conducted using UKB 2016–2023, to discuss the data that have been used, trends and considerations. It can also serve as a resource for future researchers to identify and build upon previous findings. We will explore UKB's strengths and limitations as an observational data source and signpost to further advice.

Methods

Selection and extraction

We conducted a review of published research utilizing UKB to investigate mental health themes. The UKB website, which records published papers that have been submitted as part of the data access agreement, served as a primary source. However, since the website was not updated between November 2022 and December 2023, we supplemented our search for 2022–2023 papers using Google Scholar. Specifically, for papers published in 2022–2023, we identified studies citing one of three UKB mental health methodology papers: Smith *et al.* (2013), Davis *et al.* (2019, 2020).

Papers were selected if they had a primary focus on mental health or mental disorders, as suggested by the title, and utilized UKB data or results in some form for primary research (i.e. not reviews). Selection was conducted by one researcher (L.M.), with another (K.A.S.D.) providing guidance on uncertainty and indexing. Basic characteristics of the papers were summarized using a standardized template by L.M. and A.S.K.

Classification

Four researchers were responsible for checking and categorizing the papers (K.A.S.D., A.Z., N.T.M., and S.R.C.), with paper allocation

based on individual expertise where possible. The guidelines used for categorization, iteratively produced by the researchers, are shown in [Supplementary Tables S1–S3](#). The first divide was into those papers focused on mental health disorder topics (MD papers) and those addressing broader subjects. The MD papers included specific disorders, compound mental disorders and transdiagnostic features. A compound mental disorder category was applied in cases where broad mental health measures were used. For example, Batty, Deary *et al.*'s (2020) study which hypothesized a link between mental health disorders and COVID-19 hospitalization, categorized participants based on a baseline question about whether they had ever seen a psychiatrist.

For MD papers, we documented the data used to define the disorder or trait, and the secondary topic, which were 'general/methodological', 'environmental factors', 'biomarkers', 'genetics', 'physical health', or 'other'. Papers with the secondary topic of inflammatory biomarkers were classified under biomarker research rather than physical health. Papers using genetic methods to explore relationships between mental health and another factor were classified by the factor under investigation rather than as genetic. For example, Coleman *et al.*'s (2020) study used genetics to examine depression as the mental health topic and trauma as the secondary (environmental) factor. Those papers that were categorized under physical health were further divided into 'cardiovascular and metabolic', 'pain and sensory', 'inflammatory', 'infectious', 'general health', or 'other' subcategories.

Analysis

To identify patterns, we developed a matrix summarizing MD papers by their primary mental health topic, secondary topic and the UKB data types used. In [Tables 2 and 3](#) and [Figure 2](#) mental disorder topics are rows.

In [Table 2](#) (and more detailed counterpart [Supplementary Table S16](#)), columns represent secondary topics. To assign papers to columns in [Table 2](#), we allocated a single secondary topic, prioritizing physical health, environmental factors, biomarkers, and genetics, in that order. This approach maximized sensitivity

Table 1. Mental disorder topics covered in data sources within the UKB

Source	Baseline	Enhanced baseline	Linkage	Imaging questionnaire	MHQ1	MHQ2	Other online follow-up
Mental disorder (nonspecific)	x (HS)		x		x		o Pain, HWB
Affective							
Depression	ox	x	x	ox	ox	ox	o Cognition, sleep; x digestive
Mania/bipolar	x (SR)	x	x	x	x	x	
Anxiety							
GAD or nonsp anxiety	x (SR)		x		ox	o	o Sleep; x digestive
PTSD	x (SR)		x		o	x (pSR)	
Psychotic							
Schizophrenia	x (SR)		x		x (pSR)	x (pSR)	
Other sp							
Substance/ alcohol UD	x (SR)		x		ox	o (AUD)	
Eating disorder	x (SR)		x		x (pSR)	x	
ADHD/ASD	x (SR)		x		x (pSR)	x (pSR)	
Transdiagnostic							
Self-harm			x		x	x	
Psychotic experiences					x		
Anhedonia	o	ox		o	ox	ox	
Mood instability	x	x		x			
Other notable questions	SR regular medications Neuroticism Social factors	Baseline questions also asked		Repeat of baseline	QoL, adverse experiences	QoL, social factors, COVID exposure	HWB uses same health-related quality of life as MHQ2

Abbreviations: ADHD, attention-deficit and hyperactivity disorder; ASD, autism spectrum disorder; AUD, alcohol use disorder; HS, help-seeking question 'Have you ever seen a general practitioner (GP) for nerves, anxiety, tension or depression?'; HWB, Health and Wellbeing questionnaire; o, current; QoL, quality of life; SR, self-report (open question); pSR, prompted self-report (yes/no); SUD, substance use disorder; x, lifetime.

for studies on physical comorbidities while maintaining specificity for purely genetic aspects of mental health disorders. For example, the Coleman et al.'s (2020) study is positioned at the intersection of depression and environment in the matrix. It defined depression using MHQ1 lifetime disorder status, classifying its data type as MHQ for [Supplementary Table S16](#) and [Figure 2](#). The paper is also summarized in the relevant [Supplementary Table S4](#).

