



Project Number:	284941	Project Title:	Human Brain Project
Document Title:	The HBP foresight lab: first report on future medicine (report)		
Document Filename:	SP12_D12.1.1_ForesightLab_FirstReport_FINAL		
Deliverable Number:	D12.1.1		
Deliverable Type:	Report		
Work Package(s):	WP12.1		
Dissemination Level:	PU = Public		
Planned Delivery Date:	Month 18 / 31 March 2015		
Actual Delivery Date:	Month 18 / 31 March 2015		
SP Leaders:	Jean-Pierre CHANGEUX, IP (P27), Kathinka EVERS, UU (P82)		
SP Manager:	Benjamin SIMMENAUER, IP (P27)		
Authors:	Nikolas ROSE, KCL (P32); Christine AICARDI, KCL (P32); Michael REINSBOROUGH, KCL (P32)		
STO Review:	UHEI (P45): Björn KINDLER, Sabine SCHNEIDER, Martina SCHMALHOLZ		
Editorial Review:	EPFL (P1): Richard WALKER, Guy WILLIS, Celia LUTERBACHER		
Abstract:	This report is the first Deliverable produced by the HBP Foresight Lab at King's College London (P32, WP12.1). The Foresight Report on Future Medicine outlines the major societal and ethical challenges faced by the developers of the HBP Medical Informatics Platform (MIP) in their work to federate data and translate key health findings into clinical practice. The report focuses on two key issues (data federation and privacy, plus disease signatures and personalised medicine) and it provides recommendations. In particular, we have developed a set of scenarios, based on consultation with different stakeholders, which provide frameworks for evaluating the possible impacts of the HBP research in different areas of society.		
Keywords:	Ethics, RRI, Medical Informatics Platform, biological signatures of disease.		



Table of Contents

Executive Summary	3
1. Introduction.....	5
1.1 What is the HBP?	5
1.2 The Research: King's College London and the Foresight Lab	6
1.3 Why Responsible Research and Innovation (RRI)?	6
1.4 What is Foresight?.....	8
1.5 Future Medicine in the HBP and the Focus of this Report	9
1.6 Aims and Methods of this Report	10
2. Future Medicine: Background and Challenges	11
2.1 The Work of the Medical Informatics Platform of the HBP	11
2.2 Data Federation, Data Protection and Data Privacy	14
2.3 The Problems of Diagnosis and the Search for Disease Signatures	15
3. Data Federation and Privacy: Foresight	17
3.1 Background Scenarios	17
3.2 Emerging Challenges	18
3.3 Recommendations	24
4. Disease Signatures: Foresight	26
4.1 Background Scenarios	26
4.2 Emerging Challenges	27
4.3 Recommendations	34
5. Acknowledgments	36
6. Annex 1: The Scenarios	37
6.1 Data federation and privacy	37
6.2 Disease Signatures	40
7. Annex 2: The Medical Informatics Platform	44
8. Annex 3: Collaboration with the Danish Board of Technology Foundation.....	49
8.1 A Collaborative Approach.....	49
8.2 Webinars.....	50
9. References	61

Index of Figures and Tables

Figure 1: The technical architecture of the Medical Informatics Platform 14

Figure 2: This illustrates several layers of organisation around individual patients in a form similar to layered structures in geographic information services (GIS) mapping. Image from *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*, National Academy of Sciences, 2011 p15. 32

Executive Summary

The Human Brain Project (HBP) is one of the Future and Emerging Technology Flagship initiatives, which are funded by the European Commission and promote ambitious and science-driven research initiatives. The HBP is a 10-year project in medicine, neuroscience and computing that brings together scientists and institutions from 20 nations across Europe.

As part of this, the builders of the HBP Medical Informatics Platform (MIP) are working to identify the biological basis of brain diseases, with the long-term goal of informing the development of new treatments for neurodegenerative and psychiatric disorders. Developers of the HBP MIP are federating medical data from European hospitals, and deploying statistical clustering strategies to identify the brain signatures of these disorders.

This report by the HBP Foresight Lab at King's College London (Work Package 12.1) outlines the major societal and ethical challenges faced by MIP developers as they work to federate data and translate key health findings into clinical practice. It focuses on two key issues:

- Data federation and privacy
- Disease signatures and personalised medicine

Data protection and data privacy present some major challenges for the federated organisation of the HBP Medical Informatics Platform, which our Foresight research suggests will be key to the future of the Project. We identify three such challenges that are interwoven in the multi-layered architecture of the MIP: legality, trustworthiness, and privacy. Some of these challenges may be addressed by measures for technology management; others may be addressed via community-building activities around the MIP. These activities could involve clinicians, the pharmaceutical industry and other professional stakeholders, as well as patient groups and the general public.

Based on our research and our discussions with MIP developers, we recommend the following:

- 1) Regular and systematic work on scenarios for misuse.
- 2) An Information Security Architecture Partnering project within the MIP, to be established before the end of the Ramp-Up Phase, which would preferably be extended to the whole HBP, if such provision is not already in place.
- 3) A Privacy Impact Assessment.
- 4) Evaluation of consent requirements for different functions of the MIP, with a focus on informed consent wherever appropriate.
- 5) Consideration of special issues that may arise from the use of data acquired from outside the EU.
- 6) Protocols for engaging patient, patient support and stakeholder groups in respect of the different functions of the MIP.
- 7) A Public Engagement and Research Dissemination Plan.
- 8) A Data Governance Committee for the MIP with a broad membership including representatives of users and the public.
- 9) A Research Audit structure that can identify, authorise and audit all users of the MIP.



Many challenges are raised by the interpretation of complex biological and clinical data to identify signatures that may be clinically useful in the diagnosis and treatment of mental disorders. Further challenges arise if such signatures are used to identify predispositions or susceptibilities to disorders. These challenges include: recognition and interpretation of statistical clusters within data; identification of when a particular biological signature is an indicator for a disease; translation of findings into protocols and guidelines for clinical practice; ensuring appropriate use of disease signatures by clinicians and patients, and preparing the regulatory and governance infrastructure for what is often termed 'personalised' or 'precision' medicine. Many of these can only be addressed adequately by strong consultation and engagement with affected communities.

On the basis of our Foresight work, we make the following recommendations to ensure that an effective partnership is created with patients, clinicians and other potential users of disease signatures:

- 1) Include patients and clinicians in a research advisory capacity.
- 2) Address implications for clinical practice, in particular by engaging clinicians in the assessment and verification of disease signatures and their utility in clinical applications.
- 3) Address implications of disease signatures for clinical ethics.
- 4) Reflect on the use of brain signatures in the identification of pre-clinical disease susceptibilities.
- 5) Ensure awareness that differences in brain structure or function do not equal brain disease.
- 6) Develop effective communication strategies to explain the potential clinical and research uses of brain signatures to the public at large and other audiences.
- 7) Link with other research communities and relevant regulators to develop appropriate pathways for translation of research to clinical applications.
- 8) Consider the wider implications of moving to a brain-based understanding of disorders
- 9) Consider the implications for public trust and support of the HBP, of the future intended commercialisation of the findings to generate employment and wealth creation.

These recommendations have arisen from foresight work with multiple sources, including: reviews of relevant literature and reports from professional bodies; discussions with stakeholders in formal and informal settings; webinars and seminars organised in collaboration with our partners in the Danish Board of Technology; and productive discussions with the directors and researchers of the MIP itself. Whereas we have already incorporated some of our recommendations into the developing MIP, some will require action by others as the work of the Human Brain Project progresses. These collaborations embody and enact the principles of Responsible Research and Innovation, and together aim to ensure that the work of the HBP is directed most fruitfully to meet the major challenges that psychiatric and neurological disorders pose to individuals, families, communities and societies.



1. Introduction

1.1 What is the HBP?

Understanding the human brain is one of the greatest scientific challenges in the 21st century. The Human Brain Project is based on the consideration that, by progressing in our research, we will be able to gain insights into what makes us human, to develop new treatments for brain disease and to progress in building revolutionary new computing technologies. And it argues that developments in Information and Communication Technology, for the first time, bring these goals within sight.¹

The HBP is a 10-year project co-funded by the European Commission². The Project, which will have a total budget of over EUR 1 billion, is European-led and has a strong element of international cooperation. It has grown since its initial proposal and currently consists of a Consortium of over 100 partner universities and companies, covering more than 20 countries in Europe and beyond. When it was launched in 2013 as one of the two Future and Emerging Technologies (FET) initiatives,³ the European Commission welcomed the Human Brain Project as:

“The world's largest experimental facility for developing the most detailed model of the brain, for studying how the human brain works and ultimately to develop personalised treatment of neurological and related diseases. This research lays the scientific and technical foundations for medical progress that has the potential to improve the quality of life for millions of Europeans.”

Initial funding for the HBP was given for a 30-month Ramp-Up Phase, with further negotiations envisaged for the funding of the remaining seven and a half years of the Project.

One of the major obstacles to understanding the human brain is the fragmentation of brain research and the data it produces. The HBP intends to gather international knowledge and advances in emerging information and communication technologies (ICTs) in order to integrate this data in a unified picture of the brain as a single multi-level system. The ultimate goal is to integrate existing knowledge in massive databases and in computer models of the brain.

The HBP has three main “Flagship Strategic Objectives”:

Future Medicine: Develop an objective, biologically grounded map of neurological and psychiatric diseases based on multilevel clinical data; use the map to classify and diagnose brain diseases and to configure models of these diseases; use *in silico* experimentation to understand the causes of brain diseases and develop new drugs and other treatments; establish personalised medicine for neurology and psychiatry.

Future Neuroscience: Achieve a unified, multi-level understanding of the human brain that integrates data and knowledge about the healthy and diseased brain across all levels of biological organisation, from genes to behaviour; establish *in silico* experimentation as a foundational methodology for understanding the brain.

Future Computing: Develop novel neuromorphic and neurorobotic technologies based on the brain's circuitry and computing principles; develop supercomputing technologies for brain simulation, robot and autonomous systems control and other data intensive applications.



1.2 The Research: King's College London and the Foresight Lab

From the initial conception of the Project, it became clear that many social and ethical issues would arise from the research. This is why an “Ethics and Society Programme” was included as an integral part of the Project. The aims of this programme are: to explore the Project's social, ethical and philosophical dimensions; to promote engagement with decision-makers and the general public; and, more generally, to foster responsible research and innovation by raising social and ethical awareness among Project participants.

The Department of Social Science, Health and Medicine at King's College London is part of this programme. We identify and evaluate the potential impact of the new knowledge and technologies produced by the HBP, in terms of benefits to European citizens, European industry and European society. To fulfil this aim, we established the King's Foresight Laboratory.

The Foresight Lab conducts systematic foresight exercises to identify and evaluate these impacts. We do this by adapting established foresight methods already in use in different areas of medicine and ICT, including modelling, horizon scanning and scenario planning. These methods involve recurring consultations with researchers, potential industrial and professional users of new technologies, civil society groups, regulators and other stakeholders. In particular, we have developed a set of scenarios, based on consultation with researchers, potential users of new technologies, civil society groups, regulators and other stakeholders, which will serve as frameworks and stimuli for evaluating the possible consequences of the HBP on different areas of society.

1.3 Why Responsible Research and Innovation (RRI)?

The demand for responsible research and innovation⁴ has become widespread among research and development funding agencies across Europe. At its simplest, RRI is linked to the growing belief among policy-makers that funded research should address grand societal challenges. RRI thus aims to shape the research and development trajectory of emerging technologies, to ensure that they remain aligned with the values, needs and priorities of society at large. This requires everyone to consider the potential risks of their work and how these could be identified and mitigated, as well as the short, medium and longer-term societal consequences of their work. An RRI approach also encourages citizens to be aware of the day-to-day process by which research is done, and of decisions that may have unforeseen consequences. If key aspects of our collective future are born in the laboratory, then from the perspective of RRI, citizens should pay attention to what is happening there.

“Responsible Research and Innovation (RRI) is the on-going process of aligning research and innovation to the values, needs and expectations of society.”

Source: Rome Declaration on Responsible Research and Innovation in Europe, October 2015

Of course, there are many difficulties with such aspirations. In many cases, it is only well after the introduction of any particular technology that there is a full understanding of its social, economic, environmental and health effects. This often comes only after these technologies have established themselves within the interlocking technical systems and infrastructure of society, i.e., when it may be difficult or expensive to remove or change them.⁵



In part, the demand for RRI in emerging technologies arises from the view that broadening the pool of decision-makers involved in introducing a new technology/health practice/medicine to include those directly affected by the technologies (lay input) might lead to better anticipation of potential adverse effects. In addition, RRI techniques may improve scientific researchers' capacity to consider such consequences early in the development pathway, allowing them to make better-informed judgments. Likewise, RRI is associated with improving the capacity of the public to understand contemporary developments in science and technology, so that they are better able to participate in research governance. Lastly, increasing the number and diversity of persons contributing to technology decision-making may result in better design outcomes, creating more equitable technology that meets the needs of a greater diversity of people.

RRI approaches encourage reflection among researchers, funders, specific stakeholder groups and citizens more generally, and feed that reflection back into the development pathway to shape research in the direction of public benefit. One UK-based science and policy funding body has helpfully characterised the RRI approach in terms of four dynamically interacting phases: Anticipation, Reflection, Engagement and Action (AREA):⁶

- **Anticipation:** Describing and analysing the impacts, intended or otherwise, (e.g., economic, social, environmental) that might arise. Not to predict but to explore both anticipated and unanticipated impacts and implications.
- **Reflection:** Reflecting on purposes, motivations and potential implications of the research, and the associated uncertainties, areas of ignorance, assumptions, framings, questions, dilemmas and social transformations these may bring.
- **Engagement:** Opening up such visions, impacts and questioning to broader deliberation, dialogue, engagement and debate in an inclusive way.
- **Action:** Using all of the above to influence the trajectory of the research and innovation process itself.

The phases are not distinct in timing, but necessarily integrated across the whole pathway of research, development and (where appropriate) application or commercialisation. They are also intended to be capacity building for all members of the community. Anticipation occurs by engaging researchers with stakeholders and experts to think reflectively about the research system and their position within it. Possible outcomes become apparent when people from different parts of the research system interact. For example, patient groups may be aware of issues from the perspective of the patient that a researcher not involved in clinical work, or even a clinician, might not have realised. Patients, by hearing what researchers are planning, have their first opportunity to reflect upon the proposals and give feedback. In these cross engagements, the capacity to act by modulating the priorities of research or the design of intended health technologies is developed.

While RRI may increase reflexivity among individual researchers, it does not focus only on the individual. RRI does not require individuals to anticipate research outcomes from their single vantage point, reflect upon these and engage the public, and then take action to restructure the research system. Enhancing the capacity of researchers is, in part, a side effect of bringing together actors from different parts of the research process, including bench scientists, stakeholders and members of the public, to develop the collective capacity to shape research strategy and system in a way that is informed by the dialogue between all parties. There is no guarantee that all parties will agree or that consensus will be the outcome of such a process. However, innovation choices (even the choice not to go forward with a proposed innovation) can in principle be better informed, and their likely social benefits enhanced, through the dialogue created by an RRI process.



Developing an RRI process that successfully embodies these principles is difficult. It relies on the good faith of participants, requires adequate resources, and must be carefully adapted to the particular circumstances of the research situation and relevant stakeholders. Developing and testing successful RRI practices is part of the greater research agenda of the Human Brain Project.

