What can UK Biobank's 500,000 participants teach us about chronic pain?

Eoin Maurice Kelleher[1,2], Anushka Irani[1,3]

[1]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, OX3 9DU, UK

[2] Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

[3]Department of Rheumatology, Mayo Clinic, Jacksonville, FL, USA

Correspondence to: Dr Eoin Kelleher

Department of Anesthesia, Critical Care and Pain Medicine,

Massachusetts General Hospital,

55 Fruit Street, GRB 460,

Boston, MA 02114, USA

E-mail: ekelleher11@mgh.harvard.edu

Twitter/X: @EoinKr

Keywords: Chronic pain; UK Biobank; Epidemiology; Cohort Study; Neuroimaging

Summary

UK Biobank is increasingly used for the study of chronic pain, yet clinicians and researchers may be unfamiliar with the strengths and limitations of this cohort. In this editorial we provide practical guidance for interpreting UK Biobank pain research, highlighting key examples of how UK Biobank has advanced our understanding of pain. We describe the available pain assessments, compare UK Biobank with other population cohorts, and discuss key methodological considerations including selection bias, collider effects, and temporal misalignment between data modalities. When interpreted appropriately and triangulated with complementary study designs, UK Biobank represents a valuable resource for advancing our understanding of chronic pain.

Introduction

Chronic pain affects approximately one-third of adults and represents a leading cause of disability worldwide.(1) Despite its prevalence, fundamental questions remain: what biological mechanisms underlie chronic pain, why do some individuals develop persistent pain whilst others recover, and how can we identify at-risk individuals for early intervention? Many studies have been constrained by a fundamental trade-off: small studies with detailed pain phenotyping lack statistical power to detect modest effect sizes, whilst large cohort studies have relied on rudimentary pain assessments that limit mechanistic insights or ability to distinguish between different pain phenotypes. Advancing our understanding of chronic pain requires cohorts that combine depth of phenotyping with sufficient scale for reliable detection and characterisation of risk factors.(2, 3)

UK Biobank represents a valuable resource for addressing these challenges. Its multi-modal data collection - spanning questionnaires, genomics, proteomics, neuroimaging, and health record linkage - enables investigation of pain across multiple levels. Importantly, UK Biobank's large sample size not only permits detection of small effect sizes but also provides more accurate and stable estimates, particularly for genetics and imaging studies. Effect sizes from small studies are often inflated due to overfitting and winner's curse (where only the largest, often chance-inflated, effects reach statistical significance and publication). This is particularly relevant for neuroimaging, where high dimensionality (thousands of brain features per participant) exacerbates overfitting and has historically limited reproducibility.(4) UK Biobank's ~100,000-participant imaging sub-study mitigates this problem. This may explain why UK Biobank studies consistently identify modest biological

effects for chronic pain: these estimates likely reflect true effect sizes rather than the inflated values typical of underpowered discovery studies.

However, UK Biobank's design introduces important methodological constraints that, if unacknowledged, risk spurious associations and misinterpretation. In this editorial, we examine how UK Biobank can advance chronic pain research whilst highlighting considerations essential for valid inference. This editorial aims to equip clinicians and researchers with practical guidance for conducting and interpreting UK Biobank pain research.

UK Biobank: design and pain assessments

UK Biobank recruited ~500,000 adults aged 40-69 between 2006 and 2010 across England, Scotland, and Wales.(5) At baseline, participants completed physical measurements, biological sampling, and questionnaires. Initial pain-related data were limited: participants reported pain presence and duration across seven anatomical regions (head, face, neck/shoulder, stomach/abdomen, back, hip, knee) or "all over the body" within the preceding month and persisting for three or more months, the definition used for chronic pain in ICD-11.(6)

A dedicated online Pain Questionnaire, added in 2019 and designed by a multidisciplinary expert panel, substantially enhances the level of detail about participants' pain. Of ~335,000 participants invited, ~166,700 completed this survey. The questionnaire incorporates multiple validated instruments: the 2016 American College of Rheumatology revised Fibromyalgia survey criteria(7), combining the Widespread Pain Index (WPI, 0-19 body regions) with the Symptom Severity Scale (SSS, 0-12) to generate a Fibromyalgia Index (FMI,