[Figure 2](#) and [Supplementary Table S16](#) details the data sources used to define mental health topics, again assigning one value per paper, with papers using more than one data source described under 'multiple'. In [Table 3](#), we applied the same method for studies exploring comorbidities, linking mental health disorder topics as rows with columns represent the physical health subcategorization. The use of genetic technique was something identified from relevant papers by researchers, and is summarized for the comorbidity papers in [Table 3](#).

Results

We identified 480 papers within the broad mental health domain published between 2016 and 2023. Of these, 338 were classified under mental health disorders topics (MD papers), summarized in

[Supplementary Materials Appendix 2.1: Supplementary Tables S4–S9](#). The remaining papers, covering broader mental health topics, are listed in [Supplementary Tables S10–S15](#).

MD papers are shown in [Table 2](#), categorized by disorder (rows) and secondary topics (columns). [Supplementary Table 16](#) also shows data used. Depression is revealed as the most studied disorder, with 140 papers examining depression alone, 25 combining it with anxiety, and 14 with bipolar affective disorder. There are at least 25 depression-related papers in each of environmental, biomarkers, genetic, and physical health topic categories. The next most common disorder classification was 'multiple disorders', with 51 papers (66 when including compound mental disorder definitions). Schizophrenia was the second most researched single disorder.

Ninety-two papers from [Table 2](#) had biomarkers as a secondary topic, including 41 related to brain imaging and 19 examining sleep or activity patterns. Seventy-three papers explored environmental factors as a secondary topic, with 27 considering both environmental and genetic influences. Fifty-five papers had genetics as their sole secondary topic, the majority of which (40 of 55) focused on depression alone or in combination with other disorders. Eighteen papers explored secondary topics such as cognition, reproductive traits, and personality traits, categorized as 'other'; 12 focused solely on the mental disorders they studied, categorized as 'general'.

Table 2. Matrix of papers using UK Biobank to research mental health disorder topics classified by mental health topic (rows) and secondary topic (columns) with reference numbers

N. papers ref no.	Environmental	Biomarkers	Genetics	Physical health	General	Other	Total: N. papers
Affective disorders (see also anxiety and psychotic disorders for combined disorders)							
Depression	27 1–27	36 28–63	25 64–88	42 89–131	4 132–135	5 136–140	140
Bipolar	1 141	3 142–144	0	2 145–146	1 147	0	7
Depression and bipolar	1 148	3 149–151	2 152–153	5 154–158	1 159	2 160–161	14
Anxiety and stress-related disorders, including anxiety and/or depression							
Generalized anxiety disorder	0	2 162–163	1 164	1 165	0	0	4
PTSD	3 166–168	0	1 169	1 170	0	1 171	6
Anxiety and depression	9 172–180	9 181–189	4 190–193	12 194–205	2 206–207	1 208	37
Anxiety and PTSD	0	1 209	1 210	0	0	0	2
Psychotic disorder							
Schizophrenia	4 211–214	15 215–229	0	5 230–234	0	2 235–236	26
Bipolar & schizophrenia	2 237–238	2 239–240	0	0	0	0	4
Other specified disorder							
Alcohol or substance use disorder	0	1 AUD ²⁴¹	4 SUD ²⁴² 243, AUD ²⁴⁴ 245	0	1 AUD ²⁴⁶	0	6
ADHD/autism/eating disorder	1 ASD ²⁴⁷	1 ED ²⁴⁸	0	0	0	1 ADHD ²⁴⁹	3
Multiple and combined							
Multiple separate disorders	14 250–263	9 264–272	9 273–281	12 282–293	3 294–296	4 297–300	51
Mental disorders (single category)	4 301–304	5 305–309	0	6 310–315	0	0	15
Transdiagnostic features							
Self-harm	3 316–318	1 319	6 320–325	0	0	2 326–327	12
Psychotic experiences	1 328	1 329	1 330	0	0	0	3
Anhedonia	0	2 331–332	0	0	0	0	2
Trans-diagnostic	2 333–334	2 335–336	1 337	1 338	0	0	6
Total	73	92	55	88	12	18	338

Note: See Supplementary Material, Section 1.2 for bibliography.

Abbreviations: ADHD, attention-deficit and hyperactivity disorder; ASD, autism spectrum disorder; AUD, alcohol use disorder; eBaseline, enhanced baseline; MHQ, first mental health questionnaire; Rpt, repeat; SUD, substance use disorder.