1.4 What is Foresight?

Foresight is the “practice of making ‘forward looks,’ of anticipating change, or making studies of future possibilities.”⁷ Foresight differs from forecasting, the most common approach to the future used by engineers. Foresight considers multiple possible futures, whereas forecasting is an attempt (often using numerical calculation) to predict the single most likely future.

“...Scenarios - i.e. many futures - and forecasting - one future - have different ontological and epistemological underpinnings. ...In forecasting, the emphasis is on what is knowable in advance from evidence of the past. Uncertainty is treated as a ‘lack of knowledge’. In periods characterised by rapid and stable growth, forecasting has proved to be a reliable approach to predicting the future. In situations characterised by complexity, turbulence and ambiguity, over-reliance on forecasting can be a fatal error.”⁸

A forecasting approach might describe uncertainty as ignorance or lack of knowledge, might expect or consider only one future, and might deal only with explicit worldviews. A foresight approach, however, sees uncertainty as intrinsic. There are many possible futures, and actions in the present (and indeed decisions taken at any point along the research and development pathway, as they interact with a host of external, contingent or loosely connected events) will shape which futures come into existence.

Foresight also recognises that underlying values and knowledge influence how we perceive our past and present. Foresight aims to consider how such values may influence our present, past, and future. This is particularly important in multiple stakeholder communities, where values may vary within the collective of groups that must consider possible futures. Foresight is often participatory, involving many constituencies and not just experts and specialists. The additional or alternate insight from differing constituencies is made possible by a foresight approach that takes seriously the differing values within and between communities.

To anticipate possible futures is to consider them now, in the present, and to prepare. Anticipatory approaches to foresight emphasise capacity building, either individually or collectively (societally or organisationally). Because we cannot be certain which of many possible futures we will experience, anticipatory foresight is a type of engagement with communities or individuals that helps them build, or begin to consider, the types of capacity that they might require to be successful in differing future scenarios.⁹

To be useful, foresight must be grounded in empirical research. Foresight is different from speculative ethics, where many of the situations considered (regardless of the intellectual value of the considerations in themselves) are often implausible or very far in the future, making their relevance low. Science fiction approaches do not serve the interest of effective science policy development.

Foresight is often commissioned as part of policy work. Foresight approaches have been used in areas as diverse as addiction and drugs, and photovoltaics.¹⁰ The foresight timeframe suitable for policy work is usually quite close (in our case 5-15 years). Over longer timeframes, it becomes increasingly hard to provide empirical backing to justify the

range, amount and direction of variability that the future may have; the work becomes more speculative, and hence has less value.

Foresight uses a range of methods.¹¹ These include Scenario Planning, Delphi studies, Horizon/Environmental scanning, back casting, judgements made from the previous historical range of plausible outcomes, and sometimes road mapping or (product) life-cycle analysis, particularly anticipatory life-cycle analysis.

Foresight is only one element within RRI; it should be linked to variety of other aspects of technology assessment. As such, foresight must be integrated with public engagement, reflection, as well as the actions of researchers, stakeholders, and publics in responding to anticipations of some particular future.

1.5 Future Medicine in the HBP and the Focus of this Report

The Human Brain Project has chosen Future Medicine as one of its three main research areas, the others being Future Neuroscience and Future Computing. Future Medicine is the focus of our first report. In the HBP, researchers in Future Medicine are developing an objective, biologically grounded map of neurological and psychiatric diseases based on multi-level clinical data. They are using this map to classify and diagnose brain diseases, and to configure models of these diseases. They are also using *in silico* experimentation to understand the causes of brain diseases and develop new drugs and other treatments, and they are establishing personalised medicine approaches for neurology and psychiatry.

This work takes, as its primary data, clinical and hospital records of patients treated for psychiatric and neurodegenerative diseases across Europe and other regions. The underpinning hypothesis is that it is possible to aggregate and ‘mine’ these data to derive objective biological signatures of brain disease. These will be used for diagnosis, more accurate prognosis and to open up new types of drug discovery pathway for the development of new medicines. The aim is to identify distinctive patterns of data from brain scans, genetic tests and other indicators that will distinguish individual disorders more accurately than current techniques, which are largely based on classification by observable symptoms rather than underlying pathology. Such ‘disease signatures’ could enable more accurate prognosis and diagnosis of individual patients in the future, based on a better understanding of their specific neuropathology, thus also enabling more accurate choice of the treatment that is thought to be most appropriate and effective for specific persons and their conditions. This is what is meant by ‘personalisation’ in this context. Furthermore, in the future, it is hoped that it might be possible to identify patients at risk on the basis of these disease signatures, before the disorder becomes manifest or at an early stage in its development. This should allow more opportunity to intervene to forestall or mitigate its consequences.

We consider some of the key social, legal and ethical questions raised by this endeavour, with a focus on two areas: issues of data protection and data privacy that are raised by accessing and analysing patient records; as well as the nature and consequences of the search for brain-based ‘signatures’ of psychiatric and neurodegenerative diseases, and their use in personalised medicine. Our Foresight Lab will explore time horizons of up to 25 years during and beyond the HBP’s Operational Phase; however, in the current Ramp-Up Phase, we have focussed on issues that may arise during the 10-year Project itself.

1.6 Aims and Methods of this Report

We have explored some of the key social and ethical issues that are emerging, or have the potential to emerge, from the HBP's strategic objectives concerning 'Future Medicine.' For this first report, we focus on a relatively short timescale. We consider the issues that could emerge over the life of the Project, and that have fairly short-term implications for strategic decisions concerning the management of that part of the HBP's work.

This primarily concerns the work of the HBP Medical Informatics Platform. In particular, it concerns the MIP's aim to 'federate' clinical data on patients with psychiatric and neurological diseases from hospitals and research sites in Europe and elsewhere, and to 'mine' these data in the hope of identifying 'brain signatures' of disorders which could inform 'personalised medicine.' These twin aspects - data federation and brain signatures of disorder or disease - are the two foci of this report.

We have used a mixed method loosely built around scenario or vignette construction. The scenarios were developed through detailed engagement with the technical and social literature and discussions with experts. This process was further supported by two webinars with external stakeholders and HBP researchers, organised jointly with our colleagues at the Danish Board of Technology (DBT), to discuss key future medicine issues arising from data federation, data mining, the search for brain signatures, and the development of personalised medicine. Based on these diverse interactions, we developed a set of brief narratives concerning different plausible aspects of HBP futures. These were used to spark conversation in another seminar that brought together researchers from the MIP and various stakeholders and experts.

Our scenarios were used in a seminar organised by our colleagues at the DBT on data federation and personalised disease signatures.¹² The collaborative process was designed to collect data for this future medicine report, to facilitate interaction and reflexive debate between different levels of the research community, and to identify key issues for further discussion between researcher and stakeholder communities.

We identified number of key issues based on this work that are relevant to the strategic development of this part of the HBP. We formulated a number of recommendations, which were then fed back to the senior management of the Human Brain Project for deliberation and potential action.

It is not the Foresight Lab's role to act as the social and ethical regulator of the HBP, nor to require or mandate particular actions. This would be inconsistent with the dialogic and capacity-building ethos of responsible research and innovation. In consonance with the principles of RRI, the role of our reports is to enhance capacity among the HBP managers and researchers at all levels, across the HBP's widely dispersed laboratories and research projects, so that they may make more informed, socially and ethically aware and responsible decisions regarding short-, medium- and long-term research planning.



2. Future Medicine: Background and Challenges

2.1 The Work of the Medical Informatics Platform of the HBP

The Human Brain Project (HBP) is, in part, a response to the fragmentation of brain research and the data it produces. To accelerate the pace of brain research, the HBP aims to provide an integrated system of ICT Platforms offering services to neuroscientists, clinical researchers and technology developers. The Medical Informatics Platform (MIP), which is being developed in Subproject 8 (SP8), is one of six ICT Platforms being set up in the Ramp-Up Phase of the HBP. The MIP was conceived, in part, to address the problem that neurological and psychiatric diagnosis, which currently relies largely on symptomatic classification, is beset by problems of reliability and validity, and also by the very considerable co-morbidity between different diseases, as currently classified.¹³

“The way we understand and treat disease is changing rapidly. The list of conditions for which there is no satisfactory treatment is increasing and, even when treatments are available, many patients either do not respond or experience unacceptable side effects... It is clear that we must move away from a one-size-fits-all approach and towards healthcare that is tailored to the needs and characteristics of the individual. Personalised medicine is a new approach to classifying, understanding, treating and preventing disease based on data and information on individual biological and environmental differences. It seeks to integrate data on the entire dynamic biological makeup of each individual as well as the environmental and lifestyle factors that interface with this makeup to generate a complex, individual phenotype. Using this information, models can be generated to identify the most appropriate healthcare choices, from treatment to prevention, in individual citizens.”

Source: European Science Foundation (ESF) (2012) Personalised Medicine for the European Citizen - towards more precise medicine for the diagnosis, treatment and prevention of disease. Strasbourg: ESF.

For many years, researchers in psychiatry and neurology have sought to identify the precise neurobiological mechanisms that many increasingly believe underpin, or ‘subserve’ the symptoms that patients exhibit in their everyday lives and in clinical settings. As far as psychiatric disorders are concerned, a well-funded diverse programme of research carried out by research groups in many countries has tried to discover ‘biomarkers’ - patterns of genetic data, brain imaging or other physiological measures - that would enable them to precisely diagnose the problem from which the specific patient was ailing, and to target therapies for an underlying neurobiological pathology.

It was hoped that identifying such biomarkers would enable a diagnosis to be made at an early point, or even pre-symptomatically, to enable preventive medicine. This approach, often referred to as personalised, stratified or precision medicine, is not confined to psychiatry, but has been a more general aspiration among research funders and policy makers for the last decade.¹⁴ However, in psychiatry, success has been negligible: advanced genetic studies have not been able to identify a single genetic variation that can be linked to a high probability of a specific psychiatric diagnosis. The same is true of research using brain scanning or other methods. It had been hoped that the two major international diagnostic systems in psychiatry - the International Classification of Disease (ICD) and the Diagnostic and Statistical Manual (DSM) - both of which have been in a process of revision - would be able to include some biomarkers to underpin their disease



classifications. However, despite much research and high aspirations, by the time the fifth edition of the DSM was published in 2013, not a single clinically validated biomarker was available for any psychiatric disease.¹⁵ The same has been true of attempts to find clinically reliable biomarkers for the major neurodegenerative diseases, notably the dementias. Despite some promising results that are providing the basis for further research, clinically reliable and robust biomarkers that can be used to identify unequivocally the presence of a disease are lacking, and diagnosis continues to rely upon symptoms and psychological tests, rather than the precise identification of an underlying neuropathology.

It is in this context that many have suggested an alternative approach.¹⁶ They have suggested that a major problem with previous research is that research subjects have initially been classified on the basis of their evident symptoms (i.e., classified by the DSM or ICD as falling into one or other diagnostic category), after which research has attempted to find biomarkers for these symptomatically defined categories. But if the symptomatic categories are misleading - for example, if the category of schizophrenia includes patients who may appear symptomatically similar, but whose conditions arise from a whole variety of different underlying neurobiological causes - results will inevitably be confused. Perhaps, then, similar symptoms arising from diverse genetic or neurobiological pathways are conflated into a single diagnostic category. On the other hand, perhaps conditions that appear quite different symptomatically arise in fact from rather similar underlying brain anomalies. Indeed, a recent paper published by the Cross-disorder Group of Psychiatric Genomics Consortium suggests that from a genetic point of view there is only one psychiatric disorder but multiple manifestations.¹⁷

To overcome such problems, many now suggest that one should abstain from any initial division of patients by symptomatic diagnosis. Instead, one should simply explore the genetic and other neurobiological data, to see if one can identify clusters of anomalies at this level - at the level of the genetics or the neurobiology itself. Then, one can see whether these might map symptoms in a new way at a second stage, subdividing some 'traditional' diagnoses and grouping together others. One might then explore whether therapies targeted at these underlying neurobiological anomalies would be more effective than those that we already possess.

For those who take this view, the possibilities for identifying such disease clusters lie in the analysis of large quantities of data. This is because large quantities of data from many patients are required to give the data analysis sufficient power to sift through all the different variables and find statistically significant correlations that can enable the identification of these clusters. To gather such data afresh would be a daunting and expensive undertaking, although there is already research seeking to do just that. However another approach is possible: to use the vast quantity of existing data - genetic, physiological and brain imaging data - that has been gathered from many patients and that often lies unexplored in their medical records in hospitals and clinics. Public research groups and pharmaceutical companies also gather such data in the course of research. If such data could be brought together and standardised through a common platform, and 'mined' for correlations and clusters, it is hoped that patterns which are impossible to see in a few individual patients will become apparent - 'brain signatures' of psychiatric and neurological diseases and disorders.

This is the context of the MIP's work. The goal of the MIP is to allow researchers to identify the complex biological mechanisms that explain brain diseases. It aims to bring together - or federate - imaging, genetic and other clinical data that reside in individual hospital, research and clinical trials archives and databases. All this is to be done while guaranteeing protection for sensitive patient information. It is hoped that these resources will permit the identification and constant updating of unique biological signatures of brain



diseases. These will be used for diagnosis, more accurate prognosis and to open up new avenues of drug discovery for the development of new medicines.

The MIP will provide the research community with tools for epidemiological exploration, interactive analysis of new models and methods for diagnosis and treatment, and identification of biological signatures of diseases using data mining algorithms. The services of the MIP will be accessible online via the HBP Unified Portal (UP), which will provide a single point of access to all the HBP ICT Platforms. The MIP hopes to provide end-to-end solutions ranging from processed data to advanced analytical tools. Researchers would then be able to investigate questions requiring data correlation, distribution and interaction in the context of disease processes and epidemiological factors. By using the analytical and predictive tools provided, creating new such tools, or developing data mining algorithms for the validation of new biological signatures of disease, researchers will be able to investigate the relationships between biological and demographic variables and clinical phenotypes. Simultaneously, as further data accrue and new hospitals and data generators are recruited, data mining tools will allow exploration of all the data, to detect recurrent patterns with the aim of identifying biological signatures of disease. It is hoped that such biological signatures of disease will generate a 'diseaseome', by forming the basis for a new disease space that neuroscientists and clinicians can explore.¹⁸

The MIP will integrate large volumes of clinical and research data for mining by tapping into public and research databases and hospital data, federated by novel data management and querying techniques. MIP developers are creating a complex technological platform to achieve this, and hope that their federation software and hardware will allow researchers to query and analyse a very large volume of data without moving them from local servers or compromising data privacy.

Prior to the engagement with our Foresight work, the considered view of the leading researchers in the MIP was that the technical design of the federation architecture, as illustrated in Figure 1 below, was sufficient to ensure the complete anonymity of data subjects, and would meet both present and any future reforms of data protection law. Participating hospitals, clinics and other organisations would have the responsibility of maintaining and protecting data as required by their local Institutional Review Boards and institutional regulations, and all local and national laws. They would place data from their clinical records, stripped of directly identifying details (such as name, date of birth, address) into a local 'data store mirror' which would have no 'feedback' connection to the original data itself. The level of detail copied in the local mirror would be under the control of the local institutions themselves, and could thus vary from institution to institution. This depersonalisation would constitute the initial step in the anonymisation process. The hospital data managers would control the type and quantity of data they provide to the data store mirror, and would be able to alter these with fine-grained control, thanks to the development of novel configurable security mechanisms. The MIP would provide the hospital (at the hospital's expense) with both the local data store mirror and a local node server, which would be able to interrogate the data mirror in response to a query coming through the Data Federation infrastructure.

