

MRC Mental Health Data Research Hub Survey

A survey was developed to seek insight from the research and innovation community regarding a draft specification for the Mental Health Data Research Hub. The survey was live from 12th - 25th of November and was advertised through the UKRI website and social media channels as well as through direct email to a breadth of stakeholders.

Introduction:

Mental health problems have a huge impact on the UK's health, wealth and wellbeing. An estimated one in six of us are affected by mental illness each year in the UK.

Mental health and illness stem from a multidimensional interplay of biological, psychological, social and environmental factors which influence risk and resilience to illness and response to treatment. Integrating biological, clinical, epidemiological, lifestyle-related, social, educational, environmental and other data sources offers the potential to accelerate the development of new treatments and approaches for mental health problems, facilitate biomarker and target discovery, deliver insights into the aetiology of mental ill-health, increase our understanding of the complex interactions between physical and mental health, and inform early intervention and prevention strategies for mental illness.

The MRC is planning a call for funding to support the establishment of a Mental Health Data Hub in partnership with Health Data Research UK. We expect this to build on, and add value to, other investments and activities in mental health data science, including the MRC funded Mental Health Data Pathfinder initiative. The aim is to develop a centre of excellence for curation of mental health data at scale that provides UK wide expert research data services

We would like to share with you the draft specification for the Hub as part of the consultation phase. We are seeking your feedback and comments on specific aspects of the specification. To do this we have created the below survey to gather your responses, which should take just under 15 minutes to complete. We will be closing the survey **4pm 25th November 2021**.

Your comments and feedback will help to align the call specification with the key needs of users and the priority opportunity areas of mental health research.

We look forward to hearing your thoughts.

Please provide your:

Name (optional)

Organisation (optional)

Role (optional)

Type of organisation (multiple choice) – Academic institution, Industry, Charity, NHS, Other (please specify)

- A total of 67 stakeholders from 31 different institutions responded to the survey. 61 were from an academic/research institution, 3 from an NHS institution and 3 from charities. No industry stakeholders responded to the survey. As such, this consultation is limited in its ability to scope the full breadth of user needs. Applicants will be expected to demonstrate additional evidence of engagement with a range of users to understand the full needs across the research and innovation community.

Please briefly outline your mental health research interests (limited free text box)

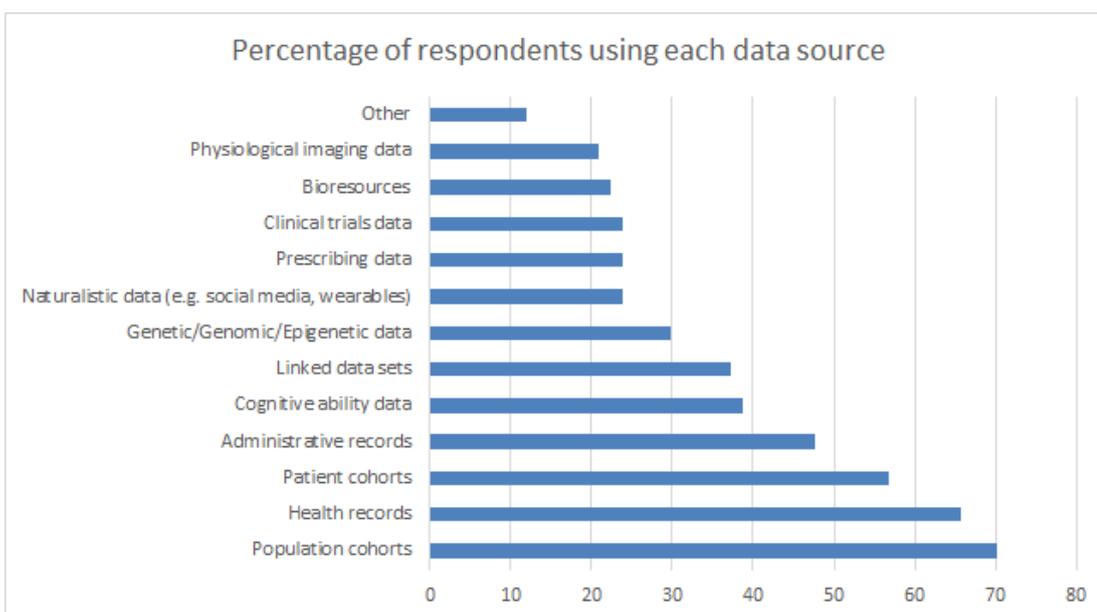
Responses are summarised below:

- Psychiatry
- Psychiatric genetics
- Gene/environment interactions
- Cognitive science
- Psychology
- Clinical informatics
- PPIE/service user research and co-production
- Population based research
- Cohort data management
- Prevention research

MRC Mental Health Data Research Hub Survey

- Epidemiology
- Health services research
- Inequalities in mental health
- Maternal mental health
- Child and adolescent mental health
- Digital solutions for mental health
- Psychopharmacology
- Public mental health
- Psychological intervention and evaluation research
- Social determinants of mental health
- Policy research
- Educational research
- Economics
- Comorbidities and physical health interface
- A wide range of mental health conditions
- Imaging
- Neurodegeneration
- Health inequalities
- Global mental health
- Methodology
- Developmental psychopathology
- Vulnerable groups
- AI/machine learning
- Sleep/circadian rhythms
- Wellbeing

What types of data sets do you currently use in your research (tick all that apply)? Percent of respondents who used this type of data



[Mental Health Data Research Hub draft specification - Consultation](#)

[Draft specification](#)

MRC Mental Health Data Research Hub Survey

This call for funding is to support the establishment of a Mental Health Data Hub over 3 years. It is anticipated that the Hub will continue to grow in scale and ambition beyond the 3 years, evolving into a coordinated and integrated resource for mental health research.

It is expected that the Hub will:

- Act as a centre of excellence for curation of mental health data at scale and providing UK wide expert research data services*
- Improve the quality, breadth and depth of datasets and provide a repeatable curation approach that can be used to enable UK-scale research.*
- Create globally leading data sets and make these discoverable and accessible to researchers and other users in areas of greatest opportunity. The Hub will be expected to define the key research questions and needs that will drive its strategic approach to data selection.*
- Provide services to meet the needs of researchers, NHS, industry, charities, innovators and ultimately patients and the public.*
- Enable real world insights, innovations and improvements in care not currently possible without data at this scale.*
- Improve secure access to data through membership of the UK Health Data Research Alliance (the 'Alliance') and use of the UK Health Data Research Innovation Gateway (the 'Gateway').*
- Build on the outputs and learnings generated by the MRC Data Pathfinder awards as appropriate and link to/complement existing investments, nationally and internationally. Harness opportunities to work across the HDR UK Hub network.*
- Involve PPI in development of the Hub to ensure it is driven by the needs of the population and considers key ethical issues pertinent to mental health data.*
- Have strong leadership, deep domain expertise and governance, including plans for including patients and members of the public in decision-making and governance and how the Hub will ensure that research outcomes from its work are linked to patient benefits.*
- Contribute to efforts for harmonisation of procedures for prospective measurement of mental disorders and symptoms in cohorts and longitudinal studies and measurement of the outcomes of interventions.*
- Designed with capacity to evolve and be responsive to changing needs and opportunities. Incorporate plans for sustainability of the investment, including through grants and service provision.*
- Meet delivery milestones and to sign up to the [Digital Innovation Hub Programme Principles for Participation](#) – this includes the FAIR data principles (Findable, Accessible, Interoperable and Reusable).*