0-31); the Douleur Neuropathique 4 (DN4) for neuropathic pain screening;(8) and select items from the Brief Pain Inventory assessing pain intensity and interference. The FMI functions both as a diagnostic criterion for fibromyalgia and as a potential continuous marker of nociplastic pain severity.(9, 10) A follow-up Pain and Sensation 2 questionnaire, closely modelled on the 2019 instrument, was launched in 2024. These data are anticipated to be released to researchers in the near future, enabling longitudinal assessment of pain trajectories.

Beyond pain-specific assessments, UK Biobank's breadth of complementary data expands research possibilities. All participants have genome-wide genotyping data enabling genome-wide association studies (GWAS) that have identified novel genetic loci for chronic pain.(11) Over 100,000 participants underwent multi-organ imaging including brain MRI, permitting investigation of central nervous system correlates.(12) Linked health records to both primary and secondary care also provide longitudinal data on diagnoses, prescriptions, and healthcare utilisation. Recent additions include proteomic profiling and accelerometry data. The complete range of assessments available in UK Biobank are summarised in **Figure 1**.

UK Biobank in context: comparison with other population cohorts

Of course, UK Biobank is not unique in enabling pain research at the population-level. A range of population-based cohorts worldwide have collected pain data, though with varying depths of phenotyping and range of complementary data (**Table 1**). An understanding of UK Biobank's position relative to its peers clarifies which research questions it is best suited to address.

It should be noted that several cohorts offer capabilities UK Biobank lacks. The Tromsø and Rotterdam Studies, for example, include quantitative sensory testing enabling experimental pain validation. The HUNT cohort's repeated assessments across four decades permit prospective trajectory analysis. Generation Scotland's family-based design enables within-family analyses with improved controlling for unmeasured confounding.

Nevertheless, UK Biobank's scale and range of data remain unmatched. The 2019 Pain Questionnaire represents the most comprehensive assessment of pain at population level. The 100,000-participant imaging sub-study dwarfs imaging samples in other cohorts (HUNT n~1,000; Rotterdam n~6,000; Generation Scotland STRADL n~1,085, **Table 1**), although the German-based NAKO health study comes closest with ~32,000 participants.(13) For GWAS requiring hundreds of thousands of participants to detect variants of modest effect, only UK Biobank, Million Veteran Program, China Kadoorie Biobank, and All of Us approach sufficient power, but only UK Biobank offers the same depth of pain assessment and corresponding imaging data.

Triangulation across cohorts can help strengthen causal inference and address limitations inherent to any individual study. For example, Tanguay-Sabourin and colleagues replicated their prognostic risk score for pain spreading derived from UK Biobank in two smaller independent cohorts, the Northern Finland Birth Cohort and the PREVENT-AD cohort.(14)

Advancing pain research: examples from UK Biobank

Collectively, research from UK Biobank has given us novel insights about chronic pain.

First, work from UK Biobank has demonstrated that pain conditions share substantial genetic architecture. Johnston and colleagues identified 76 independent genetic variants across 39 loci associated with multisite chronic pain, with gene enrichment in neurogenesis and synaptic plasticity pathways, and demonstrated strong genetic correlation with depression (rg=0.53) alongside bidirectional causal relationships with Mendelian randomisation.(11) Subsequent work focusing more specifically on chronic back pain identified additional loci including *SOX5*, implicating developmental pathways, and found genetic overlap with intervertebral disc degeneration and psychiatric traits.(15, 16) These foundational studies have been followed by numerous site-specific GWAS examining pain at individual anatomical locations, including neck/shoulder,(17) hip,(18) and knee pain,(19) as well as investigations of chronic postsurgical pain,(20) pain medication response,(21) and sex-stratified analyses.(22) The large sample size of UK Biobank enables detection of novel genetic associations impossible in smaller cohorts. However, genetic instruments for "chronic pain" may capture pleiotropy (associations with correlated traits such as depression, neuroticism, or general distress) rather than pain-specific mechanisms. Without experimental validation or careful triangulation using methods like Mendelian randomisation with well-defined pain subtypes, distinguishing whether genetic associations reflect true nociceptive pathways, confounding by psychological comorbidities, or reverse causation remains difficult. However, the more detailed pain assessments in the 2019 and 2024 questionnaires may help ameliorate these issues.