The pattern of data use across the mental health disorder topics illustrated in Figure 2. Eighty-eight of 338 papers (26%) used MHQ1 alone and the next most popular source (72 papers) was genetic risk, broadly defined as any genetic instrument proxy, such as polygenic risk scores (PRS). Sixty-eight used the baseline questions alone, including 30 papers using the universal baseline questions, 26 using enhanced baseline questions, and 12 using a repeat

of the baseline questionnaire items at the time of imaging. Linkage was used as the only data source by 18. Seventy-eight papers used multiple or combined UKB sources.

The pattern of use of sources between mental health disorder topics partly reflects the availability of data, as shown in Table 1, but also other patterns. For instance, for the two most commonly studied of depression and schizophrenia, the use on genetic risk

Table 3. Matrix of UK Biobank published research (papers) that involve both mental health and physical health aspects, sorted into mental health topic and physical health topic, highlighting the role of genetic epidemiological techniques (genetic tx).

N. papers Physical topic, (ref)							
MH dx	Cardiovascular and metabolic dx and related traits	Pain and sensory	Inflammatory	Infectious	General Health	Other	Summary: N. papers
Affective • Depression 42 • Bipolar 2 • Depression + Bipolar 5	21 Body mass index ¹²⁷ CardioM dx ^{93, 113, 125, 156, 158} CVD ^{92, 96, 97, 109, 124, 317} Diabetes ^{111, 114, 122, 145} CardioM traits ^{100, 108, 119, 123, 128}	8 Pain ^{90, 95, 105, 117, 155} Tinnitus ¹⁰⁷ Hearing ^{98, 99}	6 IBD ^{112 131} Inflammatory dx ^{91, 115, 116, 126}	2 Herpes Simplex ¹¹⁸ Severe infection ¹²¹	5 Hospitalization ¹²⁰ Multiple ^{89, 104} Health ^{110, 157}	7 Screening uptake ⁹⁴ Gynecological ¹⁰¹ Sleep and diet ¹³⁰ Sleep apnea ¹⁰³ VTE ¹⁰² Fitness ^{129, 146}	49 17 (35%) used genetic tx
Anxiety+ • Anxiety 1 • PTSD 1 • Anxiety + Depression 12	3 CVD ^{170, 196, 198}	4 Pain ²⁰⁴ Tinnitus ^{194, 205} Vision ¹⁶⁵	2 Asthma and eczema ¹⁹⁵ Eczema and psoriasis ¹⁹⁹	0	1 Multimorbidity ²⁰¹	4 CRP and gut microbiome ²⁰³ Irritable bowel syndrome ¹⁹⁷ Fitness ^{200, 202}	14 3 (21%) used genetic tx
Psychotic dx • Schizophrenia 5	3 CardioM dx ²³¹ CVD ²³³ CardioM traits ²³²	0	0	0	1 Multiple ²³⁴	1 Dental ²³⁰	5 2 (40%) used genetic tx
Multiple and combined • Multiple sp. 12 • Combined 6 • Trans- diagnostic 1	7 CardioM dx ²⁹⁰ CVD ³¹⁵ Congenital CVD ²⁸⁶ Diabetes ³¹¹ Obesity ²⁸⁸ CardioM traits ^{285, 292}	1 Tinnitus ³¹²	2 Asthma ²⁸³ Inflammatory dx ²⁸²	4 COVID- related ^{287, 289, 310, 313}	4 Hospitalization ²⁹¹ Mortality ³³⁸ Health ²⁸⁴ Frailty ²⁹³	1 Gut microbiome ³¹⁴	19 9 (47%) used genetic tx
Summary (N. papers)	34 (35% genetic tx)	13 (31% genetic tx)	10 (50% genetic tx)	6 (0% genetic tx)	11 (36% genetic tx)	13 (46% genetic tx)	87 31 (36%) used genetic tx

Note: Reference numbers refer to bibliography in [supplementary materials](#), Section 1.2.

Abbreviations: CardioM, cardiometabolic; CVD, cardiovascular disorder; dx, disorder(s); IBD, inflammatory bowel disorder; PTSD, post-traumatic stress disorder; tx, technique(s); VTE, venous thromboembolism.

to define the disorder was markedly different: for depression 14% (20/140); for schizophrenia 69% (16/26). Two further schizophrenia papers used UKB as a control cohort for clinical schizophrenia samples. This is likely related to the much lower prevalence of schizophrenia in the cohort.

Eighty-seven studies examined both mental health disorders and a physical health aspect, as shown in the matrix in [Table 3](#). The physical health aspect included both disorders (e.g. diabetes) and traits (e.g. blood pressure). The most common physical health category was cardiometabolic (34 of 87), followed by sensory issues and pain, inflammatory disorders, and infections. General health, included indicators of overall health or ill health, such as mortality. Among mental health disorders, affective disorders – primarily depression – were the most frequently studied, appearing in 49 of the 87 comorbidity papers. Genetically informed techniques were widely used to investigate the relationships between mental and physical health aspects, with 31 studies (36%) employing such methods, as noted in [Table 3](#).