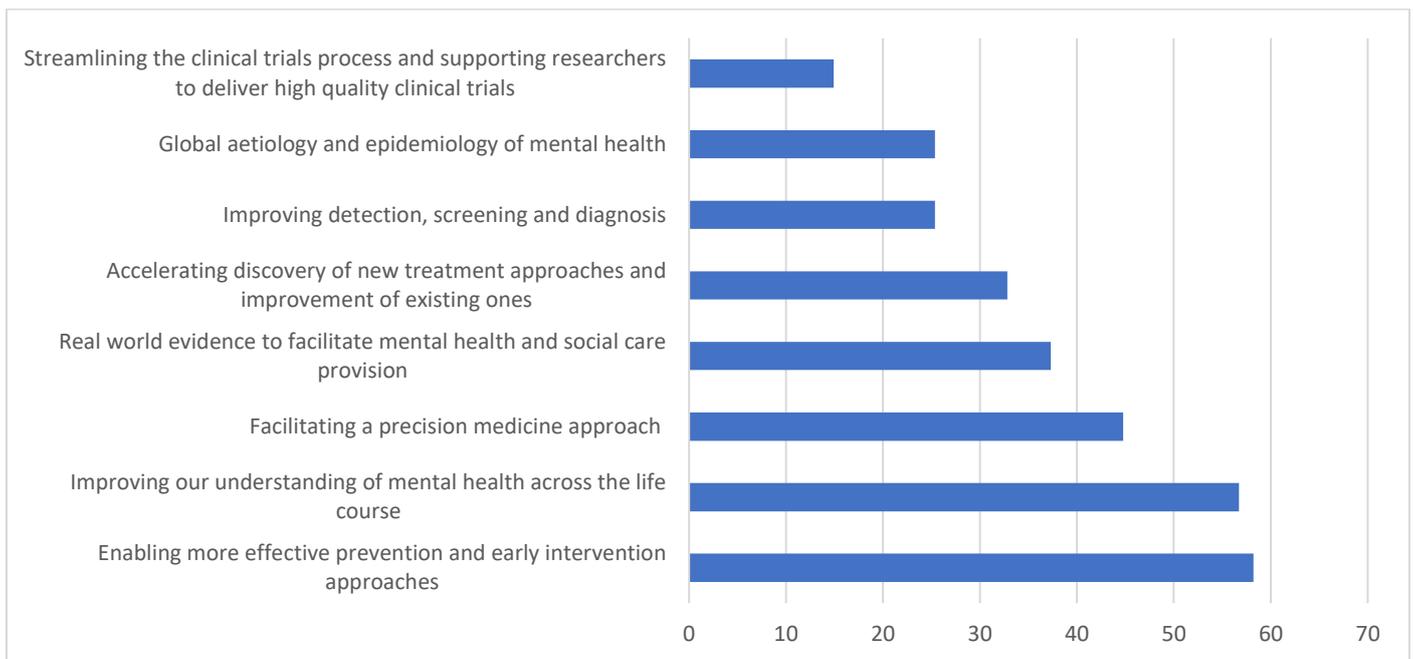
*The purpose of the Hubs is **not** to perform original research. The funding available to applicants will be to create "data as infrastructure" that improves access to and quality of data and increases the UK's future research and innovation capability. While Hubs will support research and innovation activities through the creation of new data and services they provide, the funding from this call will not support specific research projects.*

Questions

It is proposed that the Mental Health Data Hub should initially prioritise the data it makes available and the services it provides based on a defined vision that addresses areas of greatest need/opportunity. Please consider the areas of focus below and prioritise which could be most impactfully advanced by increased access and usability of data across multiple sources during the 3 year establishment phase of the Hub. Each of these areas align with the [MRC Strategy for Lifelong Mental Health Research](#) and with prior work identifying opportunities for mental health data science (including a [paper](#) published by the MQ Data Science group in 2019).

Please select up to three areas. Percentage of respondents who selected each of the areas are shown in the graph below.