Second, proteomic signatures observed in UK Biobank distinguish pain from pain-free states. Analysis of 2,923 plasma proteins identified 474 associated with chronic pain across body sites, with 11 proteins common to all sites with enrichment in immune and metabolic

pathways.(23, 24) Mendelian randomisation suggested potential causal roles for immune-related proteins including Tumour Necrosis Factor family members, though the phenotypic and pleiotropy challenges outlined above apply here as well.

Third, psychosocial characteristics, rather than biological factors, dominate pain prediction in UK Biobank. In a landmark pain study, Tanguay-Sabourin and colleagues developed a prognostic model using 99 biopsychosocial features, and found that sleep disturbance, low mood, stressful life events, and high BMI predicted future pain spreading more strongly than biological markers (AUC 0.68-0.78), with a simplified 6-item tool showing comparable performance in independent cohorts.(14)

Fourth, neuroimaging provides insights on pain phenotypes. Imaging-derived phenotypes (quantitative measures of grey matter volumes, white matter tract integrity, and resting-state network components) are provided to facilitate analyses.(12) However, researchers can also derive novel phenotypes from raw imaging data. For example, in a recent study from our group, higher scores on the Fibromyalgia Index (derived from the 2016 American College of Rheumatology criteria for fibromyalgia(7)) are associated with altered structural and functional connectivity in the descending pain modulation system, supporting its use as a continuous marker of nociplastic pain severity.(25) Fillingim and colleagues have also applied machine learning to imaging modalities to develop predictive models for chronic pain conditions.(26)

Given that participants were imaged after potentially years of chronic pain, observed associations may reflect consequences rather than causal mechanisms. However, ongoing repeat imaging assessments will enable longitudinal analyses as sample sizes increase, addressing these limitations. The phenotyping and pleiotropy issues discussed above apply

equally here: brain differences attributed to "chronic pain" may instead reflect comorbid depression, sleep disturbance, physical inactivity, or medication effects, particularly given the rudimentary baseline pain assessments (though analyses using the more detailed 2019 Pain Questionnaire may address this). Furthermore, technical considerations warrant caution, head motion artefact (potentially greater in pain patients), scanner effects across imaging centres, and other confounds may influence findings and should be accounted for in analyses.(27)

Beyond UK Biobank, collaborative initiatives such as the ENIGMA-Chronic Pain working group offer complementary approaches by pooling neuroimaging data across independently collected datasets worldwide, enabling cross-validation of findings and investigation of brain correlates across heterogeneous pain conditions.(28)

Of relevance to anaesthesiologists is whether UK Biobank can inform our understanding of perioperative pain. Chronic post-surgical pain (CPSP) affects 20-30% of surgical patients and represents a major public health burden.(29) Surgical procedure data is available in UK Biobank via healthcare record linkage, and the 2019 Pain Questionnaire specifically enquired about CPSP. Recent studies have leveraged these data to identify genetic risk factors and pre-operative predictors of CPSP.(20) However, the 2019 CPSP question relies on retrospective self-report of surgery as pain cause, which may be unreliable years after procedures.

This small selection of findings from UK Biobank outlined above demonstrate the value of its combination of scale, depth of pain phenotyping, and breadth of data including genetics, proteomics, and imaging.