The key part of the anonymisation procedure would consist in two further steps in the anonymisation process, as only aggregate queries could be executed on the de-identified data with results going through a double round of aggregation, firstly at the level of the hospital local node server, then at the federation level. Nothing but these aggregated results would come out from the MIP into the outside world. A suitably accredited researcher could then file a query via the web portal; this would be broken into pieces and distributed to the hospitals which data mirror contains data relevant for answering; the local servers collect, aggregate and send the data back to the data federation layer for

consolidation and secondary aggregation, which is returned to the researcher's terminal. There would be no way back from the de-identified and doubly aggregated query result to the individuals whose clinical data participated in the result to the query.

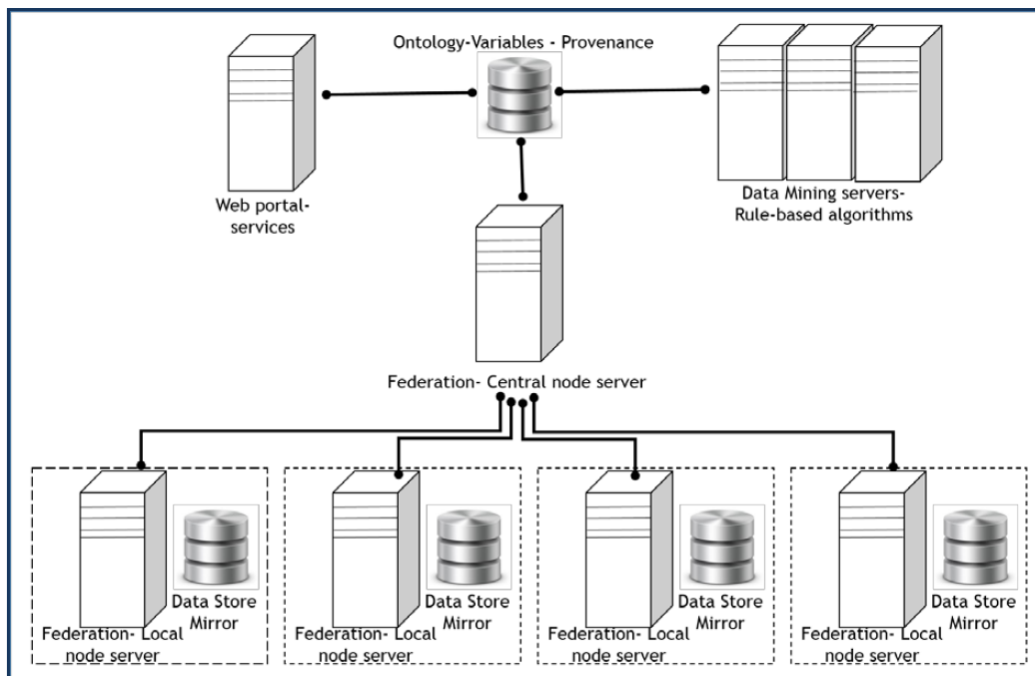


Figure 1: The technical architecture of the Medical Informatics Platform

2.2 Data Federation, Data Protection and Data Privacy

The first major social and ethical issues that we consider in our report concern data protection and data privacy. While much potentially valuable medical information on psychiatric and neurodegenerative disorders remains unexamined and locked away in the clinical records of patients in hospitals across Europe and beyond, accessing, aggregating and interrogating these data presents a number of challenges. In most cases, these data are historical - that is to say they were collected in the past in the course of medical treatment or, in some cases, in the course of clinical trials for matters unrelated to the work of the HBP.

Have patients given their consent to the use of these data for this kind of research? If not, what steps need to be followed to ensure consent, or are there any circumstances where the data can be used without requiring such consent, for instance if it is completely anonymised and aggregated? Can absolute privacy be assured, such that the obligations enshrined in data protection legislation in all the relevant jurisdictions can be met? Is there a conflict between the ethical principle that might insist that these data be used for the public interest, and the ethical principle that might insist that these data remain the property of the individual, and can only be used if that individual gives specific informed consent?

Concerns about data protection have become highly salient in Europe (and elsewhere) recently, and these have major implications for data federation in the HBP. There are two linked issues here - legality (i.e., compliance with legal norms and provisions at EU level and at country level) and trustworthiness. As far as legality is concerned, in 2012, the European Commission proposed a comprehensive reform of the EU's 1995 data protection rules, partly because the legislation had been implemented differently in different

member states leading to fragmentation and additional bureaucracy, and partly because technological progress had changed the way that data was collected and accessed. The initial draft legislation required specific and explicit consent for the use and storage of personal data, but allowed various exemptions for medical and health-related research, enabling it to be processed for medical and epidemiological research, without specific consent from each individual, provided that the data was ‘pseudo-anonymised’, that is to say, the individual’s identity was masked to protect privacy, and provided that the research was subject to strong ethical and governance safeguards, for example as approved by a competent research ethics committee. However, in the wake of the revelations about the mining of electronic data by the US National Security Agency, the draft was amended by LIBE, the Civil Liberties, Justice and Home Affairs Committee of the Parliament, and the new draft legislation, currently being debated at the time of writing, prohibits the use of such personal medical data without specific consent by each ‘data subject’ for each particular use of the data.¹⁹ This amended draft is strongly contested by a large number of medical and scientific research organisations across Europe, on the grounds that it would seriously damage medical research.

In the case of the HBP, such a modification of the legislation could make it exceptionally difficult to federate and mine data as proposed by the MIP, where broad consent for the research use of their data has not been obtained from the patients concerned, or from their families or guardians, if they were deceased or otherwise unable to provide such consent. While the MIP procedures described earlier aim to provide complete anonymity, and it is formally the case that fully anonymised data falls outside the remit of the Data Protection regulations, as we shall see, some argue that full and complete anonymisation is technically impossible, and hence may question the claims made for the anonymisation procedure adopted by the MIP.²⁰ The question of the legality of the MIPs data federation technology, in the light of current and forthcoming legislation and regulation, provided one key theme in our foresight work on Future Medicine for this report.

It should be pointed out that legality does not in itself ensure trustworthiness; public trust in research may require more than simply compliance with existing legislation. This recently became very evident in the controversy about the care.data programme in the UK (<http://www.england.nhs.uk/ourwork/tsd/care-data/>). This project differed from the MIP, because it proposed aggregation outside the firewall of participating data providers, whereas the MIP proposes that this happens inside the firewall. Nevertheless, it clearly illustrated that trustworthiness cannot be assured solely by technical means. The care.data proposal to allow personal data from general practitioners and hospitals to be aggregated in electronic form and mined for the purposes of medical research was mired in a storm of controversy. Although entirely legal under current UK legislation, there was a distinct lack of adequate consultation with the population whose data was to be shared for the purposes of research. There were also inadequate procedures for patients to ‘opt out’ of the programme; in fact, the default option was that patients were automatically ‘opted in’ for data sharing. A further concern among the many critics of the proposal was that, in certain unspecified circumstances, patient data would be made available to commercial entities. Even in a cultural environment of strong trust in the UK National Health Service, care.data failed to gain the trust of those whose data it would use, and implementation was postponed until new procedures could be put into place to secure that trust.

2.3 The Problems of Diagnosis and the Search for Disease Signatures

We consider the issue of biological signatures of disease. As we have said, previous research on ‘biomarkers’ of psychiatric or neurobiological disorders of current diagnostic categories - i.e., objectively measureable biological characteristics that can indicate the



presence and nature of a disease in a living individual - has proved remarkably unsuccessful. At present, despite many hopes and premature claims, there are no clinically validated biomarkers of any neurological or psychiatric disease that can be used in living patients to guide diagnosis, treatment or early intervention.

Recently, a new approach has been advocated that avoids initial classification by current diagnostic categories, such as those used in the DSM or ICD. This approach involves gathering large quantities of biological data on patients (e.g., from brain scans and various tests) and then analysing these data to find clusters of these biological measurements that might help to re-classify patients, not on the basis of observable symptoms, but of underlying neuropathology. Clusters derived in this way will, in a second stage, be related back to symptomatology in order to assess whether an approach based on biology alone will provide better clinical specificity and sensitivity in determining effective treatment.

This approach could enable researchers to investigate the relationship between biological variables and clinical phenotypes. The goal is to produce biologically based diagnostic categories and procedures to enable clinicians to identify the precise pathology that underpins the illness of each individual patient, to personalise treatment, and to eventually develop screening tests to identify and intervene on patients at risk before disease develops.

This approach also faces social and ethical challenges. Most biomarkers are probabilistic rather than deterministic, and this raises dilemmas for treatment. For example, a probabilistic biomarker might indicate that an individual was in a population group that was unlikely to respond well to a particular and expensive treatment, and hence that treatment should not be made available to them.²¹ Screening for biomarkers, with intervention on the basis of biomarkers, raises questions concerning potential false positives and false negatives, and the consequences for self-identity, stigma and social expectations.²² In the case of neurodegenerative disease, issues also arise if pre-symptomatic biomarkers identify a risk of a condition for which no effective treatment is available. Hence, the second focus of our research for this first report on Future Medicine examines the social and ethical dilemmas that are raised by the search for brain signatures in psychiatric and neurological diseases.



3. Data Federation and Privacy: Foresight

3.1 Background Scenarios

The first set of major social and ethical issues considered in our report relate to data protection and data privacy in the federated MIP architecture. Together, they present three main types of challenges to the research team developing the MIP, which we think are key to the future of the Project: legality, trustworthiness and privacy in relation to Big Data expectations.

To explore these issues, the Foresight Lab developed three narrative and fictional short scenarios (vignettes) and sets of questions for a stakeholder seminar organised in collaboration with the Danish Board of Technology Foundation (DBT) in Copenhagen in October 2014. It is important to stress that the scenarios are not predictions of what will or might happen, or anticipations of likely problems or eventualities. As we discussed earlier, the scenarios were intended to illustrate and dramatise potential futures, incorporating external events and contingencies that could affect research and development, to provoke debate and discussion. The scenarios focussed on problematic or negative possibilities in order to provoke debate, not to suggest that these would in fact occur. Indeed, in these discussions, it became clear that some of the futures pictured in the scenarios were unlikely; that other contingencies not embodied in the scenarios were actually more probable; and that for some eventualities, measures of mitigation were already in place.

The analysis, conclusions and recommendations in this report have been informed by these lively debates, and the constructive discussions that occurred in this seminar and in other interactions that were sparked off by this event. As intended by our Foresight work, these discussions have already fed into clarifications and some modifications of the procedures and processes adopted by the MIP. The webinars and workshops for this report were concluded by November 2014, and we present the current protocols of the MIP as of January 2015 in the Annex.

3.1.1 *Scenario 1: The Price of Success*

By 2024, the HBP Medical Informatics Platform has succeeded beyond expectations. The data federation protocol has been widely accepted by hospitals and large amounts of clinical data have been federated, without the need to obtain individual consent from ‘data subjects.’ Over the past few years, despite the significant human and material investment required, research hospitals as well as general hospitals across Europe have been flocking to enrol as collaborating institutions, to give their researchers and clinicians access to the wealth of resources offered by the Platform. However, wide adoption brings some sloppiness of procedures.

This scenario considered the possibility that a widely implemented, successful MIP could lead to participating institutions and their ethical boards becoming hasty in their eagerness to deploy a local MIP server in their IT infrastructure. The scenario develops a possible consequence of such hastiness: that personal data could enter the public domain because of inadequate implementation of the de-identification procedures in the hospital concerned.²³ What other implications could it have?²⁴ What could be the consequences for the MIP? In particular, could the MIP be held co-responsible with the incriminated institutions for the breach of anonymity? What steps could be taken to minimise the risk of such occurrences?²⁵ Is this a technical problem, or (as some suggest) is it that complete anonymity is impossible to guarantee?



3.1.2 Scenario 2: The Costs of Early Adoption

By 2024, the HBP Medical Informatics Platform has achieved only partial success in getting European hospitals to collaborate. Only a few research hospitals are participating, and the MIP has entered into partnerships with other, non-European brain projects as sources of research data.

This scenario was concerned with the possible implications that the failure to enrol a sufficient number of European public hospitals could have for the MIP and participating institutions. It develops a set of ethical and economic consequences that could impact hospitals that have joined the MIP. What changes in strategic choices could such a partial failure mean for the MIP roadmap? What could be the consequences for other actors of the wider MIP, including its sub-contractors? What steps could be taken to ensure that the 'early joiners' do not get penalised by possible changes of strategy? And what might be the implications of the MIP, and the HBP, using data that comes from countries and regions whose ethical protocols are widely considered to be less robust?

3.1.3 Scenario 3: The Citizen-Patient Alternative

By 2024, the HBP Medical Informatics Platform has failed in its objective of enrolling hospitals to collaborate. Despite sophisticated anonymisation protocols, getting ethical approval for accessing existing patient data without their expressly informed consent has proven to be a fatal hurdle. Another major cause for the difficulties of the MIP's partnering with hospitals was that the overall investment proved very costly, while the returns in terms of clinical benefits were dubious.²⁶ In the meantime, another consortium of neuroscientists has chosen a different strategy that is proving successful, but this strategy also has its downsides.

This scenario considered the possibility that the MIP could fail at implementing its planned federated infrastructure of public hospital, in part because of the problems of consent.²⁷ It explores the idea of an alternate distributed infrastructure, implemented with the collaboration of patients and patients support groups. How could such an alternative model develop? How could it interface with the general public health services and in particular the public hospitals? What would be its economics? Would it represent a preferable model to be adopted by the HBP?

3.2 Emerging Challenges

The challenges of legality, trustworthiness and privacy in relation to expectations concerning 'Big Data' were discussed at length in our webinars, seminars and workshops. They have also been much explored in the published literature, including in other reports from social and ethical enquiries. In this section, we distil the main challenges from these various sources, and then, in the next section, we make some recommendations for best practice that have emerged from our discussions and debates. Legality, trustworthiness and privacy are very much interwoven in the ethical and social issues raised by data federation, protection and privacy in the multi-layered architecture of the HBP Medical Informatics Platform. But distinguishing them, as we do here, brings to light different perspectives on these issues.

3.2.1 Legality

To have a chance at succeeding, the Project must conform to the law. Yet in the case of the MIP - a data federation platform with a multi-layered architecture, which may be deployed across diverse organisations in many different countries, each with their own



legal requirements - we are faced with a level of complexity that poses its own legal challenges.

In the MIP's federated architecture, participating local hospitals and other collaborating institutions would each house a local MIP server, which would only store a pre-processed, anonymised and partial mirror of the records database held by the institution, protected by a firewall. These data would never leave the local MIP server, which would be responsible for pre-processing and aggregating the data in response to queries coming through the MIP central federation server. Users of the MIP would interact with the central federation server via the web services offered by the HBP Unified Portal.