MRC Mental Health Data Research Hub Survey



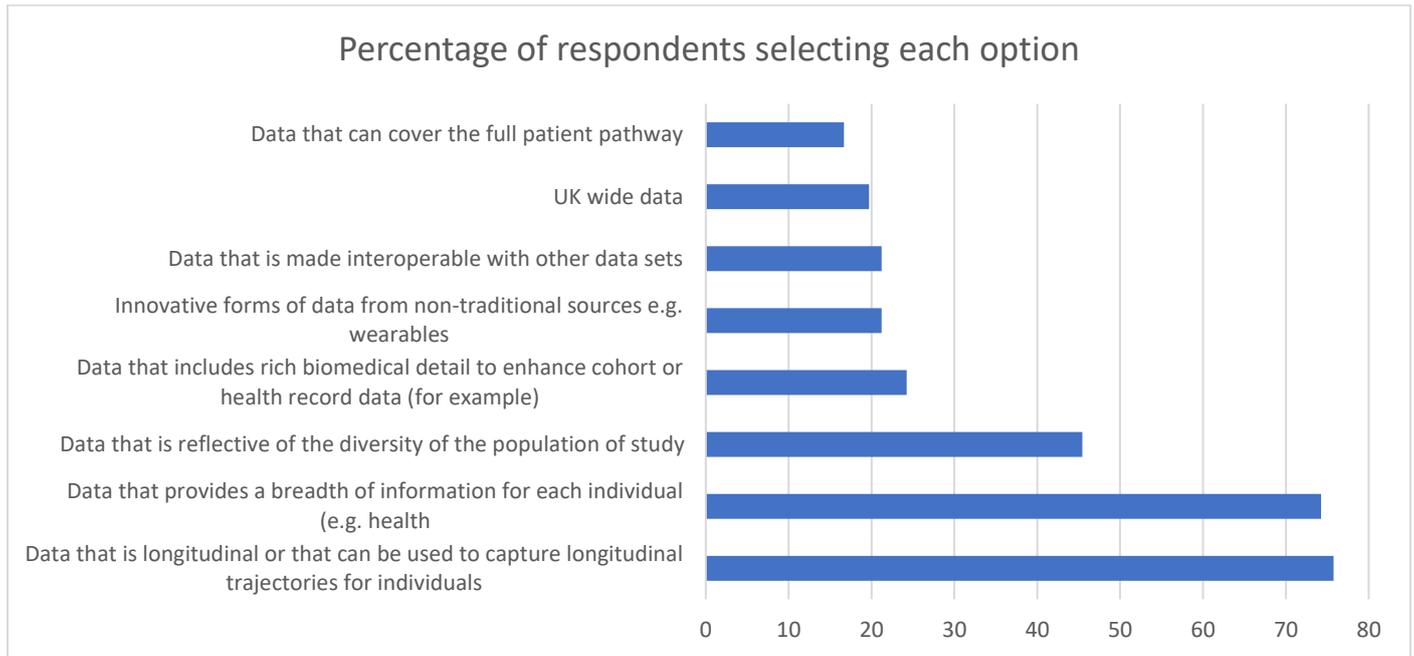
Please outline your rationale for your selection (optional, free text)

- Major opportunity in precision/stratified mental health.
- Understanding how best to target existing treatments – high potential from routine data to progress from the current ‘one size fits all’ approach.
- Data from those with severe mental illnesses represents an unmet need. These groups are poorly represented in population cohorts. Heterogeneity in these conditions argues for precision medicine approach.
- Prevention is the key step to improving outcomes of mental disorders and, combined with early intervention, is the most cost-effective approach.
- Supporting systems approaches to prevention through linked datasets about people, places and communities is important.
- Evidence across the life course from multisectoral data linkages is necessary for prevention.
- Early intervention is critical but there is a need to be able to identify who is at risk. Understanding the causal determinants of poor mental health is very important, particularly in early life, to evidence these interventions.
- We know little about how psychopathology changes across the lifespan.
- A life course approach to aetiology allows for more resource efficiency and personalisation.
- Access and linkage of multiple data sources allows the reconstruction of individual life courses at a population scale. This is extremely powerful for life course research and can produce causal evidence for prevention and intervention.
- Data from young people is important given that a number of mental health problems emerge during this life stage.
- Opportunities to capitalise on advances in genomic progress.
- Longitudinal and clinical cohorts are readily available in the UK and combining these would create powerful datasets.
- There is a need for discovery of more effective interventions
- Delivery of mental health care can be supported by harnessing the vast amounts of real-time routine clinical data, linked to data about people and places.
- Data can be used to help improve overall mental health management as well as intervention and prevention.

MRC Mental Health Data Research Hub Survey

- Learning from how existing digital mental health technology helps people, searching that for patterns of insight for new treatments and personalisation in partnership with users.

The initial data sets to be included within the Hub will be driven by the prioritised challenges above. However, please comment on what data qualities would add most value to mental health research for the themes you selected in the previous question. Please select up to three responses. The graph below shows the percentage of respondents selecting each data attribute option.



- Other (2%)

Innovative forms of data from non-traditional sources e.g. humanoid robots.