Discussion

Using a combination of the UKB website and citation checking, we identified 480 papers that used UKB to investigate mental health and related topics from inception to the end of 2023. Of these,

338 papers focused on mental disorders, with the remainder addressing areas such as cognition and personality. Depression was by far the most frequently researched disorder, appearing as the sole focus in 140 papers and commonly included in studies covering multiple disorders. This emphasis aligns with depression's high lifetime prevalence in UKB – affecting approximately 24% of participants (Davis et al., 2020). Schizophrenia was the second most studied disorder, appearing in 26 papers. However, its much lower prevalence in the cohort – comprising hundreds rather than thousands of cases – meant that studies typically employed different methodologies, often relying on hospital linkage, genetic risk scores, or imaging rather than self-report or questionnaire data.

Several papers explored multiple disorders or transdiagnostic features, with self-harm being the most commonly studied of these. Among the 87 studies addressing physical – mental comorbidity, cardiometabolic illness, or traits emerged as the most frequent physical health component. Approximately 36% of papers employed genetic techniques to investigate the relationship between physical and mental disorders.

To support future users of the UKB resource, we have provided guidance on navigating the data showcase, sharing code, and handling linked datasets in [Appendix 3](#) of the [Supplementary Material](#). A critical methodological issue concerns how mental disorder variables are defined within UKB. Multiple data sources are available, each

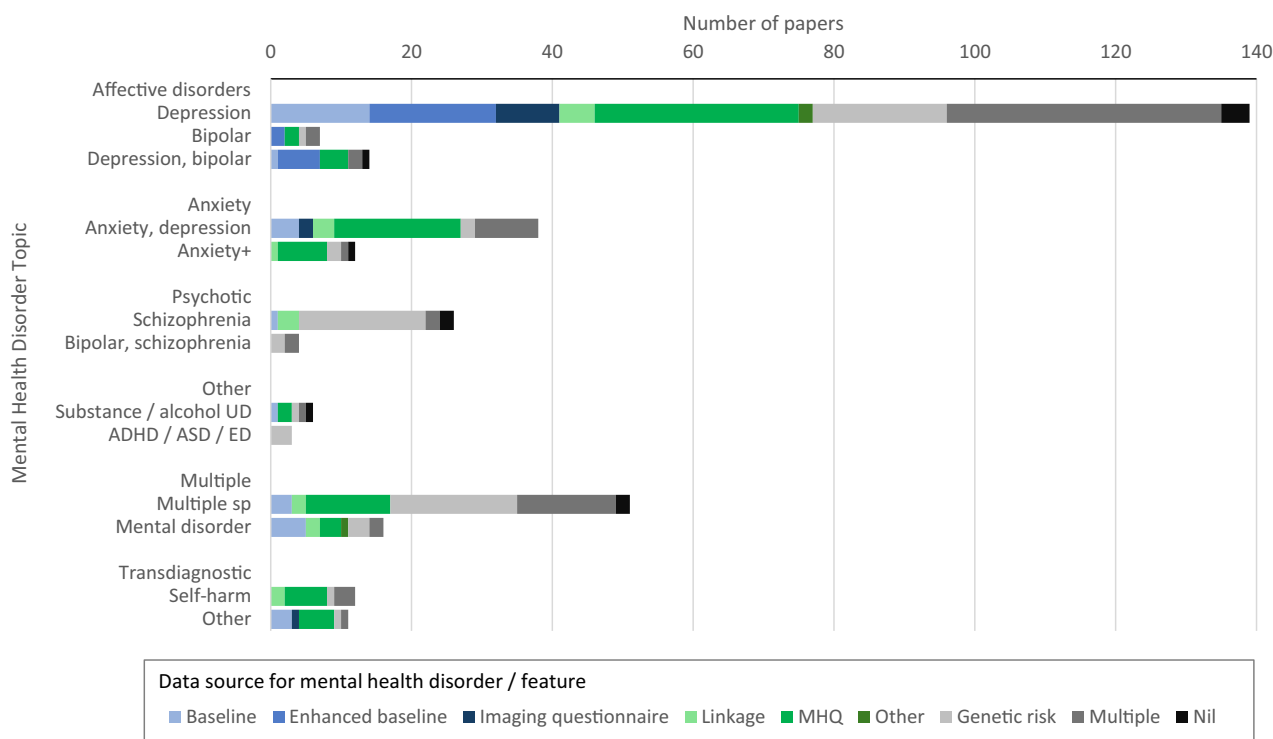


Figure 2. Numbers of papers categorized by mental disorder topic showing the data source used for ascertainment. Data sources, including ‘multiple’ are exclusive, with each paper represented only once. Nil data source occurs when disorder is not specified in UKB, for example, when used as control cohort or GWAS results from UKB used on external cohort. Note: ADHD, attention-deficit and hyperactivity disorder; ASD, autism spectrum disorder; ED, eating disorder; Multiple sp, multiple specific disorders of ‘mental disorder’ which is one variable for multiple disorders; Other transdiagnostic, psychotic experiences, anhedonia, mood instability and other; Anxiety+, generalized anxiety disorder and/or PTSD; UD, substance or alcohol use disorder.