Limitations of UK Biobank for pain research

Phenotyping pain mechanisms

UK Biobank's greatest strength – scale – also introduces challenges. An important limitation is difficulty phenotyping pain mechanisms reliably in such a large cohort. Without detailed clinical evaluation, physical examination or quantitative sensory testing (available in Rotterdam Study and Tromsø Study; **Table 1**), researchers cannot reliably distinguish nociceptive (from tissue damage), neuropathic (from nerve injury), and nociplastic (altered central processing) pain, distinctions that are important for understanding pathophysiology. Whilst the 2019 Pain Questionnaire includes the DN4 for neuropathic screening, this instrument was developed as a screening tool rather than a diagnostic one, and its positive predictive value in population-based settings without clinical correlation remains uncertain. Furthermore, many people with chronic pain likely experience mixed pain mechanisms, yet UK Biobank's assessment cannot disentangle these components.

If different mechanisms have distinct genetic architectures, neuroimaging signatures, or proteomic profiles, combining them attenuates signals and may obscure biologically meaningful associations. This mechanistic heterogeneity underlies many limitations discussed in the previous section: genetic pleiotropy may partly reflect different pain mechanisms having different genetic correlates; neuroimaging associations may differ by mechanism; treatment responses may vary by mechanism. The 2019 Pain Questionnaire represents a substantial improvement over baseline assessment, and over other population cohorts, enabling mechanistic differentiation through validated instruments. Furthermore,

UK Biobank's utility extends beyond mechanistic discovery. Healthcare record linkage, self-reported diagnoses, and prescription data provide indirect phenotypic information that can differentiate pain aetiologies (e.g., rheumatoid arthritis versus fibromyalgia versus post-surgical pain). These data enable investigation of pain trajectories, prognosis, and development of prognostic models. These clinically valuable applications do not require experimental or clinical validation of mechanisms.

Selection bias and collider bias

UK Biobank's 5.5% response rate results in a cohort that, although large, differs substantially from the general UK population: participants are more educated, wealthier, healthier, with approximately 50% lower all-cause mortality.(30) The cohort is also >94% white, limiting investigation of pain disparities and population-specific effects across ethnicities. This contrasts with All of Us (50% non-white) and MVP (**Table 1**), for example, though those cohorts have less comprehensive pain assessments.

Beyond limiting prevalence estimates, this selection introduces collider bias, a frequently under-appreciated problem that can generate spurious associations. Collider bias occurs when two variables independently influence study participation, creating spurious associations when conditioning on that participation.(31) For example, if both mental health and chronic pain independently influence participation likelihood, analysing only UK Biobank participants may induce a spurious association between these variables even if none exists in the general population. Similarly, if both socioeconomic position and pain severity affect participation, collider bias may distort observed socioeconomic gradients in pain. Munafò and colleagues demonstrate that even modest selection effects (OR=1.5) combined with moderate genetic associations (r=0.1) can generate substantial bias.(32) The imaging sub-

study faces compounded selection as chronic pain may prevent participation through travel limitations or MRI contraindications.

Thus, although UK Biobank's sample size generates precise estimates for small associations (OR 1.05-1.20), this precision does not overcome selection bias. This is particularly problematic for pain research where pain itself impacts health-seeking behaviour and disability, factors that may have influenced participation.

Temporal considerations

UK Biobank's data collection timeline creates challenges for pain research. The Pain Questionnaire was administered in 2019-2020, yet brain imaging occurred over the decade beginning in 2014 and proteomic sampling predominantly at baseline (2006-2010). This temporal misalignment introduces measurement error when linking pain phenotypes with other data, such as genetics, imaging or proteomics. Additionally, mid-life recruitment (ages 40-69) precludes prospective assessment of early-life pain experiences that may predispose to later chronic pain, limiting investigations of pain development across the lifespan. The 2024 Pain and Sensation 2 questionnaire, with data release pending, will partially address these temporal limitations by enabling longitudinal analyses of pain trajectories and transitions over a five-year interval.

Treatment effects and confounding by indication

UK Biobank contains prescription data linked through primary care records (although at present these are only available until 2016-2017), offering opportunities to examine medication-pain associations. However, confounding by indication is inevitable in observational analyses, and genetic instruments for medication use may reflect treatment-

seeking propensity rather than medication effects *per se*. The target trial emulation framework has been applied to UK Biobank for cardiovascular outcomes,(33) and may provide a solution to overcome some limitations for pharmacoepidemiology of pain in UK Biobank.