Please outline your rationale for your selection (optional, free text)

- UK wide data
 - UK wide data would allow interrogation of regional variation, which is crucial to understand mental health inequalities.
 - UK wide to provide sufficient data.
 - UK-wide collaboration is vital.
- Data that is reflective of the diversity of the population of study
 - Those with the most severe forms of mental illness are under and mis-represented within existing large-scale cohorts and to supplement efforts to represent this group would be valuable.
 - Overcoming limitations and biases due to institutional and clinical selection
 - Great to bring together biological data with other types of data, but if samples are not representative a lot of data are wasted (e.g. genetic data on ethnic minorities even when collected often not used in psychiatric genetics studies).
 - Increasing diversity in our samples is a high priority as we don't do that well in many longitudinal British cohorts at present.
 - Arguably results from any study simply reflect the sample available - therefore we need whole population data sets with little missingness, high quality/reliability of variables and as many potential causal or confounding or mediating variables as possible so we can rely on results with an acceptable degree of certainty

MRC Mental Health Data Research Hub Survey

- Diversity is important for healthcare equality and is lacking to date.
- Electronic health records offer a real possibility to understand health outcomes in people who are under-represented in research, e.g. people with severe mental illnesses and multimorbidity's, ethnic minority groups etc.
- Population studies are often prone to bias, excluding the more vulnerable. It is crucial to develop more effective processes for collecting, sharing, linking and analysing diverse and informative population data whilst protecting the identity of the young data subjects.
- Diversity is lacking in many datasets and needs improving.
- Data relating to mental health has been unrepresentative of the population at large, and it should be more representative of underserved populations.
- Need to ensure that the digital divide doesn't create further health inequalities, so covering the hardly reached and learning from people who don't use NHS services but may use others or self-care.
- Data that can cover the full patient pathway/Data that is longitudinal or that can be used to capture longitudinal trajectories for individuals
 - Rich attributes and longitudinal both important to understand the cumulative impact of life circumstances on mental health trajectories over time.
 - Clinical trials can only tell us about large treatment effects (due to limited power) over short timespans (weeks or months, not years or decades). However, treatments are often taken for decades and the most important treatment outcomes (such as death) are rare and require decades of follow up.
 - We lack longitudinal data which is essential to understand trajectories.
 - Linkages between GP, mental health trusts, NHS hospital trusts, and non-NHS mental health providers would generate evidence on the totality of mental health care across the life course.
 - Better long-term data is needed to improve long term outcomes.
 - Longitudinal in relation to importance of lifespan and intergenerational impacts.
 - Data is very limited in value if not longitudinal.
 - Rich attributes and longitudinal both important to understand the cumulative impact of life circumstances on mental health trajectories
 - Longitudinal data are essential, linked routine data characterize the visible part of the clinical iceberg, cohort data characterize the invisible part, linkage between the two gives the whole picture.
 - Longitudinal understanding of pathways and how these can be improved is important, and interoperability allows this.
 - In determining causal mechanisms longitudinal data capture will be key.
- Data that provides a breadth of information for each individual (e.g. health, administrative, social, biological)
 - Data that is already linked with biological, social and service access would be most helpful in providing a continuous understanding of the individual.
 - Data should be integrated where possible and should cover as many different aspects of study as possible.
 - The key to exploiting the power of data is completeness, linkage and chronology.
 - These problems are caused by a host of factors and they develop over time. Multifaceted longitudinal studies are essential to their understanding.
 - It is imperative that, if at all possible, social and environmental factors such as income/employment status are captured within the data hub in a longitudinal sense to allow for causal inference.
 - Health is largely dependent on social and environmental aspects of life.
 - All aspects of an individual need to be considered due to the unique presentation of mental health disorders.
 - EHRs frequently lack detail on social indicators therefore longitudinal linked datasets provide an ideal possibility.
 - We need breadth to be able to link across different datasets to disentangle bio and social determinants.
 - Knowledge advancement is most likely to arise from breadth of information.