This federated architecture raises two main legal challenges. The first is that of the protection of privacy under both EU Data Protection and Human Rights laws. The technical solution chosen by the MIP team consists of combining techniques of data de-identification and double aggregation in such a manner that the MIP would guarantee anonymity, while retaining enough meaningful information to present an interest for research through statistical power and data mining. This would allow the MIP to use historical data from hospitals' patient records and, possibly, past clinical trials by pharmaceutical companies, without having to re-obtain specific consent from each individual. A widespread view of many information security experts and researchers at the Copenhagen seminar was that, from a technical perspective, complete anonymisation cannot be ensured. According to Opinion 05/2014 on Anonymisation Techniques of Article 29 Data Protection Working Party, adopted 10 April 2014,²⁸ no existing technique can ensure full anonymity. The Opinion has analysed the robustness of each technique based on three criteria:

- 1) Individual identification: Is it still possible to single out an individual?
- 2) Linkability: Is it still possible to link different types of records relating to an individual or a group of individuals?
- 3) Inference: Can information be inferred concerning an individual or a group of individuals?

This last criterion opens the door to the possibility of targeted inference attacks, where dates (such as hospital admission/discharge, consultation) are particularly weak links.

Experts in data security at the Copenhagen seminar thought that despite the extensive precautionary measures taken by the MIP team in designing the data anonymisation and aggregation process, malicious misuse of the Platform services could not be entirely ruled out. Many felt that the scenario presented in the first of our vignettes, which highlighted the possibility that with success could come sloppiness, and that de-anonymisation might lead to class action against hospitals participating in the MIP, was quite realistic.

The k-anonymity technique retained for the anonymisation module of the MIP can prevent individual identification, but linkability and inference remain recognised risks. The motivation of a hacker to conduct inference attacks on a large medical database could be for instance to find the medical records of an important person, or the identities of a group of people affected by a specific condition, in order to blackmail them. Moreover, with our increasingly networked and complex information infrastructures, accidental blunders and security breaches are on the rise and cannot be discounted.



‘The Information Commissioner’s Office (ICO) has served [a £180,000 penalty on the Ministry of Justice](#) over serious failings in the way prisons in England and Wales have been handling people’s information.

The penalty follows the loss of a back-up hard drive at HMP Erlestoke prison in Wiltshire in May 2013. The hard drive contained sensitive and confidential information about 2,935 prisoners, including details of links to organised crime, health information, history of drug misuse and material about victims and visitors. The device was not encrypted.

The incident followed a similar case in October 2011, when the ICO was alerted to the loss of another unencrypted hard drive containing the details of 16,000 prisoners serving time at HMP High Down prison in Surrey.

In response to the first incident, in May 2012 the prison service provided new hard drives to all of the 75 prisons across England and Wales still using back-up hard drives in this way. These devices were able to encrypt the information stored on them. But the ICO’s investigation into the latest incident found that the prison service didn’t realise that the encryption option on the new hard drives needed to be turned on to work correctly.’

Source: UK Information Commissioner’s Office, 26 August 2014 <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2014/08/repeated-security-failings-lead-to-180-000-fine-for-ministry-of-justice/>

However, there are additional factors of risk for the protection of privacy in the MIP federated infrastructure. Following ‘Europe 2020 Flagship Initiative Innovation Union,’ EC funded research projects are strongly encouraged to contract out to the European private sector parts of the work for time and cost efficiency, for avoiding the waste of duplicated efforts and for boosting business growth. In this context, the MIP is subcontracting two areas of its work: privacy protocols²⁹ and data federation³⁰. According to the security experts who gave their opinions in our webinars and seminars, the subcontracting of the software for data anonymisation and data federation is introducing an additional level of uncertainty and associated risks.

The second major legal challenge concerns the responsibility for controlling data. The legal experts we consulted were of the opinion that it would be wrong to assume that, since data never physically leave the premises of the local institutions (hospital, etc.) where they are managed, unless they are anonymised through de-identification and double-aggregation, the entire legal responsibility for controlling data rests with these local institutions. The view of the MIP is that, since the MIP never deals with data, but only with the results of queries and that it holds only depersonalised, standardised, normalised data transformed to be addressable by queries, it cannot be deemed a ‘data controller.’ However, in relation to current EU data protection law, the MIP projected architecture could raise the problem of ‘data processing by assignment,’ which is the European law term for the co-controlling of the data.

According to European data law experts, in view of the MIP architecture, the legal entities that can be held responsible under EU law for data and human rights protection are the ‘points of contacts’ between the MIP and the external world, especially as the MIP is not itself a legal entity. This means that the legal entity owning the point of entry into the MIP may be considered to be processing data by assignment, and thus sharing the data controlling responsibility with local institutions. The point of entry into the MIP is the HBP Unified Portal, developed and managed by EPFL, which provides access to all HBP Platforms. It is possible that in the future, once the MIP is set up and enters its operational



phase, the MIP may ‘spin off’ a private venture under the umbrella of an HBP Foundation. This private company would be in charge of selling and maintaining turnkey packages (the local servers’ hardware and software plus training, support and consultancy) to the hospitals and other institutions like pharmaceutical companies signing-up to participate in the MIP. The legal status of the various entities involved in the management of the HBP would then have to be thoroughly re-examined, as well as any significant change in the management and organisation of the MIP federated infrastructure. To address these challenges, we make a number of recommendations below for the information security architecture of the HBP. Many of these are already in development and/or being implemented, as a result of the discussions that form the basis of this report.

3.2.2 Trustworthiness

A focus on the potential legal challenges can lead to privileging technological fixes to the problem of privacy, which do not address the critical challenge of trustworthiness. Conforming to the law is necessary, but it is not sufficient. Projects developed in all legality can flounder because they fail to generate trust among the public and their users. In the domain of public health, a prominent and recent example of such failing has been the case of the NHS care.data programme (<http://www.england.nhs.uk/ourwork/tsd/care-data/>). We now consider how the issues raised by the HBP MIP can challenge trustworthiness, according to the participants who were involved in the stakeholder and expert seminar in Copenhagen.

There was a consensus among participants that trustworthiness is key to obtaining informed consent from individual participants. Ensuring trust is therefore paramount for the MIP to succeed. Identifying and collecting suitable data required for developing data mining algorithms has already proved more difficult than thought initially, due to the impact of privacy protection concerns on the strategy initially adopted by the MIP (anonymisation to circumvent the need to re-obtain specific consent). Moreover, if a process for re-obtaining consent is put in place, the risk is that given the opportunity, people could start to opt-out in large numbers. The resulting loss of research power could then be crippling for the MIP, and it is impossible to force people not to opt-out. While it may be far-fetched, one could imagine a scenario where, in the kinds of national health services that characterise European nations, opting in might be a condition for obtaining the health benefits of an insurance-based system. The rationale for this would be that in such population-wide insurance-based systems, everyone contributes, and risk and cost are distributed. Most of every Euro you pay for medical care in a socialised system is for you, but a tiny bit is also for everyone because it is used for public health.

This argument, which is felt strongly by those who are directing the MIP, has an important base in moral philosophy, but the proposition that one should be obliged to share one’s private data as a condition for receipt of health care is a troubling one for many.

‘Someone can predict whether you are pregnant or not by your shopping activities - it is done already.’

‘Would you object to the fact that with smart metering, your electricity company may know when you are getting a divorce before you know it yourself?’

- What do you think this big data feeding frenzy is about?
- It is about prediction.
- It is about optimising corporations’ bottom line by essentially no morally directed or ethically scrutinised fact-discovering strategy, at



all. It is looking for correlations that can then be exploited.

- It is prediction and as we all know prediction is an inexact science ...
- The trouble is, it doesn't have to be exact to be used, but even if it was exact, that doesn't solve all the ethical problems.'

Source: HBP-SP12 Stakeholder and Expert seminar, Copenhagen, 8-9 October 2014

The third vignette addressed the possibility of involving patients associations and support groups in collaborating with the MIP, turning the notion of informed consent from an obstacle to be overcome into a platform for active engagement. Although fear of discrimination and stigma has real foundations, evidence suggests that many patient groups and engaged patients would be keen to be involved in such a project. The pioneering example was that of AIDS activists when they fought to involve themselves into the medical governance of AIDS. AIDS patient groups are now controlling how experimental drugs are being used, which few would have dreamt of fifteen years ago.

In an area closer to the MIP, at a recent international Huntington disease conference in Barcelona, people with the mutation have declared their genetic status and become among the biggest advocates for further research in Huntington disease; they have encouraged others carrying the genetic sequences predictive of HD to do the same and get involved.³¹ However, there are significant differences between countries in terms of discriminatory attitudes and fear of stigmatisation. When the Alzheimer Europe organises conferences, in certain countries they find few individuals willing to stand up in public and say, "I have dementia." This could have an impact on the representativeness of the groups collaborating with the MIP, with the possible self-exclusion of entire segments of the European population. In this context, and because the MIP research team is deeply convinced that securing the collaboration and involvement of patients and support groups is the way forward, we think that public engagement work with such bodies as Alzheimer Europe is necessary.

3.2.3 Privacy and Big Data Expectations

Legality and trustworthiness are common challenges to all projects involving human subjects' data and in particular health records. The HBP MIP faces yet another kind of challenge, which is that of the expectations, and associated worries about privacy, that Big Data projects can arouse among various constituencies.

The question of privacy was at the centre of the issues debated during the stakeholder seminar in Copenhagen, with diverse positions on the matter. One view was that there is currently a gap in the attitudes of people *vis à vis* how they relate to knowledge in the context of health and medical databases, and how they relate to knowledge collected in social media and the budding 'internet of things' (smartphones, smart watches, fitness apps, etc.). The number of mixed networked structures is increasing rapidly, and it is very hard to predict how these structures will develop and how attitudes will evolve with regard to privacy.

Overall, participants in the seminar thought that attitudes towards privacy and individual health data are becoming more relaxed. However, a recent *Science* review summarising and drawing connections between diverse streams of empirical research on privacy behaviour highlights that people are very certain about the consequences of privacy-related behaviours; that their concerns (or lack thereof) are very much context-dependent; and that privacy concerns may be malleable and open to manipulation by commercial and governmental interests.³² From this perspective, minorities with



psychological or psychiatric conditions may justifiably fear the possibility of discrimination and remain very keen on protecting their personal health records.³³ It was also pointed out that in the area of human brain science in particular, patients could be rationally apprehensive about whether brain scans might, in the future, provide a window into their ‘soul’ - their psychological make-up.

Trust in those who would have access to or would use the data was considered to be a major factor that would influence people’s willingness to share what they would regard as ‘their personal data’ through the MIP. This is despite the fact that, in the firm belief of those who have established the MIP, this concern is misplaced because the de-identification and aggregation procedures ensure that no personal data are shared. The discussions in our webinars and seminars clearly showed that currently, this assertion is either not understood or not accepted, even by relatively expert stakeholders, and hence might be even less well accepted by patients and others contributing data to the MIP.

Trustworthiness is all the more important since current data protection legislation, with its focus on personal information, is ill-suited to providing adequate privacy protection when it comes to Big Data applications. There, the risk may not be so much about identifying specific individuals than about associating identifiable groups of individuals with specific characteristics. This particular issue has been flagged in the report on “The collection, linking and use of data in biomedical research and health care: ethical issues” by the Nuffield Council on Bioethics, released in February 2015.³⁴ The Nuffield report also examines the important question of balancing individual privacy against public good. For individuals to be willing to relinquish some of their privacy by sharing their medical data, they have to trust that doing so will actually benefit the public; hence the importance of engaging publics in the work of the MIP at an early stage.

In fact, technical difficulties may render some of these issues less significant, because it may be the case that much of the data stored in hospital and clinic records is difficult or impossible to process appropriately for use by the MIP. At this stage in the Project, the MIP has been experiencing technical difficulties in processing data sets from hospital providers, due largely to issues with the quality of the data held in hospitals (difficulty mapping local hospital records into the projected standardised MIP data management structure; heterogeneity and scale of the data sources; for brain images, major differences between research scans and the clinical scans performed in hospitals).³⁵ Although such difficulties in standardisation have been anticipated, and may be circumvented by careful selection of variables to be extracted from clinical records, the strategy initially envisioned of relying extensively on historical data held in hospitals may not work out as well as initially thought; this particular reservoir of data, despite its anticipated potential use for the public good, proves difficult to deploy for the proposed research. Indeed the MIP is already exploring the possibility of utilising other data sources, such as those gathered in other research projects or in clinical trials conducted by pharmaceutical companies; indeed such data may present fewer difficulties for use, given the kinds of consent normally obtained in research.

In view of the risks of malicious misuse that we have highlighted, we recommend the systematic exploration of technical scenarios for misuse, and that the testing and auditing of data protection arrangements. Such tests need to be conducted at regular intervals in the future, as the risks posed to privacy by the networking of data will evolve over time. These reviews will form part of the concerns of the Privacy Impact Assessment recommended below.



3.3 Recommendations

The HBP Consortium is a large, decentralised, and distributed organisation that aims to build six ICT Platforms, and to make them accessible to the wider scientific community via the Unified Portal. In line with the conclusions reached at the Copenhagen seminar for addressing the major legal issues facing the MIP, we make the following technical recommendations concerning the Information Security Architecture, which must be a key element of the information security governance process for any large organisation, and which is increasingly seen as essential to the governance and management of IT.³⁶

- **Provide for regular and systematic work on scenarios for misuse.** Misuse can occur in many ways. For example, intruders with malicious intent have demonstrated that they have the capacity and the will to enter a variety of networked databases and to overcome cyber security provisions. It is therefore important to regularly update scenarios for misuse in the light of latest evidence, and ensure that the security architecture of the MIP is robust and protected to the greatest extent possible against misuse.
- **Set up an Information Security Architecture Partnering Project to the MIP.** We recommend that an Information Security Architecture Partnering Project to the MIP be set up before the end of the Ramp-Up Phase. Preferably, it should be extended to the whole of the HBP if such provision is not already in place.
- **Conduct a Privacy Impact Assessment of the MIP before the end of the Ramp-Up Phase.** Such an assessment should preferably cover the whole HBP (if such provisions are not already in place) since the point of entry into the MIP is the HBP Unified Portal developed by SP6 that regulates access to all HBP Platforms.
- **Evaluate the consent requirements for different functions of the MIP, giving priority to informed consent wherever possible.** Despite technical procedures for anonymisation, it is ethically desirable and socially prudent to ensure that either broad consent for research use or full informed consent for specific use is obtained from all patients or research subjects whose data are to be federated and used in the MIP
- **Consider special issues that may arise in the use of data acquired from outside the EU.** Where data are to be obtained from sources (hospitals, research organisations, pharmaceutical companies) outside the EU, special attention needs to be paid to ensure that appropriate informed consent has been obtained and is properly recorded and archived. Prior to any such data being utilised, a request for advice should be made to the Research Ethics Committee of the HBP, and data should only be used if it has their approval.
- **Establish protocols for engagement with patient, patient support and stakeholder groups regarding different MIP functions.** Strenuous efforts should be made in all participating countries to involve patient associations and support groups. As a first step, the AIDS community's involvement in medical research should be investigated, to evaluate how translatable it could be to dementia/Alzheimer for the MIP. The Work Package looking at user support and community building (WP8.5) focuses exclusively on three constituencies: pharmaceutical industry, research centres in mental health prospective and longitudinal studies, and clinicians. We recommend that the community building activities should be broadened to involve more diverse publics. As we have suggested above, overcoming possible prejudices and discriminatory attitudes, both general and country-specific, will require coordinated engagement work.³⁷
- **Develop a Public Engagement and Research Dissemination Plan.** Inspiration for public engagement and participation in research may be found, for instance, in the INVOLVE



initiative set up by the NHS in the UK, which has recently released guidance on the use of social media to actively involve people in research. Investigating these good practices and exploring their translatability to the MIP could be the goal of a HBP Partnering Project under FLAG-ERA.³⁸

- **Create a Data Governance Committee for the MIP with a broad membership including representatives of users and the public.** We recommend following best practice in biobanks and related organisations, by creating a Data Governance Committee, in order actively to involve various publics in the governance structures of the MIP.
- **Ensure a Research Audit structure that can identify, authorise and audit all users of the MIP.** We recommend that a structure be developed to identify, authorise and audit all users of the MIP. Users would be given permissions and each research project application could be elaborated in collaboration between the applicant and a 'MIP consultant' who would help to define precisely the type of data required. The evaluation of each application would involve a privacy risk assessment procedure, and would weigh the privacy risk involved against the potential benefits of the research, before approving, rejecting, or requiring amendments to the application. This process would report regularly to the Data Governance Committee recommended above.