with distinct strengths and limitations. [Supplementary Table S17](#) in Appendix 3 summarizes these. Self-reports are prone to recall bias; hospital and general practitioner (GP) linkage data are often insensitive; brief questionnaires may lack diagnostic specificity; and genetic data capture only a fraction of mental health heterogeneity. Davis *et al.* (2019) demonstrated that prevalence estimates for depression, Generalised Anxiety Disorder (GAD), bipolar disorder, and psychosis vary substantially depending on the data source. Their findings revealed poor overlap between self-reported diagnoses, symptom-based measures like the Composite International Diagnostic Interview - Short Form (CIDI-SF), and hospital records, with the latter identifying only 5–10% of cases detected by other methods. The latest enhancement, MHQ2, introduces longitudinal data, new disorder ascertainment, and additional social factors, thus expanding the scope for future research (<https://osf.io/c65t7>).

Incorporating multiple data enhancements can enrich research but also narrows the sample size. For instance, while 157,000 participants completed MHQ1 and 169,000 completed MHQ2, only 111,000 completed both. Applying further filters – for example, restricting to those with GP, imaging, or actigraphy data – may reduce the sample size by up to 80%, thereby limiting statistical power, particularly for advanced genetic and longitudinal analyses.

Selection effects must also be considered. Despite UKB’s structured sampling via NHS primary care, participation was low (around 6% of those invited), leading to a cohort that is disproportionately well-educated, health conscious, and socioeconomically advantaged (Fry *et al.*, 2017). Minoritized ethnic groups are underrepresented, with most participants of White European ancestry. This demographic skew compromises the evaluation of ethnic disparities and limits generalizability. Genetic studies have further

restricted diversity by historically excluding non-European ancestry participants, though more recent efforts aim to leverage the available diversity to improve relevance and impact (Carress, Lawson, & Elhaik, 2021; Singh *et al.*, 2025).

These selection effects are particularly relevant to mental health research, as people with mental disorders tend to be underrepresented in low-participation cohorts (Knudsen, Hotopf, Skogen, Øverland, & Mykletun, 2010). Self-selection into enhancements compounds this bias. For example, MHQ1 participants have significantly lower neuroticism scores than the broader UKB cohort, even though higher neuroticism is a known risk factor for common mental disorders (Davis *et al.*, 2025). While UKB’s estimates for common disorders align reasonably well with population-level data (Davis *et al.*, 2020), linkage with secondary mental health records shows that individuals with severe mental illness are significantly underrepresented (Li *et al.*, 2022). Consequently, UKB is not suitable for estimating prevalence, even with statistical weighting. However, analyses focused on causes, mechanisms, and consequences are less susceptible to these biases, and UKB remains highly valuable for such research aims (Allen *et al.*, 2024; Batty, Gale *et al.*, 2020).

Genetic techniques have been central to much of the mental health research using UKB. The cohort’s size and high-quality genotyping support detection of small genetic effects and exploration of gene–environment interplay (Bycroft *et al.*, 2018; Garg *et al.*, 2024). Beyond examining individual Single Nucleotide Polymorphism (SNPs), researchers have used PRSs, Mendelian randomization, and genetic correlations to study complex traits. Notable examples include the use of PRS as a proxy for unmeasured gut microbiota in studies of mental health (Qi *et al.*, 2021), and multiple studies linking schizophrenia PRS to brain structure

(Grama et al., 2020; Neilson et al., 2019). Mendelian randomization has helped infer causal directionality, such as identifying depression as a risk factor for inflammatory bowel disease (Luo, Xu, Noordam, van Heemst, & Li-Gao, 2021). Genetic correlation analyses using linkage disequilibrium score regression have illuminated relationships between mood disorders, biological rhythms, and physical health (Chen et al., 2022; Sirignano et al., 2022). Candidate gene studies have explored mechanisms such as the relaxin pathway in depression (Wong, Arathimos, Lewis, Young, & Dawe, 2023) and cytochrome P450 genotype in relation to antidepressant-related falls (Pronk et al., 2022).

These approaches help triangulate findings, reduce confounding, and strengthen causal inference (Fallin, Duggal, & Beaty, 2016; Lawlor, Tilling, & Davey Smith, 2017; Power et al., 2024). Nonetheless, interpretation must be cautious. Genetic factors account for only a modest proportion of variance in complex human traits (Tanksley, Motz, Kail, Barnes, & Liu, 2019), and the predominance of White European ancestry in UKB's genomic data again limits generalizability (Carress et al., 2021). Moreover, the benefits of combining multiple enhancements must be weighed against the resultant drop in usable sample size, which can erode statistical power for genetic and other advanced techniques.