MRC Mental Health Data Research Hub Survey

- More ambitious cross-sectoral linkages (eg, MH data from school pupils linked to exams, social security, parental health records, prescribing).
- A wide range of datasets collected over long time periods offer the best chance of determining association of factors with mental health.
- Data that is already linked with biological, social and service access would be most helpful in providing a continuous understanding of the individual.
- Data should reflect the multi-dimensional determinants of mental health - biological, social, psychological and cultural.
- Data that includes rich biomedical detail to enhance cohort or health record data (for example)
 - There is a real need for linkage to biological data for mechanistic understanding etc
 - The community craves linked biological and longitudinal (or lifetime) clinical measures and outcomes and this would provide a game changing research resource.
 - I would focus on biomedical data and ability to link.
 - Biomedical detail is most likely to be a game-changer in international leadership.
- Data that is made interoperable with other data sets
 - Interoperability to future-proof data and drive future projects.
 - Interoperability generally important - we don't want mental and physical health data silos when the two are so closely linked
 - interoperability is key (but should be trivial).
- Innovative forms of data from non-traditional sources e.g. wearables, apps
 - The problems with this data as it stands is often that, it isn't detailed and we often rely on proxies for certain things. Better if we could link with more data across/beyond health - with ethical input from service users of course.
 - The Hub must be cutting edge in terms of new forms of data and be innovative in establishing quality standards.
 - Data collection via wearables, actigraphy, ecological momentary assessments on mobile phones and light sensors - these data could have real value in many different areas.
 - Engineers and computer scientists can play an important role in generating valuable data including data from wearable and fidgeting tools

The Hub will make data discoverable via the HDR UK Gateway and will also curate data and provide services to improve its use for research. What services should the Hub make available to users? Please provide a ranked order (with 1 being the highest priority). The below reflect the average ranking order by respondents

- Establishing key data linkages to allow key outcomes to be identified and a more comprehensive picture of the factors influencing mental health. (3.1)
- Harmonising key data sets to enhance their interoperability. (3.9)
- Help navigate UK mental health data and services. Assist with regulatory and data access procedures. (4.1)
- Facilitating multi-site data research across the UK. (4.4)
- Providing (and developing) open access tools and methods for use in mental health data science. (4.9)
- Supplementing data sets e.g. with annotation of mental health related data. (5.2)
- PPI engagement workshops. Bringing the lived experience perspective to facilitate legitimate data access requests. (5.5)
- Work with data controllers and custodians to make safe sharing of mental health data easier (5.6)
- Providing test and validation data sets. (5.8)
- Assess feasibility and provide support on study design and analysis. (6.2)
- Curation of bespoke data sets. (6.6)
- Identification of individuals for clinical trials or research studies (6.8)
- Other – please specify
 - To sign point researchers to the current data available for mental health research

MRC Mental Health Data Research Hub Survey

- Broader national consensus created to establish a social licence for use of linked anonymised or otherwise mental health datasets
- Supporting service user researchers
- innovation - looking beyond the four walls!
- Easing the processing of developing data collection tools and wearables

If a reasonable cost was associated with these services, would you be happy to pay it? (Yes/No/Don't know).

Percentage of respondents who selected this alternative.

- Yes – 54%
- Don't know - 33%
- No – 13%

It is intended that the Hub infrastructure will advance the potential for mental health data science broadly, accelerating impact for research and innovation across the UK. Where do you consider the biggest potential for impact from this investment could be? What will be critical to success? (free text, optional)

Real world impacts

- It needs to impact the delivery of treatment, not just answer academic questions without real world impact
- Needs to have clinical applicability and utility as an overriding focus (i.e. driven by strong use cases, and ideally following paths with demonstrable previous productivity).
- It is important that practical translation into knowledge generation and clinical care improvement is central to any initiative, with clarity on current bottlenecks that can be most readily overcome.
- The biggest potential for impact is around the better classification and treatment of severe mental health conditions given coverage of more common mental disorders within existing infrastructures (eg UK Biobank).
- Personalised EFFECTIVE treatments.
- Helping to address mental health inequalities
- Improvement in the identification, diagnosis and treatment along pathways of care.
- Better understanding, prevention and treatment of mental disorders
- Demonstrable improvements in service provisions and pathways
- Improve population mental health through clearer understanding of stratification of risk and resilience.
- Understanding origins of disorders, identifying therapeutic targets that influence disease progression, and which populations require targeted-stratified intervention.
- Generating higher quality causal evidence on the relationship between social factors and mental health outcomes, both across population groups and between regions within the UK. This is hugely important in strengthening the evidence base for policies which act on the social determinants of mental health to reduce health inequalities and would be of direct interest to policymakers in these areas.
- Providing policy makers and other stakeholders with new and more informative resources (there is a current dependence on limited subsets due to challenges in utilisation).
- Policy-academic coproduction efforts. Allows us to create mental health focused 'what works' modelling and work with governments to consider policy solutions.
- improve cost effectiveness of services by reducing health inequalities through data science.