Conclusion: what UK Biobank can and cannot tell us

UK Biobank can identify genetic loci and neuroimaging correlates of chronic pain at a scale impossible in smaller cohorts, provide stable effect size estimates free from winner's curse, and generate hypotheses about shared biological pathways across pain conditions. It can support prognostic model development and enable investigation of pain trajectories over time, particularly as longitudinal data from the 2024 Pain and Sensation 2 questionnaire become available.

However, UK Biobank cannot establish causality from cross-sectional associations, reliably distinguish pain mechanisms without clinical examination, or fully overcome selection bias inherent to its 5.5% response rate. Observed associations may reflect collider bias or pleiotropy with psychological traits rather than pain-specific biology.

For clinicians interpreting UK Biobank pain research, two principles apply: first, modest effect sizes are expected and likely accurate, not evidence of failed studies. Second, findings should be triangulated with complementary cohorts before informing clinical practice. Used appropriately, UK Biobank represents a substantial advance in our understanding of chronic pain at the population level.

Funding & Declaration of Interests

The authors have no funding relevant to this editorial or interests to declare. EMK is an editorial fellow for the *British Journal of Anaesthesia*.

References

1. Mills SEE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population-based studies. Br J Anaesth. 2019;123(2):e273-e83.
2. Bycroft C. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562.
3. Dahl A, Thompson M, An U, Krebs M, Appadurai V, Border R, et al. Phenotype integration improves power and preserves specificity in biobank-based genetic studies of major depressive disorder. Nature Genetics. 2023;55(12):2082-93.
4. Mwangi B, Tian TS, Soares JC. A Review of Feature Reduction Techniques in Neuroimaging. Neuroinformatics. 2014;12(2):229-44.
5. Sudlow C. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12.
6. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. Pain. 2015;156(6):1003-7.
7. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RL, et al. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum. 2016;46(3):319-29.
8. Bouhassira D, Attal N, Alchaar H, Boureau F, Brochet B, Bruxelle J, et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). Pain. 2005;114(1-2):29-36.
9. Brummett CM, Janda AM, Schueller CM, Tsodikov A, Morris M, Williams DA, et al. Survey criteria for fibromyalgia independently predict increased postoperative opioid consumption after lower-extremity joint arthroplasty: a prospective, observational cohort study. Anesthesiology. 2013;119(6):1434-43.
10. Brummett CM, Urquhart AG, Hassett AL, Tsodikov A, Hallstrom BR, Wood NI, et al. Characteristics of fibromyalgia independently predict poorer long-term analgesic outcomes following total knee and hip arthroplasty. Arthritis Rheumatol. 2015;67(5):1386-94.
11. Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, et al. Genome-wide association study of multisite chronic pain in UK Biobank. PLOS Genetics. 2019;15(6):e1008164.
12. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JLR, Griffanti L, Douaud G, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. Neuroimage. 2018;166:400-24.
13. Bamberg F, Schlett CL, Caspers S, Ringhof S, Günther M, Hirsch JG, et al. Baseline MRI Examination in the NAKO Health Study—Findings on Feasibility, Participation and Dropout Rates, Comfort, and Image Quality. Dtsch Arztebl Int. 2024;121(18):587-93.
14. Tanguay-Sabourin C, Fillingim M, Guglietti GV, Zare A, Parisien M, Norman J, et al. A prognostic risk score for development and spread of chronic pain. Nature Medicine. 2023;29(7):1821-31.
15. Suri P, Palmer MR, Tsepilov YA, Freidin MB, Boer CG, Yau MS, et al. Genome-wide meta-analysis of 158,000 individuals of European ancestry identifies three loci associated with chronic back pain. PLoS Genet. 2018;14(9):e1007601.
16. Freidin MB, Tsepilov YA, Palmer M, Karssen LC, Group CMW, Suri P, et al. Insight into the genetic architecture of back pain and its risk factors from a study of 509,000 individuals. PAIN. 2019;160(6):1361-73.
17. Tao Y, Pan Q, Cai T, Lu ZH, Haque M, Dottorini T, et al. A genome-wide association study identifies novel genetic variants associated with neck or shoulder pain in the UK biobank (N = 430,193). Pain Rep. 2025;10(3):e1267.
18. Pan Q, Tao Y, Cai T, Veluchamy A, Hebert HL, Zhu P, et al. A genome-wide association study identifies genetic variants associated with hip pain in the UK Biobank cohort (N = 221,127). Sci Rep. 2025;15(1):2812.