Strength and limitations

This review offers insight into the breadth of mental health research conducted using UKB. In selecting studies, we assumed that all relevant papers would appear on the UKB website and, for 2022–2023, that researchers citing UKB data would reference one of our identified methodology papers. While we sought consistency by assigning each paper a single secondary topic, some studies may have been missed due to categorization constraints. We did not include studies using the second mental health questionnaire (MHQ2), as doing so would have delayed publication; however, this review remains a useful reference point for researchers exploring the MHQs. The summaries provided in the [Supplementary Material](#) are intended as a practical guide rather than a comprehensive listing. In the course of the review, we have created a timeline, a comparison of data types and other resources. We hope that this will inspire high-quality future research and the interpretation of UKB results.

Conclusions

Our review of mental health disorder research in UKB highlighted common themes and demonstrated the diverse approaches used to study mental health. It is encouraging to see research exploring mental health from multiple perspectives, as a comprehensive understanding of etiology, risk factors, and mechanisms is essential for advancing prevention and treatment strategies (Patel et al., 2018). UKB provides a unique opportunity to investigate many theories within a single resource, including associations with diet, inflammation, brain connectome types, and genotypes. Moreover, it serves as an important resource for replication, triangulation, and sensitivity analyses, helping to reinforce or refine findings from other studies (Sheehan & Didelez, 2020; VanderWeele, 2021).

A deeper understanding of complex multimorbidity, where mental and physical long-term conditions accumulate over time, is crucial for improving population health (Ronaldson et al., 2021). UKB remains a valuable tool for addressing these questions, offering researchers the ability to explore links between mental and physical health.

As UKB continues to grow in both the scope of its data and the volume of research it supports, it remains adaptable to researchers' evolving priorities and technological advancements. However, this expansion also increases the dataset's complexity, requiring careful navigation to maximize its potential. To fully leverage the opportunities UKB offers, researchers must continue refining their methodologies while acknowledging their inherent limitations.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S0033291725101359>.

Acknowledgement. K.A.S.D. and M.H. are supported by the National Institutes of Health Research (NIHR) KCL/South London and Maudsley NHS Trust Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Funding statement. K.A.S.D. and M.H. were part-supported by the NIHR KCL/Maudsley Biomedical Research Centre.

Competing interests. The authors declare no conflicts of interest.

References

- Allen, N. E., Lacey, B., Lawlor, D. A., Pell, J. P., Gallacher, J., Liam, S., Elliott, P., Matthews, P. M., Lyons, R. A., Whetton, A. D., Anneke, L., Hurles, M. E., Chapman, M., Roddam, A. W., Fitzpatrick, N. K., Hansell, A. L., Hardy, R., Marioni, R. E., O'Donnell, V. B., Williams, J., Lindgren, C. M., Effingham, M., Sellors, J., Danesh, J., & Collins, R. (2024). Prospective study design and data analysis in UK Biobank. *Science Translational Medicine*, *16*(729). <https://doi.org/10.1126/scitranslmed.adf4428>.
- Batty, G. D., Deary, I. J., Luciano, M., Altschul, D. M., Kivimäki, M., & Gale, C. R. (2020). Psychosocial factors and hospitalisations for COVID-19: Prospective cohort study based on a community sample. *Brain, Behavior, and Immunity*, *89*, 569–578. <https://doi.org/10.1016/j.bbi.2020.06.021>.
- Batty, G. D., Gale, C. R., Kivimäki, M., Deary, I. J., & Bell, S. (2020). Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ*, *368*, m131. <https://doi.org/10.1136/bmj.m131>.
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., & O'Connell, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209.
- Carress, H., Lawson, D. J., & Elhaik, E. (2021). Population genetic considerations for using biobanks as international resources in the pandemic era and beyond. *BMC Genomics*, *22*(1), 351. <https://doi.org/10.1186/s12864-021-07618-x>.
- Chen, G., Xie, J., Liu, W., Liang, T., Liao, X., Liao, W., Song, L., & Zhang, X. (2022). Association between depression and sleep apnoea: A Mendelian randomisation study. *ERJ Open Research*, *8*(1), 00394–02021. <https://doi.org/10.1183/23120541.00394-2021>.
- Coleman, J. R., Peyrot, W. J., Purves, K. L., Davis, K. A., Rayner, C., Choi, S. W., Hübel, C., Gaspar, H. A., Kan, C., & Van der Auwera, S. (2020). Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Molecular Psychiatry*, *25*(7), 1430–1446.
- Davis, K. A., Coleman, J. R., Adams, M., Allen, N., Breen, G., Cullen, B., Dickens, C., Fox, E., Graham, N., & Holliday, J. (2020). Mental health in UK Biobank—development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open*, *6*(2), e18.
- Davis, K. A. S., Coleman, J. R. I., Adams, M., Breen, G., Cai, N., Davies, H. L., Davies, K., Dregan, A., Eley, T. C., Fox, E., Holliday, J., Hübel, C., John, A., Kassam, A. S., Kempton, M. J., Lee, W., Li, D., Maina, J., McCabe, R., McIntosh, A. M., Oram, S., Richards, M., Skelton, M., Starkey, F., ter Kuile, A. R., Thornton, L. M., Wang, R., Yu, Z., Zvrskovec, J., & Hotopf, M. (2025). The UK Biobank mental health enhancement 2022: Methods and results. *PLoS One*, *20*(5), e0324189. <https://doi.org/10.1371/journal.pone.0324189>.