Stimulating further investment

- Encouraging crowdfunding for research
- Dramatically enhancing commercial (not just pharmaceutical) investment in mental health discovery/use for clinical trial recruitment of truly representative samples.

Improving access to data

- More streamlined and timely access to data

MRC Mental Health Data Research Hub Survey

- Fast, inexpensive access to publicly generated datasets.

Improved data

- Providing linked biological, social and longitudinal clinical data
- The main benefits are likely to be scale and facilitating linkage/approvals
- Addressing barriers to the use of routine clinical data, including data quality, completeness and interoperability
- There remains huge potential in the breadth and depth of NHS data resources, particularly when linked to administrative data and enhanced with meta-data and derived detail (including linkages between healthcare sectors, and ideally with social care as well). We currently have national-level resources with limited depth and local-level resources with limited breadth.
- More work on the linkage of routine healthcare and prescribing data to administrative data from other sectors
- A coordinated approach to the large-scale collection and genotyping of DNA for psychiatric genomics research (and allowing access to these data to researchers who may not have had resources to collect the samples);
- The ability to easily integrate biological, particularly genomic, and longitudinal (from childhood) clinical data importantly including outcomes of mental and physical health

Public and patient involvement and engagement

- PPI work across the UK to ensure public trust in the sharing of mental health data
- Participants in studies can see results as they emerge; encourage 'data for all' approaches and citizen science;
- Researchers being able to approach participants directly through a linked portal for consent
- Empowering citizen science by 1) encouraging donation of app usage and wearables data 2) supporting service users to lead research

Factors critical to success

Collaboration

- Better collaboration across the UK.
- Helping to develop multi-site from single-site resources would be transformational.
- Understanding that we're all in this together & that nobody wins unless we all win
- UK wide collaboration building on the work of the Pathfinders and open to new centres and researchers
- The Hub needs to be inclusive across the country and not be set up in a way (or costed) that disadvantages research centres.

Data access

- Transparency (who gets access, how, and what they are aiming to do with the data)
- Study PIs being willing to share their data. Which means the Hub infrastructure needs buy in from those PIs.
- It needs to be accessible to clinical researchers without requiring extensive big data training.
- Straightforward procedure for access.
- Keeping datasets up to date.
- Open access data.
- Critical to this success will be overcoming the regulatory hurdles to link, share and access data.
- Development of more streamlined legal and security safeguards to prevent data breaches, to adequately cover future situations in which data shared turn out to be more identifiable than first assumed.
- High quality proposals- peer reviewed analysis plans and protocols.

MRC Mental Health Data Research Hub Survey

- Simplify data linkage processes across the UK.
- Interoperability, discovery, and security of data are key

Sources of data

- Balance of biological and social approaches, common and severe mental disorders and life course stages.
- Focus on a cross-disorder approach.
- Ensuring the data included can be harmonised (and is actually equivalent).
- Clean, complete data with detailed (and easy to read) data dictionaries.
- A collaborative effort to curate and make available mental health data collected in schools to drive forward adolescent mental health research.
- Representativeness across UK and minority groups.
- Reliable variables covering regional and individual level key confounders.
- Regionally sensitive and agile additions of data
- Linking of mental and physical health data for multimorbidity research

Engagement with key stakeholders

- Effective communication with key stakeholders, users and policy makers. Ensuring that its developments are effectively driven by stakeholder needs (i.e. a two-way process).
- Critical to success is to work closely with all data custodians of large mental health data sets.
- The work needs to be undertaken with clear buy-in from data providers
- Sharing of expertise between data providers (those already providing and those who would like to be more active in this area).
- The thing that will be critical for success is public buy-in and the ability to opt out -- ideally centrally (as preferred by patients and the public)

Skills

- Enhancing data science capacity and career paths is essential. We have a national skill shortage and problems with retention,
- Collaboration, inclusion, inter-disciplinary and investment in ECRs who are the future of mental health research.
- To involve Data science experts, technologists and engineers in the steps to help in developing an effective data analytics and provide valuable datasets based on innovative data collection tools.
- Likely to require up-skilling and investment in NHS digital services staff.