4. Disease Signatures: Foresight

4.1 Background Scenarios

The second set of social and ethical issues considered in this report relates to the search for disease signatures and how these, if found, might be brought into clinical use and what their societal impact might be.³⁹ It should be pointed out that the MIP does not envisage the promulgation of such signatures for clinical practice until the operational phase of the HBP, and so formal plans and strategies for this aspect of their work are currently underdeveloped. To explore these issues, the Foresight Lab developed three more short narrative scenarios, which contributed to further stakeholder discussion in the aforementioned seminar organised in collaboration with the Danish Board of Technology (DBT), which took place in Copenhagen on 8 and 9 October 2014.

4.1.1 *Scenario 4: The Psychiatric Conundrum*

The MIP has been highly successful in identifying brain signatures of neurological disorders, especially Alzheimer's disease and some other dementias, and pre-symptomatic tests are coming into routine clinical practice to identify those forms of Mild Cognitive Impairment that are likely to progress to dementia. But a group of researchers working with the MIP now has claimed to be able to identify the disease signature of depressive disorders and are proposing that this be developed into a screening test to be used in early adolescence. Their University has put out a very optimistic press release claiming that 'depression has been proved to be a brain disease; tests developed by our researchers can now predict who will get depressed, and a new generation of drugs is just around the corner.'

This scenario considers some of the routes to success, including those involving the aggregation of data from a range of sources outside European hospitals, and the combination of methods. It also considers the implications of the extension of the identification of brain signatures of disorders from neurological disorders to the much more contentious area of common psychiatric disorders, and the likely pressure for screening and pharmaceutical intervention. To what extent are the controversies suggested in this scenario realistic? If so, what are the best routes to address them? What might be the implications of developing screening when only probabilistic indicators derived from a small sector (white males) of the general population are available? Will the identification of 'brain signatures' end the debate about social versus biological causes of mental illness?

4.1.2 *Scenario 5: The Award Ceremony*

The highly successful Medical Informatics Platform has been funded by the European Union for another 10 years to continue refining and validating disease signatures that have come into the public domain and, indeed, routine clinical use through the combined efforts of HBP researchers, pharmaceutical companies, social media companies, organised patient networks, and bioinformatics experts. But success brings its critics and its dilemmas.

This scenario presumes that limited success in establishing identifiable disease signatures drives the need for more data and for more refined information to make the disease signatures less probabilistic, and more precise and useable in clinical work with individuals. It considers some of the ways in which this search for more and better information might proceed and some of the costs. And it addresses some of the key issues



concerning the pathways to translation via alliances with commercial companies which can take the research from the lab into the clinic. It also raises the possibility that disease signatures might be used in the criminal courts in determining culpability (comparable to the disputes in the USA about the use of genetic evidence and brain scans in criminal trials).

4.1.3 *Scenario 6: Beyond the Brain*

It is five years after the Human Brain Project formally ended, having accomplished some minor advances in neuroscience and neuromorphic chip design. However, although various disease signatures have been proposed, validation testing and further research shows examples of healthy persons with the same brain signature as those with clear manifestations of the disease in the clinical setting. What would be the implications of this apparent inability to extrapolate from brain signatures to the experience of disease? More data seems to be the solution, but how can the rich qualitative environmental and experiential data that seems pertinent be made compatible with the advanced techniques of data mining required to identify brain signatures?

This scenario considers the possibility that the disease signature identification task set by the Medical Informatics Platform may prove too complex to resolve within the very short funding cycle of the HBP. However the scenario is optimistic in that it suggests that the HBP has built the Project in such a way that the medical research community and young people still feel inspired to take up the challenge. What kind of data collection standards could be comprehensive yet flexible enough to cover the diverse social and geographic situations from which it will be collected? What are the research investment choices that the HBP can make that will benefit the medical neuroscience community even after the HBP has ended? How might the HBP legacy relate to public health strategies in the regions of the world where individuals widely make use of treatments outside the formal medical system?

4.2 Emerging Challenges

Many have suggested that the inability of research to identify the biological basis of mental illnesses has led to a lack of progress in developing targeted treatments for psychiatric or neurological disorders. With the strength of modern computing, it is hoped that more specific knowledge of the biological basis of disease can be identified from the cross analysis of large amounts of patient data, genetic, neuro-imaging, lifestyle and other forms of data federated together in the HBP Medical Informatics Platform. The suggestion is that this biological knowledge, linked to a greater understanding of the specific individual, could be used to create healthcare specific to the circumstances of the individual. This is the approach known as personalised medicine (elsewhere termed stratified or precision medicine) which some believe would transform treatment possibilities and the identification of treatment targets for patients with 'brain disorders.'

The first target for work using the MIP is neurological disorders, and dementias in particular. Many researchers at the HBP and elsewhere believe that disorders currently classed as neurological share many basic mechanisms with those currently classed as psychiatric. Therefore, this approach could illuminate common mechanisms and potentially lead to a radical reformulation of our classificatory systems. Yet, as we have suggested, a number of ethical and social dilemmas are raised by such approaches.



4.2.1 Recognition and Interpretation of Statistical Clusters within Data

One approach used by the MIP to identify disease signatures analyses data in three stages: categorisation, clustering, and classification. Using this “3-C strategy”, an initial proof of principle study by researchers within the MIP sought to develop early diagnostics from groups of biomarkers.⁴⁰ As an initial proof of concept study, available aggregate data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) was categorised into three types: the patient's assigned disease diagnosis, clinical measurements and potential biological markers (for example, genomic or brain imaging data).⁴¹ In the next step, a learning algorithm was used to find hidden structure in the data; that is to say, to identify ‘clusters.’ The number of clusters was decided using a combination of existing medical knowledge and looking at the size of gaps between potential clusters. In the case study, there were 10 clusters (10 new “clinical manifestation classes”) that were to be linked to the clinical measures. In the final classification stage, potential groups of biomarkers were then linked to these 10 new classes. The 10 sets of grouped biomarker values ranges (each set linked to a clinical manifestation class) were then set out using hierarchical decision trees, so that they could easily be described in a clinically relevant way to assist diagnosis.

A first problem with algorithm-based approaches emerged in our discussions with our informants: How can we be certain of the results? The clusters, which are derived by a form of ‘machine learning,’ are unlikely to be self-explanatory to clinicians, patients and other stakeholders, who may well need to be convinced of the precise reasons that the computer has delivered particular clusters. In the seminar discussion, participants raised concerns that decision-making was here being embedded in technical processes. Would there be a human in the room when a decision was being made? Which human? Our informants believed that human beings must make clinical decisions, rather than delegating them to algorithms.

When clustering results indicate new ways to predict disease, clinical research will then be required to test these results. In some cases, the association of a person with a statistical cluster will suggest that they may at some unspecified time in the future develop a problem or that their relatively minor symptoms are precursors for a fully-fledged mental illness. The verification process for proposed disease signatures in some circumstances may take years or simply continue to be ambiguous. More fundamentally, for those seeking a plausible biological account of the aetiology of each disorder, the causal reasons for associations will not be apparent in the results of statistical clustering. Further, there is no reason to believe, *a priori*, that statistical clustering maps onto sets of biological pathways. The existence of statistical clusters may provoke space for new theorising and investigative research into mechanisms. For some mechanisms, where a particular biomarker can result in any of several different disease presentations,⁴² it may be particularly difficult to identify causal pathways from the neurological level to the presentation level given the current state of understanding of the brain. These statistical approaches may be linked to craft-based medical health practices that emphasise the study of the developmental pathways in disease progression, in order to develop practical healthcare responses. Indeed, one webinar participant argued that although Alzheimer's disease patients would be thought of as having a neurodegenerative disorder, psychoanalysis is still very effective in slowing progression and maintaining memory in the first phase of the disease.⁴³

4.2.2 Predictive Medicine

Advocates of this approach are hopeful that it will enable better diagnosis, and hence targeting of treatment. They also hope that it will prove useful for prediction, that is to say, for early identification of a disorder and intervention for the purposes of prevention, or at least mitigation. Thus, for example, testing at the time of first mild symptoms or first



behavioural problems may lead (through the use of biomarkers) to the conclusion that the individual has a higher than average probability of developing a psychiatric or neurological condition, and hence to early intervention.

Predictive medicine raises a number of concerns in that it inverts the normal relationship between doctor and patient.⁴⁴ Traditionally, a patient consults a doctor because they have symptoms and seek treatment. In preventive medicine, however, an individual who is subjectively healthy is advised that they may need to undertake a regime of preventative intervention, usually medication, so that some future illness or predisposition to an illness does not manifest itself. This preventive intervention treats not an illness, but a risk of illness that has often been established by epidemiological studies. Recently, however, with developments in genomics and biomedicine, these risks are established by identifying biomarkers that are associated, probabilistically, with an increased likelihood of developing a disease at some time in the future - identifying a susceptibility or predisposition to a disorder.

Integrated analysis of many data sets in large populations via the MIP could greatly increase such predictive powers. Nonetheless, in the view of many of those involved in our discussions, the predictive powers emerging from large scale statistical work associating biological, clinical and lifestyle data will remain ambiguous, and based on probabilities rather than certainties. Further, in many cases, they are likely to be insufficient definite causal mechanisms to explain the statistical associations. Commercial applications of preventative medicine may also complicate the issue, perhaps suggesting to the proto-patient that their diagnosis, and the treatment offered, is not so much for their benefit as for the commercial benefit of the innovator.

The view emerging from our Foresight activities is that it is necessary to consider the rights of the patient in the patient/health care provider relationship that might develop around these new predictive technologies. An understanding shared by clinicians and patients of issues of risk and probability amongst both clinicians and patients, and relationship of trust between patient and doctor, will be crucially important if advances in mental health care knowledge from such research are to gain public acceptance in clinical practice.

Hence, when considering the ways in which statistical classification from data mining and clustering may give rise to new diagnostic categories and procedures, and novel forms of clinical decision making, our interlocutors suggest that it will be important to link a broad range of communities into this discussion and make the process as understandable as possible. If decisions with potential societal and cultural impact are made only on the basis of medical expertise, this already shifts the doctor/patient relationship in a more technocratic direction. This may also fail to generate the trust necessary for public acceptance.

Seminar participants agreed that it was important to consider how to make the difficult mathematics of statistical clustering intelligible to a broader range of stakeholder publics, so that the latter might be included in a debate about the implications. Evidence from related public dialogues and investigations of public responses to the prospect of 'personalised medicine' suggest that many people - at least in publicly funded health services - have much trust in their medical practitioners and will readily delegate their authority to them. However, there is evidence that some groups are less trusting; younger people and those from ethnic minorities, who are concerned about possibilities of discrimination in the allocation of treatments on the basis of such tests.⁴⁵

In addition, there is a long history of disquiet about relations between psychiatrists and patients, and much evidence that the trust placed in non-psychiatric practitioners is less readily obtained, or easily forfeited, in this particular medical specialty, perhaps for



understandable historical reasons. Further, there are often concerns expressed about commercial involvement and motivation, which can undermine the trust that seems so readily placed in doctors. Hence, there is an important role for an engagement process in establishing institutional trust and ensuring transparency of the process by which these novel approaches have been developed, the methods for their application in clinical settings, and their implications.

4.2.3 What is the Difference Between a Disease Signature and a Brain Signature?

Many of the participants in our discussions raised the question of when a particular pattern of data that correlates to a brain signature should be taken as an indication of a disease. Neurobiological differences do not in themselves indicate disease status. This raises the question of who should decide at what point variance from a norm indicates the presence of a disorder, especially in cases where the issue is one of prediction and preventive intervention. It is clearly very important that such questions are recognised and debated openly, not just with psychiatric and neurological clinicians - many of whom have conflicting views on these issues - but especially with affected or potentially affected patient groups - for example, where parents with a particular psychiatric or neurological condition might be concerned about whether their children are going in turn to develop such a condition.

In this context, linking with affected and potentially affected communities can enable them to play a role in the development of definitions and treatments. In particular, our participants argued that working with patient and disability groups will be of particular value in improving the outcomes of this process. Participation by affected communities may provide insights for effective implementation that researchers or clinicians might not initially see. While it is important to consider concerns with possible stigmatisation of persons with abnormal brain signatures, there are also potential opportunities for identifying neural difference and protecting those persons with it. For example, many would argue that the identification of dyslexia and subsequent procedures to recognise this neural difference and protect those with it from unfair practices or discrimination had positive consequences for those so diagnosed.⁴⁶ Several of our participants argued that we need to learn from history, and to work with potentially affected communities, to ensure that identification of brain signatures in the Human Brain Project in conjunction with affected communities will lead to positive consequences rather than stigma and discrimination. This is clearly a case where outcomes depend on how brain signatures will be communicated to the public. Seminar participants suggested the need for a communication plan that was audience-specific, with different strategies adopted for communication with statisticians, clinicians, patients, the public, journalists, and science policy professionals.

A further issue raised by one of our webinar participants concerned the potential impact of such research on 'cultural ability expectations' - i.e., the ways in which beliefs about the value of certain abilities enable access to resources such as income, political influence and employment. He pointed to historical evidence that changes in expectations about human abilities can be triggered by scientific research, technological development or other events, and might even shape decisions about which types of abilities should be valued in the future. He suggested that it was important to undertake research in this areas using a disability studies lens.⁴⁷ Many different developments are envisioned as contributing to personalised medicine: active health consumers, 'the quantified self', patient-driven healthcare and research models, as well as health social networks and participatory medicine.⁴⁸ We need to consider the potential implications for ability expectations of such developments in predictive and personalised medicine.⁴⁹



4.2.4 Clinical Use and Acceptance of Disease Signatures

Even if it is possible to identify brain signatures of disease that are potentially useful for clinical practice, the process of translation will be lengthy and complex. The incorporation of relatively simple new research findings into clinical best practice guidelines typically lags the publication of research findings by several years and the widespread uptake of new practices usually follows more than 15 years after the initial research was published.⁵⁰ The kind of foundational change in how we diagnose and treat mental health problems suggested by the MIP may take much longer to incorporate. The process of transition to personalised medicine in complex clinical health systems requires careful consideration.⁵¹

Clinical acceptance of new research is a slow process. Early claims of success in a research setting can only be judged in the clinical setting over time. Rapid change in complex institutions such as national health care systems can potentially result in a reduction in the overall quality of health service, especially if change is implemented from above before a broader consensus has been agreed. System-wide changes that are linked to a change in the status of those involved in health provision (e.g., if diagnosis on the basis of brain signatures rather than clinical judgment was linked to deskilling of mental health professionals) may lead to resistance.