- Davis, K. A. S., Cullen, B., Adams, M., Brailean, A., Breen, G., Coleman, J. R. I., Dregan, A., Gaspar, H. A., Hübel, C., Lee, W., McIntosh, A. M., Nolan, J., Pearsall, R., & Hotopf, M. (2019). Indicators of mental disorders in UK Biobank—A comparison of approaches. *International Journal of Methods in Psychiatric Research*, **28**(3), e1796. <https://doi.org/10.1002/mpr.1796>.
- Fallin, M. D., Duggal, P., & Beaty, T. H. (2016). Genetic Epidemiology and Public Health: The Evolution From Theory to Technology. *American Journal of Epidemiology*, **183**(5), 387–393. <https://doi.org/10.1093/aje/kww001>.
- Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American Journal of Epidemiology*, **186**(9), 1026–1034.
- Garg, M., Karpinski, M., Matelska, D., Middleton, L., Burren, O. S., Hu, F., Wheeler, E., Smith, K. R., Fabre, M. A., Mitchell, J., O'Neill, A., Ashley, E. A., Harper, A. R., Wang, Q., Dhindsa, R. S., Petrovski, S., & Vitsios, D. (2024). Disease prediction with multi-omics and biomarkers empowers case-control genetic discoveries in the UK Biobank. *Nature Genetics*, **56**(9), 1821–1831. <https://doi.org/10.1038/s41588-024-01898-1>.
- GBD Mental Health Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Psychiatry*, **9**(2), 137–150. [https://doi.org/10.1016/S2215-0366\(21\)00395-3](https://doi.org/10.1016/S2215-0366(21)00395-3).
- Grama, S., Willcocks, I., Hubert, J. J., Pardiñas, A. F., Legge, S. E., Bracher-Smith, M., Menzies, G. E., Hall, L. S., Pocklington, A. J., Anney, R. J. L., Bray, N. J., Escott-Price, V., & Caseras, X. (2020). Polygenic risk for schizophrenia and subcortical brain anatomy in the UK Biobank cohort. *Translational Psychiatry*, **10**(1), 309. <https://doi.org/10.1038/s41398-020-00940-0>.
- Knudsen, A. K., Hotopf, M., Skogen, J. C., Øverland, S., & Mykletun, A. (2010). The Health Status of Nonparticipants in a Population-based Health Study: The Hordaland Health Study. *American Journal of Epidemiology*, **172**(11), 1306–1314. <https://doi.org/10.1093/aje/kwq257>.
- Lawlor, D. A., Tilling, K., & Davey Smith, G. (2017). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, **45**(6), 1866–1886. <https://doi.org/10.1093/ije/dyw314>.
- Li, Z., Kormilitzin, A., Fernandes, M., Vaci, N., Liu, Q., Newby, D., Goodday, S., Smith, T., Nevado-Holgado, A. J., & Winchester, L. (2022). Validation of UK Biobank data for mental health outcomes: A pilot study using secondary care electronic health records. *International Journal of Medical Informatics*, **160**, 104704. <https://doi.org/10.1016/j.ijmedinf.2022.104704>.
- Luo, J., Xu, Z., Noordam, R., van Heemst, D., & Li-Gao, R. (2021). Depression and inflammatory bowel disease: A bidirectional two-sample Mendelian randomization study. *Journal of Crohn's and Colitis*, **16**(4), 633–642. <https://doi.org/10.1093/ecco-jcc/jjab191>.
- Neilson, E., Shen, X., Cox, S. R., Clarke, T. K., Wigmore, E. M., Gibson, J., Howard, D. M., Adams, M. J., Harris, M. A., Davies, G., Deary, I. J., Whalley, H. C., McIntosh, A. M., & Lawrie, S. M. (2019). Impact of polygenic risk for schizophrenia on cortical structure in UK Biobank. *Biological Psychiatry*, **86**(7), 536–544. <https://doi.org/10.1016/j.biopsych.2019.04.013>.
- Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., Chisholm, D., Collins, P. Y., Cooper, J. L., Eaton, J., Herrman, H., Herzallah, M. M., Huang, Y., Jordans, M. J. D., Kleinman, A., Medina-Mora, M. E., Morgan, E., Niaz, U., Omigbodun, O., Prince, M., Rahman, A., Saraceno, B., Sarkar, B. K., De Silva, M., Singh, I., Stein, D. J., Sunkel, C., & Unützer, J. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet*, **392**(10157), 1553–1598. [https://doi.org/10.1016/S0140-6736\(18\)31612-X](https://doi.org/10.1016/S0140-6736(18)31612-X).