19. Meng W, Adams MJ, Palmer CNA, Agee M, Alipanahi B, Bell RK, et al. Genome-wide association study of knee pain identifies associations with GDF5 and COL27A1 in UK Biobank. Communications Biology. 2019;2(1):321.
20. Li S, Toneman MK, Diatchenko L, Parisien M, Vissers KCP, Ten Broek RPG, et al. Genome-wide association study on chronic postsurgical pain in the UK Biobank. Br J Anaesth. 2025;134(3):783-92.
21. Li S, Poelmans G, van Boekel RLM, Coenen MJH. Genome-wide association study on pharmacological outcomes of musculoskeletal pain in UK Biobank. Pharmacogenomics J. 2023;23(6):161-8.
22. Johnston KJA, Ward J, Ray PR, Adams MJ, McIntosh AM, Smith BH, et al. Sex-stratified genome-wide association study of multisite chronic pain in UK Biobank. PLoS Genet. 2021;17(4):e1009428.
23. Li Z-Y, Ma Q, Zhang J, Yin R-Y, You J, Hao Q-Z, et al. Large-Scale Plasma Proteomics to Profile Pathways and Prognosis of Chronic Pain. Advanced Science. 2025;12(16):2410160.
24. Chen L, Kelleher E, Meng R, Liu D, Guo Y, Wang Y, et al. Diagnosis, Prognosis, and Drug Target Discovery for Chronic Widespread Pain: A Large Proteogenomic Study. Adv Sci (Weinh). 2025:e07691.
25. Kelleher EM, Lange F, Wanigasekera V, Rathod-Mistry T, Nichols T, Seymour B, et al. Brain signatures of nociplastic pain: Fibromyalgia Index and descending modulation at population level. Brain. 2025.
26. Fillingim M, Tanguay-Sabourin C, Parisien M, Zare A, Guglietti GV, Norman J, et al. Biological markers and psychosocial factors predict chronic pain conditions. Nature Human Behaviour. 2025;9(8):1710-25.
27. Alfaro-Almagro F, McCarthy P, Afyouni S, Andersson JLR, Bastiani M, Miller KL, et al. Confound modelling in UK Biobank brain imaging. Neuroimage. 2021;224:117002.
28. Quidé Y, Jahanshad N, Andoh J, Antoniou G, Apkarian AV, Ashar YK, et al. ENIGMA-Chronic Pain: a worldwide initiative to identify brain correlates of chronic pain. Pain. 2024;165(12):2662-6.
29. Rosenberger DC, Pogatzki-Zahn EM. Chronic post-surgical pain - update on incidence, risk factors and preventive treatment options. BJA Educ. 2022;22(5):190-6.
30. Fry A. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol. 2017;186.
31. Day FR, Loh PR, Scott RA, Ong KK, Perry JR. A Robust Example of Collider Bias in a Genetic Association Study. Am J Hum Genet. 2016;98(2):392-3.
32. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. International Journal of Epidemiology. 2017;47(1):226-35.
33. Zhou S, Chen R, Liu J, Guo Z, Su L, Li Y, et al. Comparative Effectiveness and Safety of Atorvastatin Versus Rosuvastatin : A Multi-database Cohort Study. Ann Intern Med. 2024;177(12):1641-51.

Figure 1. UK Biobank timeline showing evolution of pain assessment.
Pain phenotyping progressed from basic screening (7 body sites, 2006-2010) to comprehensive online questionnaire with validated instruments (2019-2020), which was repeated in 2024-2025 (data is pending release). Note temporal gap between imaging (2014-2018) and detailed pain assessment. Approximate sample sizes and key assessment modalities shown.