As has been argued in work on personalised medicine in other contexts, developments in medical training for clinicians will be required to integrate the use of disease signatures into clinical health. However, the feedback process needs to be bidirectional - the integration of clinical knowledge and practice into the design of health technologies to support clinical use of disease signatures is crucial to their success and acceptance. Even simple things such as the interface design for a computer system should continue to support clinician practice and professional responsibility, or they might contribute to a gradual deskilling of health professionals. There is a danger that this might remove not only the doctor from health decision-making but also the patient, as the latter relies upon subjective interaction with the former to participate in decisions about their own health. Algorithm-based decision aids can support doctors and other health professionals, but it is unlikely that they could replace them, even over longer time scales than are considered in this report.

4.2.5 Personalised Medicine: Implications for Knowledge Infrastructure

There has been a gradual move from one-size-fits-all ‘blockbuster medicine’ (i.e., medicine effective for the whole population of those diagnosed with a disorder) to more stratified medicine (where treatment is specific to different groups of persons with similar biological responses to treatment). While stratification relates to the division of potential patient groups into sub-groups according to their likelihood of responding to particular treatments, personalisation implies tailoring treatment to each specific individual. While personalised medicine was initially thought of specifically in relation to genomic markers, the definition has now broadened: “consideration of individual characteristics, molecular or otherwise, at every stage of medical practice, from prevention, diagnosis, therapy to monitoring.”⁵²

There are currently no clinically validated biomarkers in mental health that could underpin a personalised medicine regime in psychiatry or even in relation to the dementias. However, the search for brain signatures is certainly linked to an aspiration for personalised medicine; hence it is worth considering various images of what a future regime of personalised medicine might look like.

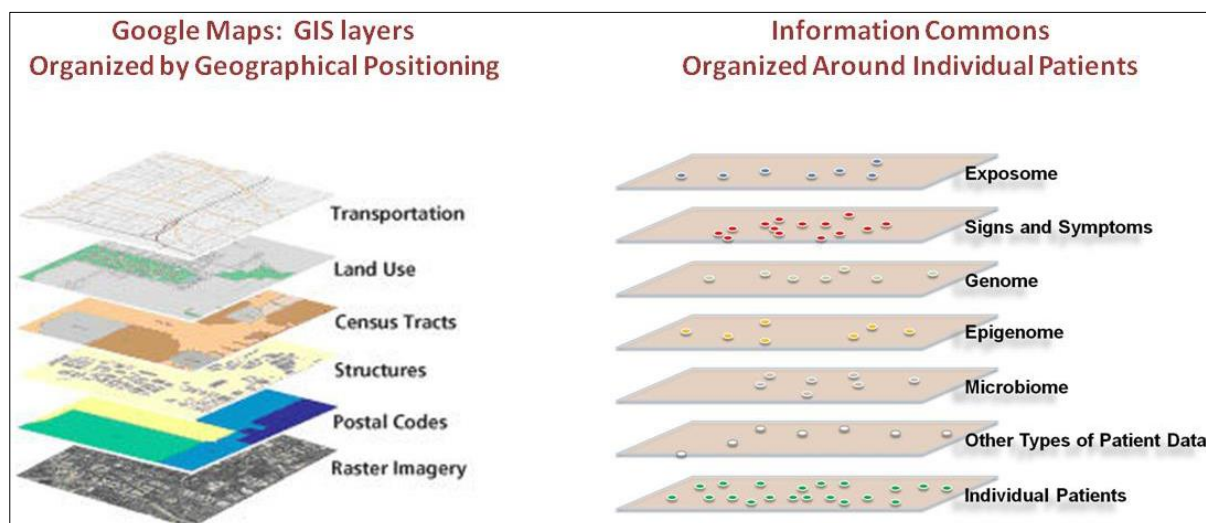


Figure 2: This illustrates several layers of organisation around individual patients in a form similar to layered structures in geographic information services (GIS) mapping. Image from *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*, National Academy of Sciences, 2011 p15.

In current conceptions of personalised medicine, symptom-based disease taxonomies will give way to data-rich biological characterisations of individuals in various stages of their lives. Every disease is unique to the individual, although similar patterns drawn from large amounts of data from other individuals help to explain how the disease manifests itself within the individual in question. In this vision, randomised control trials (RCT) would gradually be replaced with more in-depth studies; first of more stratified groups, but ultimately of one person, coupled with *in silico* modelling. Biological signatures would be developed for the various progressive stages of a disease, and treatments developed and targeted to these biological substrates. But, as was pointed out in our discussions, while some contrast ‘personalised’ medicine to public health preventive medicine that focuses on widescale changes in health environment, there is no reason in principle that social determinants could not be part of the personalisation data mapping structures. As indicated in an earlier chapter, there is already much on-going discussion about the implications for medical research ethics, and it is important to consider what could be the HBP’s contribution to this larger discussion.

Regardless of whether stratified or personalised medicine will become technically possible, there are certainly challenges relating to how this might be implemented within existing health care systems. Participants in our webinars and seminars pointed to many of these changes. Challenges include data integration and interoperability, gatekeeping of data, and protection/privacy. While the previous section of this report discusses many of these issues from the specific perspective of the MIP, we should note that, as big data aspirations become more central to medical and health research, with a wider move towards personalised medicine, these issues will need be revisited more intensively. Our participants also pointed to other necessary changes. There is even now a greater need to develop data literacy amongst a greater number of users. Researchers, doctors and patients must better understand what big data can and cannot do. Many of our interlocutors stressed that, for doctors, algorithm-based decision aids should aim to support rather than replace them.

Furthermore, there is the question of who has claim to the data. Some argue that in a democratic health regime, patients should have access and control of their own data.⁵³ Inclusion of patients in their own data evaluation is not only an issue of patient rights. It



may also improve data analysis. For example, patients might be able to recognise which data might have been left out of a particular data architecture; in particular, the types of things that are hard to digitise.⁵⁴

4.2.6 Patient Involvement

Early ‘upstream’ involvement of patients in the translation of the research of the HBP to clinical medicine process may be able to address potential problems that would arise without their involvement. Managing patient expectations and giving them all the information up front (e.g., where their data could ultimately go) is important. There is a need for good quality information for lay audiences, and working with patient organisations can be helpful in ensuring that patients are fully informed. Many patient organisations, such as the Alzheimer’s Society in the UK, have developed effective ways to engage patients and/or their families in the research process. As indicated by one of our informants, many patients start off being involved in an observational study and transition to being part of an experimental study. However, patient involvement is not merely to be encouraged in the form of active commitment to being research subjects, important as that might be. Involvement of patients as participants in research decision-making or strategy is also important; for example, being on or advising governance committees.

We suggest that, to the greatest extent possible, involvement of patients prior to the clinical translation process itself is desirable. Communication should be a two-way partnership with patients who may be able to help in developing protocols and research policies, and will be knowledgeable about issues of data use, governance, and de-identification. Better communication with patients naturally allows them to transition from being informed research subjects to the next level of assisting with research governance. As the MIP moves towards developing its work on brain signatures towards clinical applications, they might consider patient inclusion in an on-going monitoring or ‘steering’ group. In addition to potential benefits to the research process this gives patients and other stakeholders’ confidence in the integrity of the study. Ultimately, all evidence suggests that patient endorsement is paramount for success.



4.3 Recommendations

Many of the social and ethical challenges that the MIP faces can potentially be alleviated by having a strong consultation and engagement process with communities that may be affected by the work of the HBP. Integrating a broad range of communities and stakeholders into the research process of the HBP MIP (in some cases, with the aim of building long-term working relations) can potentially improve the quality of HBP research methods, outcomes, and their translation into future health technologies and clinical practice. Thus, on the basis of our Foresight work on disease signatures, we make the following recommendations:¹

- 1) Include patients and clinicians in a research advisory capacity. Remember that communication is two-way. Consider patient and clinician inclusion on an advisory group. In addition to assisting in the translation of research results into clinical practice and new health care technology, this body might help in developing research protocols and data usage protocols.
- 2) Address implications for clinical practice; in particular, by engaging clinicians in the assessment and verification of disease signatures and their utility in clinical applications.
- 3) Provide more information on what big data can and cannot accomplish in disease signature medicine. Algorithm-based decision aids can support doctors and other health professionals, but not replace them.
- 4) Address implications of disease signatures for clinical ethics.
- 5) Let patients know up front how their data might be used - covering all the different possible research uses. Consider to what extent patients will have control of and access to their own data. Consider how incidental findings will be handled.
- 6) Reflect on the use of brain signatures in the identification of pre-clinical susceptibilities.
- 7) Facilitate discussion and solicit input from user and stakeholder communities on the development of disease signatures that are clinically effective and socially acceptable. This is especially important where such signatures may be used in a pre-clinical phase, i.e. before a disorder is manifest or subjectively experienced by a patient. The involvement of clinicians and patients in the interpretation of results is especially important here.
- 8) Ensure awareness that neural difference does not equal brain disease.
- 9) Consider how brain signatures will impact upon cultural expectations about people's ability. Will this create more or less security (in self-identity, employment, etc.) for differently abled people? How will brain signature impact upon neural diversity?
- 10) Develop effective communication strategies to explain the potential clinical and research uses of brain signatures to the public at large and other audiences.
- 11) Develop a communication plan for complex science, allowing for different possible interpretations. This could benefit from being audience-specific: statisticians, neuroscientists, clinicians, patients, legal and policy professionals, journalists and the public at large all require different types of information to understand what we know

¹ These are numbered in sequence with the recommendations given in Chapter 3.



about unique brain data patterns and how this knowledge can best be used to improve human health and well-being. It may also be important to communicate what limits our knowledge has and what we can't do with brain signatures.

- 12) Consider the wider implications of moving to a brain-based understanding of disorders.
- 13) Observe how new findings are interpreted by researchers, clinicians, pharmaceutical companies, media, public opinion and the legal system. Some suggest that these developments may raise questions about what it means to be human.⁵⁵ History shows that any such changes are slow, complex and dependent on many other social and cultural conditions. Nonetheless, there is a role for historical, philosophical, legal and anthropological enquiry into the medium- and long-term implications of the advances in our understanding of neural processes that will be brought about by the new brain projects.
- 14) Link with other research communities and relevant regulators to develop appropriate pathways for translation of research to clinical applications.
- 15) Develop a detailed strategy for a translation pathway to clinical use as the work of the MIP moves toward clinical applications. This will enable progress toward personalisation in psychiatric and brain disorders.
- 16) Consider the implications for public trust and support of the HBP, of the future intended commercialisation of the findings to generate employment and wealth creation.
- 17) Consider how early stage research and development can be better directed towards publically desirable outcomes that address major societal challenges, as we have seen how public trust can be easily forfeited if biomedical research is seen as primarily directed towards commercialisation and private profit rather than public good.



5. Acknowledgments

We would like to thank all the individuals and organisations who have contributed generously to the work presented in this report. There are too many to name here, but we would like particularly to thank the following:

- All of our colleagues in SP12, Ethics and Society, in particular the Danish Board of Technology Foundation, Lise Bitsch and Lars Kluwer, WP12.3
- Richard Frackowiak, Ferath Kherif and Bogdan Draganski (Hospices Cantonaux, Centre Hospitalier Universitaire Vaudois) as members of the MIP of the HBP SP8
- The HBP Science Writing and Editorial Services Office
- All participants in the webinars co-organised with the Danish Board of Technology Foundation: Dennis-Kenji Kipker (Universität Bremen), Dirk Lanzerath (European Network of Research Ethics Committees - EUREC), Paul Quinn (Vrije Universiteit Brussel), Caspar Bowden, Emilio Mordini (Responsible technology), Doug Brown (Alzheimer's Society), Gregor Wolbring (University of Calgary); all participants in the HBP seminars
- Colleagues at King's College London: Claire Marris, Barbara Prainsack, Tara Mahfoud, Loretta Platts
- John Ashburner (University College London)
- David Ford, (Professor of Health Informatics and leader of the Health Informatics Group at Swansea University College of Medicine)



6. Annex 1: The Scenarios

6.1 Data federation and privacy

6.1.1 Scenario 1: The Price of Success

By 2024, the HBP Medical Informatics Platform has succeeded beyond expectations. The data federation protocol has been widely accepted by hospitals, and large amounts of clinical data have been federated, without the need to obtain individual consent from ‘data subjects.’ Over the past few years, despite the significant human and material investment required, research hospitals as well as general hospitals across Europe have been flocking to enrol as collaborating institutions to give their researchers and clinicians access to the wealth of resources offered by the Platform. However, wide adoption brings some sloppiness of procedures.

‘Today is 21 September 2024, and it promises to be a momentous day in Mia’s young career as an ambitious lawyer. This is her conviction, anyway, as she slaps the alarm clock off.

Mia has picked World Alzheimer’s Day to file her first ever class action lawsuit. It will send a message that the media should love, and that should make her career. She is going to file a class action lawsuit in the European Court of Justice against one of the largest general hospitals in the country, for violation of the European data protection law. Mia thinks back on the events - unfortunate for some but lucky for her - that have made it possible.

A year ago, there was a short story in the news that caught her attention. The identity and health records of an early onset Alzheimer’s patient had been made public through an alleged blunder at the hospital in which neurology service she was treated. The diagnostic capacities based on brain signatures that had been pioneered by the Human Brain Project were now being widely used in clinics for the early identification of neurodegenerative disorders based on brain signatures, and in this case, the diagnostic brain signature pointed at a specific type of dementia that was highly correlated with an inheritable genetic marker.

Soon after the information about her mother’s condition was in the public domain, her only daughter was made redundant on supposedly economic grounds and got into trouble with her mortgage-protection insurance for not having divulged her mother’s condition. Mia had contacted the daughter as a result of the story, pitching to her the possible benefits of filing a suit against the hospital, and discovered that there was scope for much more than an individual lawsuit. Her potential client’s misfortunes originated not so much from a blunder than from a weakness in the hospital data management architecture, which meant the problem could occur again. The weakness could be traced to the fact that the anonymisation protocols established by the MIP some years before had apparently been inadequately implemented by the hospital database manager, so that unfortunate cases of individual re-identification could occur, exposing entire families through genetic linkage. To add to that, patients had never been made aware of this use of their data or been asked for their consent.

At this point, Mia had felt that she had the makings of a highly successful landmark case, and one that would question the very legal and ethical foundations of this new diagnostic procedure.’



6.1.2 Scenario 2: The Costs of Early Adoption

By 2024, the HBP Medical Informatics Platform has achieved only partial success in getting European hospitals to collaborate. Only a few research hospitals are participating and the MIP has entered into partnerships with other, non-European brain projects as sources of research data.

‘Today is 9 November 2024, and it promises to be an utterly nightmarish day. This, anyway, is Paul’s conviction as he swallows the last dregs of his coffee and nervously makes his way towards the boardroom. It is the first time in his seven years as chief executive director of one of the country’s prominent teaching and research hospitals that he has so dreaded attending a meeting with the Board of Trustees. In all likelihood, before this is over there will be blood on the carpet and it will be his. How could things have turned out so badly?’