- Power, G. M., Sanderson, E., Pagoni, P., Fraser, A., Morris, T., Prince, C., Frayling, T. M., Heron, J., Richardson, T. G., & Richmond, R. (2024). Methodological approaches, challenges, and opportunities in the application of Mendelian randomisation to lifecourse epidemiology: A systematic literature review. *European Journal of Epidemiology*, **39**(5), 501–520.
- Pronk, A. C., Seppala, L. J., Trajanoska, K., Stringa, N., van de Loo, B., de Groot, L. C. P. G. M., van Schoor, N. M., Koskeridis, F., Markozannes, G., Ntzani, E., Uitterlinden, A. G., Rivadeneira, F., Stricker, B. H., & van der Velde, N. (2022). Candidate genetic variants and antidepressant-related fall risk in middle-aged and older adults. *PLoS One*, **17**(4), e0266590. <https://doi.org/10.1371/journal.pone.0266590>.
- Qi, X., Guan, F., Cheng, S., Wen, Y., Liu, L., Ma, M., Cheng, B., Liang, C., Zhang, L., Liang, X., Li, P., Chu, X., Ye, J., Yao, Y., & Zhang, F. (2021). Sex specific effect of gut microbiota on the risk of psychiatric disorders: A Mendelian randomisation study and PRS analysis using UK Biobank cohort. *The World Journal of Biological Psychiatry*, **22**(7), 495–504. <https://doi.org/10.1080/15622975.2021.1878428>.
- Ronaldson, A., Arias de la Torre, J., Prina, M., Armstrong, D., Das-Munshi, J., Hatch, S., Stewart, R., Hotopf, M., & Dregan, A. (2021). Associations between physical multimorbidity patterns and common mental health disorders in middle-aged adults: A prospective analysis using data from the UK Biobank. *The Lancet Regional Health Europe*, **8**, 100149. <https://doi.org/10.1016/j.lanepe.2021.100149>.
- Sheehan, N. A., & Didelez, V. (2020). Epidemiology, genetic epidemiology and Mendelian randomisation: More need than ever to attend to detail. *Human Genetics*, **139**(1), 121–136. <https://doi.org/10.1007/s00439-019-02027-3>.
- Singh, M., Chatzinakos, C., Barr, P. B., Gentry, A. E., Bigdeli, T. B., Webb, B. T., & Peterson, R. E. (2025). Trans-ancestry genome-wide analyses in UK Biobank yield novel risk loci for major depression. *medRxiv*. <https://doi.org/10.1101/2025.02.22.25322721>.
- Sirignano, L., Streit, F., Frank, J., Zillich, L., Witt, S. H., Rietschel, M., & Foo, J. C. (2022). Depression and bipolar disorder subtypes differ in their genetic correlations with biological rhythms. *Scientific Reports*, **12**(1), 15740. <https://doi.org/10.1038/s41598-022-19720-5>.
- Smith, D. J., Nicholl, B. I., Cullen, B., Martin, D., Ul-Haq, Z., Evans, J., Gill, J. M., Roberts, B., Gallacher, J., & Mackay, D. (2013). Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. *PLoS One*, **8**(11), e75362.
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., & Landray, M. (2015). UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*, **12**(3), e1001779.
- Tanksley, P. T., Motz, R. T., Kail, R. M., Barnes, J. C., & Liu, H. (2019). The genome-wide study of human social behavior and its application in sociology. *Frontiers in Sociology*, **4**. <https://doi.org/10.3389/fsoc.2019.00053>.
- VanderWeele, T. J. (2021). Can sophisticated study designs with regression analyses of observational data provide causal inferences? *JAMA Psychiatry*, **78**(3), 244–246.
- Wong, W. L. E., Arathimos, R., Lewis, C. M., Young, A. H., & Dawe, G. S. (2023). Investigating the role of the relaxin-3/RXFP3 system in neuropsychiatric disorders and metabolic phenotypes: A candidate gene approach. *PLoS One*, **18**(11), e0294045.
- Wykes, T., Bell, A., Carr, S., Coldham, T., Gilbody, S., Hotopf, M., Johnson, S., Kabir, T., Pinfold, V., Sweeney, A., Jones, P. B., & Creswell, C. (2023). Shared goals for mental health research: What, why and when for the 2020s. *Journal of Mental Health*, **32**(6), 997–1005. <https://doi.org/10.1080/09638237.2021.1898552>.