In 2015 when Paul had engaged into a partnership with the Human Brain Project, to make his establishment one of the pilots collaborating research hospitals to their Medical Informatics Platform, his daring strategic move had received full approval from the Board of Trustees and had been hailed as ‘visionary’ by expert commentators. But the idea of a European public hospital federated infrastructure is now history. For the few European hospitals, like his, which became involved in the early stages, the investment has turned out to be considerable. This has dissuaded other hospitals from collaborating. As a result, the private contractors who were providing the local servers bundled solutions have announced that they will soon stop maintaining the dedicated software because they have sold too few units.

Therefore, the hospital is going to be left with costly dedicated machines that are difficult to put to other uses. And to add to the pressure on the hospital management to pull out and write the whole thing off as a bad decision, their ethical board is less than happy to see them associate with a venture that has become controversial. This is in part because the MIP has revised its strategy, and is now using data from hospitals outside the EU, which have been less concerned about matters such as data protection and individual consent: the volume of data available in this way has made it less important to collaborate with European hospitals, but at the price of ethical and political controversy.’

This scenario is concerned with the possible implications that the failure to enrol a sufficient number of European public hospitals could have for the MIP and participating institutions. It develops a set of ethical and economic consequences that could impact hospitals that have joined the MIP. What changes in strategic choices could such a partial failure mean for the MIP roadmap? What could be the consequences for other actors of the wider MIP, including its sub-contractors? What steps could be taken to ensure that the ‘early joiners’ do not get penalised by possible changes of strategy? And what might be the implications of the MIP, and the HBP, using data that comes from countries and regions whose ethical protocols are widely considered to be less robust?

6.1.3 Scenario 3: The Citizen-Patient Alternative

By 2024, the HBP Medical Informatics Platform has failed in its objective of enrolling hospitals to collaborate. Despite sophisticated anonymisation protocols, getting ethical approval for accessing existing patient data without their expressly informed consent has proven to be a fatal hurdle. Another major cause for the difficulties of the MIP’s partnering with hospitals was that the overall investment proved very costly, while the returns in terms of clinical benefits were dubious.⁵⁶ In the meantime, another consortium of neuroscientists has chosen a different strategy that is proving successful, but this strategy also has its downsides.



‘Today is Friday 23 April 2024, and the end of a week which started as an ordinary week for Gert but has turned into a life-changing decision-making ordeal. Having turned 50 a few weeks ago, he got treated through a comprehensive check-up courtesy of the National Health Service, and he met his General Practitioner first thing on Monday morning for the routine post-results feedback - except, it ended up being not so routine. The blood exam testing for the likelihood of his developing some form of dementia in the next five to ten years has proved positive. This was a shattering blow to his retirement dreams. But although Gert’s GP has been very professional in his advice, Gert would never think of relying solely on the NHS diagnosis. The NHS blood testing is all well and good, but he wants to have access to more cutting-edge diagnostic and monitoring tools than those available through the general public health services.

Gert has heard of an EU-wide project that was been set up a few years ago, the Human Brain Project, but it appears that although research is still going on, the diagnostic procedures have not become generally available for clinical use, as few hospitals have agreed to participate, and there were some kinds of legal controversies - Gert does not quite understand why. Despite the best efforts of its public engagement work, the Human Brain Project has become bogged down in controversy, much of it coming from civil liberty groups and other radical critics, who have also suggested that its real motives are economic, as it has not made its findings public and has focussed on developing intellectual property and collaboration with private corporations who are charging exorbitant amounts for the use of the tests.

However, Gert has found other options by surfing the websites of patient support groups. These are linked to a new extension to the EU Biobank, which has been developed by other neuroscientists in close collaboration with patient groups across Europe, and where thousands of patients have agreed to all their clinical records and samples being stored and analysed, with full individual consent, and where results have been very positive.

Researchers have worked closely with patients groups and with clinicians to overcome barriers of mistrust and uncertainty about the meaning of the results, to ensure that the full benefits of early diagnostic and preventive therapies are available in the clinic. But because of the costs involved in collecting and processing the data, in particular the brain images that have to be research-grade, there is a selection process to get enrolled.

It seems that patients’ support groups are acting as gatekeepers, and here is the small silver lining: because of his prospects of getting early onset dementia, Gert has a very good chance of getting enrolled. But still, this is an elitist process and only those who happen to have a powerful and well-resourced patient group to represent them and ‘their condition’ have been able to access this technology: the prospects of this diagnostic technology becoming routinely available for less ‘fashionable’ conditions and in every general hospital in the EU are remote.

This scenario considers the possibility that the MIP could fail at implementing its planned federated infrastructure of public hospital, and explores the idea of an alternate distributed infrastructure, implemented with the collaboration of patients and patients support groups. How could such an alternative model develop? How could it interface with the general public health services and in particular the public hospitals? What would be its economics? Would it represent a preferable model to be adopted by the HBP?



6.2 Disease Signatures

6.2.1 Scenario 4: The Psychiatric Conundrum

The Medical Informatics Platform has been highly successful in identifying brain signatures of neurological disorders, especially Alzheimer's disease and some other dementias. Pre-symptomatic tests are coming into routine clinical practice to identify those forms of Mild Cognitive Impairment that are likely to progress to dementia. But a group of researchers working with the MIP now has claimed to be able to identify the disease signature of depressive disorders and are proposing that this be developed into a screening test to be used in early adolescence. Their University has put out a very optimistic press release claiming that 'depression has been proved to be a brain disease, tests developed by our researchers can now predict who will get depressed, and a new generation of drugs is just around the corner.'

'It was Saturday morning in Heidelberg in November 2020, and Philippe was still recovering from the late arrival of his flight to Frankfurt and the long transfer by train - he wondered: why on earth did they choose to have the press conference here? In fact, he was rather dreading the whole thing.

Philippe's group had shifted the MIP into a new phase. It had always been the belief of the directors that they could find brain signatures not just for neurological disorders, but also for psychiatric diseases - the goal that had eluded researchers for so many years. But his group had done it, when at last they had enough data from European hospitals, supplemented by some collected from other sources, and driven in part by hypotheses from existing research identified using the HBP's own platforms - including some hitherto obscure mouse data, and some reports from China - based on work with higher primates unfortunately - together with brain mapping data from a large cohort recruited from the favelas in Rio by a contract research organisation (very economically he was delighted to say). Data integration was paying off, and although the 'hypothesis free' route of data mining had not generated results for psychiatry, they had combined 'top down' and 'bottom up' and it worked.

The results were probabilistic, of course, but using a retrospective approach to their longitudinal data, they had shown that they could predict at age 14 with 70% accuracy - in white males only at present - whether an individual was going to have developed a major depressive disorder by 25 - well actually a subtype of major depression, as the condition had been subdivided as a result of neurobiological work undertaken by the National Institute of Mental Health in the USA.

Now, there was no doubt that depression was a neurobiological condition with a strong genetic component - those who claimed that it was environmentally caused, or a response to misfortune, or a reaction to unemployment were just plain wrong. But he could not understand the response - the organisers of the press conference had warned him that there were to be big demonstrations from protesters, ranging from scientologists to patients groups, and he had received several abusive emails - this seemed to be a very sensitive subject in Germany, for some reason.

Maybe he had been unwise to suggest that this could lead to a programme of population screening of schoolchildren for the disorder and early intervention - but surely prevention was better than cure, and if these kids could be given antidepressants to ward off the depression, what would be wrong with that? Especially since his group had entered into a very lucrative agreement with a large pharmaceutical company to develop a whole new



generation of antidepressants which would be prescribed after the identification of the brain signature - these would be smart drugs, not like those old dirty SSRIs, with all their side effects - they would really target the biological basis of the disorder. Public health officials were already excited - think of the effect of these new treatments on the 'burden of disease' - especially in less developed countries where all those faintly ridiculous talking therapies were not available. Like it or not, these are the facts and they will have to live with them, he said to himself as he pulled on his trademark black tee shirt and jeans and slung his leather jacket over his shoulder.'

This scenario considers some of the routes to success, including those involving the aggregation of data from a range of sources outside European hospitals, and the combination of methods. It also considers the implications of the extension of the identification of brain signatures of disorders from neurological disorders to the much more contentious area of common psychiatric disorders, and the likely pressure for screening and pharmaceutical intervention. To what extent are the controversies suggested in this scenario realistic? If so, what are the best routes to address them? What might be the implications of developing screening when only probabilistic indicators, which are derived from a small sector (white males) of the general population, are available? Will the identification of 'brain signatures' end the debate about social versus biological causes of mental illness?

6.2.2 Scenario 5: The Award Ceremony

The highly successful Medical Informatics Platform has been refunded by the European Union for a second ten years to continue to refine and validate the various disease signatures that have come into the public domain and routine clinical use through the combined efforts of HBP researchers, pharmaceutical companies, social media company, organised patient networks, and bioinformatics experts. But success brings its critics and its dilemmas.

'It was Saturday afternoon, August 8, 2026. Eugénie was getting dressed for the dinner and award ceremony later that evening. She would be receiving a special award for 'Social Media and Health.' It was her idea that had moved the work of the HBP in identifying brain signatures to a new level: the Project had run into trouble because although hospital records had enabled brain signatures to be identified, their precision was limited because of inadequate lifestyle information; hospitals and clinicians were either unable or unwilling to trace back to patients to supplement what data was on the electronic data records.

Eugénie had had the great idea of developing a social media company that brought together Facebook participants to share their data with the Human Brain Project, but this was alongside her real success, which was to help negotiate the contracts that would enable the HBP to gain access to a massive amount of data from outside the EU which would 'power up' the data set and help refine the disease signatures so that they were precise enough for individual clinical use. As data officer, she also had to handle some sensitive media work.

This past week, Eugénie helped the HBP to distance itself from the legal system in Hungary that which had just accepted a brain disease signature as evidence in a criminal prosecution. But the most exciting thing was that she was being offered a position on a company board! This was a giant leap forward in her career. Facebook was collaborating with one of the big pharmaceutical companies in Germany, and as part of this there would be a small research company, Pharmakon Sociale, that would serve as the intermediary. She was recommended to be on the new Board of Directors as a result of her data sharing work. This was a great opportunity for Facebook to monetize their information assets.



In the elevator, Eugénie checked her voicemail. Her secretary had called to tell her that a political columnist had written an article about the legacy of the HBP Medical Informatics Platform in *Le Monde*. While the journalist grudgingly acknowledged the significance of the diagnostic technology that the HBP had helped develop, he focussed more on the conflicts around the commercialisation of the technology and the links of the researchers with private corporations, on unresolved issues of patient's rights, and on the dubious ethical status of the collaboration with China. Eugénie was named personally in the column as someone who had benefited financially from her involvement and hence might have had conflicts of interest. "Well, I suppose," thought Eugénie, "she might as well get used to this type of thing. It was just part of being successful."

This scenario presumes that limited success in establishing identifiable disease signatures drives the need for more data and for more refined information to make the disease signatures less probabilistic and more precise and useable in clinical work with individuals. It considers some of the ways in which this search for more and better information might proceed and some of the costs. And it addresses some of the key issues concerning the pathways to translation via alliances with commercial companies who can take the research from the lab into the clinic. It also raises the possibility that disease signatures might be used in the criminal courts in determining culpability (comparable to the disputes in the USA about the use of genetic evidence and brain scans in criminal trials).

6.2.3 Scenario 6: Beyond the Brain

It is five years after the Human Brain Project formally ended, having accomplished some minor advances in neuroscience and neuromorphic chip design. However, although various disease signatures have been proposed, validation testing and further research shows examples of healthy persons with the same brain signature as those with clear manifestations of the disease in the clinical setting. What would be the implications of this apparent inability to extrapolate from brain signatures to the experience of disease? More data seems to be the solution, but how can the rich qualitative environmental and experiential data that seems pertinent be made compatible with the advanced techniques of data mining required to identify brain signatures?

'Today is Tuesday May 8, 2029. It is Indira's second day in the field and she found this summer job exciting. She would be driving all over Scandinavia and North Western Germany this summer interviewing families about the diet and exercise habits of their deceased loved one.

The Finnish neuroscientist who had hired Indira felt that if only she had additional data she might be a step closer to solving the riddle about why the hypothesised disease signatures that had been identified in the Human Brain Project, and had been the subject of dozens of research papers, only correlated with the onset or severity of neurodegeneration in some cases but not others. She had got the idea of 'verbal autopsies' from public health work in India where many people die in rural communities outside of the medical system. The census often uses information provided by surviving relatives in the household to give the cause of death.

The researcher for whom Indira was working had read about this Indian strategy, and thought that it could be adapted to provide much more information about those who lived with, and died from neurodegenerative diseases. So she was visiting their homes and conducting a survey to collect information about the deceased - about their exercise and diet, amount of social interaction, employment and work stress, habits of reading or watching television; she had even included a question about how often they listened to



music and of what type was the music. And she was also gathering information about the kinds of treatment that people had sought, much of it outside the official medical system - she found that people experiencing disturbing symptoms of loss of memory used everything from herbal remedies to prayer or meditation, and often reported real improvements.

Although she was not a neuroscientist, Indira had trained in psychology for her undergraduate degree, and had developed considerable skills in quantitative statistics. What she could not quite understand was how the stories that she was being told by the surviving relatives about the lives and habits of their loved ones, and the range of non-medical treatments that they had used, could be turned into the kinds of data that might be used in the quantitative analyses required by neuroscientists. It all seemed so woolly, so dependent on beliefs, customs, and the peculiarities of individual lives. Indira wanted to apply to graduate school and continue her education in the sciences, and now thought she would specialise in neuroscience - surely she would be able to learn how the experts managed to turn all these individual experiences, that seemed so real and important if one really wanted to understand the progress of the disease, into data.'

This scenario considers the possibility that the disease signature identification task set by the Medical Informatics Platform may prove too complex to resolve within the very short funding cycle of the HBP. However the scenario is optimistic in that it suggests that the HBP has built the Project in such a way that the medical research community and young people still feel inspired to take up the challenge. What kind of data collection standards could be comprehensive yet flexible to the diverse social and geographic situations from which it will be collected? What are the research investment choices that the HBP can make that will benefit the medical neuroscience community even after the HBP has ended? How might the HBP legacy relate to public health strategies in the regions of the world where individuals widely make use of treatments outside the formal medical system?

7. Annex 2: The Medical Informatics Platform

We reproduce here some further technical details of the MIP as the Platform has been developed in dialogue with the work of the Foresight Lab.

The HBP Medical Informatics Platform (SP8) will federate large volumes of anonymised data (genetic data, imaging data, and other clinical data) originally generated for clinical purposes, and make it available to the research community. Procedures for anonymisation are described below. Partnering Projects will mine these data for biological signatures of disease, which if found, could provide important insights into disease mechanisms, contributing to the development of new diagnostic tools and new treatments. The Project will encourage community efforts to use Platform data and tools for studies of a broad range of brain disorders. Below we discuss the implications of this approach.

Compliance with European and National Data Protection Law. At the time of writing, data protection in EU member states is regulated by EU directive 95/46/EC and by derived national legislation. Negotiations for a new Data Protection Regulation are now at an advanced stage. Given, however, the draft regulation has yet to be finalised, the discussion here will be limited to existing law.

For the purposes of data protection law, health-related data pertaining to a data subject is personal data, and can only be gathered legally under strict conditions, for a legitimate purpose. In particular, Recital 33 of the data protection directive provides that “data should not be processed unless the data subject gives his explicit consent”². However, Recital 26 states, “the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable.”

Architectural considerations. In the architecture adopted by the HBP Medical Informatics Platform (SP8), all data referring to human subjects is held in local data repositories managed by the individual hospitals that contribute data to the Project. There is no central repository and no transfer of raw data outside the hospital perimeter. Thus, the raw data are protected by the same technical infrastructure and technical measures and receives the same legal protection provided to all patient data. This implies that any attempt to re-identify patient data would constitute a criminal offense. Access to raw data is restricted to authorised personnel. Access is protected by passwords, and additional physical protection measures (e.g. use of smart cards), in line with the policies adopted by individual hospitals. Servers are protected using the same measures used to protect other hospital information systems containing patient data.

De-identification (anonymisation) of data. Given the recitals of the data protection directive, the applicability of Data Protection legislation depends on whether or not the data stored in the Medical Informatics Platform can be treated as anonymous data.

De-identification or anonymisation of data is the process whereby personal data are processed with the aim of preventing identification of the data subject. Several anonymisation techniques may be envisaged, (and) there is no prescriptive standard in EU legislation. Relevant standards and regulations include ISO 29100:2011, and the US HIPAA regulation.

² Note the ethical requirement that data subjects should give informed consent for the use of their data depend not just on data protection law but also on general principles of medical ethics, as laid down in the Helsinki Convention, and on explicit requirements of European and national law.



In its Opinion 05/2014 on anonymisation techniques, the EU “Data Protection Working Party” examines the robustness of a broad range of anonymisation techniques. The Working Party concludes that “anonymisation techniques can provide privacy guarantees and may be used to generate efficient anonymisation processes, but only if their application is engineered appropriately (...) The optimal solution should be decided on a case-by-case basis, possibly by using a combination of different techniques, while taking into account the practical recommendations developed in this Opinion.” In the light of these recommendations, SP8 has adopted a strategy of “defence in depth” which combines different technical and organisational measures. In line with the findings of the Working Group, SP8 recognises that anonymity and anonymisation are lively fields of research and will update its protection measures as the field progresses.

Local HBP data repositories contain only anonymised data. Thus, even an HBP user with system administrator rights cannot access individual patient records. The data repositories are populated using a pipeline that extracts data from hospital medical records, and applies filters to remove information that could allow the identification of a patient. The procedure includes removal of identifiers and pseudo-identifiers that could allow the re-identification of patients, as specified by the US HIPAA regulation. It also provides for addition of noise to some kinds of clinical data to prevent identification of patients. In particular, brain-imaging data are “defaced” (i.e., image parameters are altered to prevent reconstruction of a patient’s face).

No end-user access to raw data. The HBP MIP will provide hospitals with the software to post-process the raw data contained in local data repositories and extract features of interest (grey and white-matter volumes, as revealed by medical imaging). The post-processing software is controlled and run by the hospitals. End-users of the Platform will be able to query the feature data but not the raw data. De-identified raw data (e.g. imaging data) will be conserved in local repositories for use by HBP researchers involved in the development of new feature extraction algorithms. Access to the data will be restricted to researchers authorised by individual hospital data controllers.

k-anonymity. The Medical Informatics Platform will use the technique of k-anonymity. The Platform will provide end-users with descriptive statistics for particular features (or correlations among them) in a set of records matching a query, only when the number of records in the set is greater than a threshold-value. If only one or a small number of records, satisfy a query, the Platform will not respond to the query. Work is in progress to extend this approach to genetic data. The software will include filters that block suspicious queries.

Audit trail. With the protective measures in place it would be virtually possible for an attacker to infer data about an individual patient. Nonetheless, the Medical Informatics System will maintain an audit trail, recording the origin, time, date and content of individual queries and the records used to generate the response. Analysis of these data could in principle detect suspicious activity.

Software. Software implemented in the Medical Informatics Platform, whose development has been funded by the HBP, will be released under an open-source license such as BSD. The same code will be available for by privacy impact audits (see below). To extract data in primary hospital information systems, and remove HIPPA identifiers the HBP will use software provided, configured and maintained by GnuBila, a leading supplier of hospital software. The software meets HIPAA standards together with additional HBP requirements and will allow system administrators to precisely define access rights. The software will run on servers managed by hospital staff. The subcontracting company will have no access to patient data.



Pseudonymity. The policy of some hospitals requires that patients should be able to request details of the purposes for which their data have been used. This may make it necessary to maintain a table-linking patient codes and identifies. Such a table would be held by local hospital data controllers and would not be accessible to HBP staff or to users of the HBP platform. Given that researchers using the HBP platform will not have access to data for individual patients, they would not be able to use the code to trace individual patients. Nonetheless, hospitals that adopted such policies would not be able to treat the data as anonymous, and would thus be subject to data protection law. Discussions are in progress with hospital administrations to find a suitable solution before the platform comes on line.

Organisational measures and legal responsibility. For the purposes of data protection legislation, the data controllers for anonymised patient data held in local hospital repositories will be the data controllers in individual hospitals. The data controller for the overall Platform and for metadata and provenance files will be the partner responsible for the Platform.

The Medical Informatics Platform is evaluating additional procedural safeguards, regarded as best practice for large databanks of medical data (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2744675/>). An initial set of safeguards will be in place before the Platform comes online in Project Month 30. They will include the creation of a Data Governance Committee responsible for ensuring that the acquisition of data by the HBP Medical Informatics Platform and the use of the data by users of the Platform comply with relevant law, regulation and professional standards. The committee will include representatives of different stakeholders including data providers, and members of local IRB.

Privacy impact assessment. A recent stakeholder forum managed by WP12.3 and SP8 suggested that anonymisation techniques should be subjected to a Privacy Impact Assessment by an independent third party. SP8 agrees and is discussing how the recommendation should be implemented. Results will be reported in the next review.

Technical measures to enforce anonymisation. Metadata associated with clinical records stored in local hospital data repositories will routinely include details of the hospital archive from which the data was taken, the form of consent, the location of consent documents, measures to de-identify the data, and relevant data use agreements and plans (see below). The HBP provenance tracking system will make it possible to analyse these data and to retroactively exclude data that does not meet emerging ethical standards.

Informed Consent. Informed consent for use of patient data for scientific purposes is a requirement, not just of European and national data protection law, but also of general medical ethics, as codified in the Helsinki convention, and in European and national legislation on human experimentation.

There are differences among European policy-makers concerning the appropriate trade-off between the social benefits to be derived by exploiting patient records for research, and patients' right to control their personal data. These differences are reflected in differences in national law and jurisprudence. These issues were recently exposed in the controversy about the Care.data programme in the UK (<http://www.england.nhs.uk/ourwork/tsd/care-data/>), which has met severe obstacles, due to lack of adequate consultation with the citizens whose data was to be shared and inadequate procedures for patients to 'opt out.' However, Care data proposed to use



individual data. In the Ramp-Up Phase the HBP will use only aggregate data³ (see discussion of anonymity above)

Legal and regulatory differences are reflected in the different policies adopted by different European countries and in the debate on Europe's future Data Protection Regulation, where one of the themes is the use of anonymous health data for health research. SP10 will advise the BoD and the ExCo on the on-going debate. Currently, the use of patient data for research is regulated by national and local laws and regulations, as interpreted by local IRBs.

A key issue is the secondary use of anonymised historical data, collected for purposes of diagnosis or treatment, and during clinical trials. Given that, in many cases patients have not consented to the use of their data for research, this raises ethical issues. The organisations providing data to the HBP Medical Informatics Platform will apply applicable national regulations, deontological standards and accepted procedures, whose details will vary by country. At the time of writing, we have full information only for the CHUV, the Swiss hospital that leads SP8. Similar information will be collected for all participating hospitals and made available to the EC, before the data are deposited in the Platform (which will only start full operation in Project Month 30). No data will be collected from any institution without a data sharing agreement, in which the contributing institution certifies the informed consent procedures applying to the data and their approval by the relevant IRB. The metadata describing data sets and data records will always specify the form of consent given by the patient.

A second issue is how to reform models of informed consent for future patients. Many ethicists and clinicians would agree that data cannot be used unless patients have given their consent. However, there is growing recognition that it is not always possible to enumerate the different ways in which future researchers may wish to use specific classes of clinical data. This has led to suggestions for forms of "open consent" that cover a broad range of possible uses. Obviously the use of such consent procedures should itself be subject to ethical review by the relevant IRB. Clinicians and ethicists working in the HBP are participating in public and academic debate around this issue. In the meantime, they will work with local hospital administrations and ethics committees to develop informed consent procedures that provide meaningful protection to patients, while simultaneously enabling effective research.

Requests for Ethical Approval. No clinical data will be made available through the Medical Informatics Platform without the approval of the hospitals or other organisations holding the data. These organisations will be responsible for requesting ethical approval (where required) from the relevant Independent Review Boards. In the early stages, it is expected that the majority of these organisations will be located in Switzerland. At later stages, the Project will involve hospitals all over Europe. Staff in the HBP Ethics and Society Programme (SP10) and staff from the Medical Informatics SP will assist hospitals in formulating their applications and in providing any information IRBs may require.

³ In a late phase of the Project, the HBP plans to develop services for personalized medicine. These will probably be managed by a new legal entity set up for this purpose. Individual data used for such services will be covered by the confidentiality requirements of the doctor -patient relationship, and will not be made available to researchers. Informed consent procedures for patients contributing data to the service will conform to the legislation and ethical regulations in force at the time (late in the Project) when the service comes into operation.



The goal of the Project is that all necessary applications for ethical approval should be completed not later than the end of Month 24, and that all requests should be approved not later than Month 30.



8. Annex 3: Collaboration with the Danish Board of Technology Foundation

8.1 A Collaborative Approach

The Foresight Lab and the Danish Board of Technology Foundation (DBT) co-organised two webinars on 7 May 2014. One webinar focused on data federation and data protection, the other disease signatures and personalised medicine. A range of stakeholders considered the social and ethical issues raised by the plans of the MIP to federate clinical data from across Europe and mine them with the aim to isolate biological signatures of brain diseases. Details about the webinars and participant stakeholders are in Annex 8.2.

To build on the webinars and explore more thoroughly the challenges they had helped to identify, the Foresight Lab developed background material based on these challenges as well as on preparatory meetings and interviews with MIP members, MIP presentations and reports, and both scholarly and grey literature review. We elaborated two series of three narrative and fictional short scenarios, also called vignettes, one for each of the main topics: Data federation and privacy, and disease signatures.

The vignettes served as a basis for discussion in a follow-up stakeholder seminar that we organised again in collaboration with the Danish Board of Technology Foundation (DBT), which took place in Copenhagen on 8 and 9 October 2014. The DBTF planned it according to the ‘a dinner and a day’ approach to scenario-making, a concept originally pioneered by Shell. The event involved diverse external stakeholders as well as collaborators of the MIP. Participants were split in three working groups, each with a facilitator and a rapporteur. One group in particular focused exclusively on the topic of data federation and privacy. The discussions initiated from the vignettes both refined and expanded the range of issues that had been identified, and zoomed in on a number of short term challenges that were debated in greater depth.

The teams who collaborated in the organisation of the webinars and seminar have been using their results in various outputs. Besides our report on Future Medicine, the DBTF has published a policy brief on privacy and informed consent in the Human Brain Project. The results also feed into a European citizens’ convention organised by the DBTF and a citizens’ online deliberation forum organised by the Institute Pasteur (IP), Paris Institute.

8.2 Webinars

We reproduce below the programmes and background material prepared and circulated by the Danish Board of Technology Foundation prior to the webinars.

8.2.1 Webinar on data federation and protection: social, ethical and legal issues - Programme and Background material

8.2.1.1 Aims of the webinars

The webinars are a forum where Human Brain Project (HBP) stakeholders and relevant experts can have a constructive dialogue with regard to ideas, considerations and concerns arising from HBP research, discoveries and technologies. In particular, this webinar will aim at strengthening the HBP researchers' understanding of the broader social, political, ethical, legal implications of their work.

The objective of this webinar will be to discuss the ethical and social issues that may arise from the federation of data, data protection and their use to personalise medicine. This may also involve a wider consideration of the social implications related to the future of medicine.

During the webinar there will be four presentations, each followed by questions raised by a HBP panel (see timetable page 3).

8.2.1.2 The speakers

Dennis-Kenji Kipker is a member of the European Academy for Freedom of Information and Data Protection (EAID) and of the German Association for Law and Informatics (DGRI). He studied from 2006-2011 law at the University of Bremen with a focus on information and medical law, promoted by the German National Academic Foundation. Since 2011, he works as a research associate and a member of the universities Ethics Commission at the Institute for Information, Health and Medical Law in Bremen with a research emphasis on medical data protection law as well as police and intelligence agency law. After several publications and teaching assignments as a guest lecturer in Germany, Austria and Poland, he plans to finish his dissertation in autumn 2014. For more information: <http://www.jura.uni-bremen.de/personen/dennis-kenji-kipker/>

Dirk Lanzerath is General Secretary of the European Network of Research Ethics Committees (EUREC). He holds a doctoral degree in Philosophy (Dr. phil.) at the University of Bonn (1998); Habilitation (Philosophy) in Bonn (2013). He is Senior Research Assistant at the Philosophy Department (LFB II) of the University of Bonn (1993-1994); Senior Research Assistant at the Institute of Science and Ethics in Bonn (1994-1998); Head of the Research Department at the German Reference Centre for Ethics in the Life Sciences (DRZE), University of Bonn (1998-2002); since 2002 Executive Manager of the DRZE, University of Bonn. Since 1995 Lecturer of Philosophy at the University of Bonn; since 1996 Guest Lecturer at Loyola Marymount University, Los Angeles, Ca. He is also Member of the European Faculty of the Bioethics Program at the Graduate College of Union University Schenectady and Albany Medical College, NY, the Editorial Board of the Journal "Research Ethics Review," the Central Ethics Committee at the German Physician Association and the Review Board of UNESCO's Global Ethics Observatory Database (GEObs).

Paul Quinn is a researcher at the Vrije Universiteit Brussel. He has, in the past few years, conducted research on legal issues that surround legal issues in telemedicine and medicine in general focusing on patients' rights and privacy. He has worked on a number of European FP7 projects concerning such issues and has published his findings. Originally from the UK, Paul originally obtained a BSc in Biochemistry before converting to Law and



training as a Barrister. For more information:
<http://www.vub.ac.be/LSTS/members/quinn/>

Caspar Bowden is an independent advocate for information privacy rights, and public understanding of privacy research in computer science. He is a specialist in EU Data Protection, European and US surveillance law, PET research, identity management, and information ethics. He is author of 2013 EU Parliament inquiry briefing on the US FISA law, and co-authored the 2012 Note on privacy and Cloud computing (which anticipated the infringements to EU data sovereignty disclosed by Edward Snowden). For nine years he was Chief Privacy Adviser for Microsoft for forty countries, and previously co-founded and was first director of the Foundation for Information Policy Research (www.fipr.org). He was an expert adviser for UK Parliamentary legislation, and co-organised six public conferences on encryption, data retention, and interception policy. He has previous careers in financial engineering and risk management, and software engineering (systems, 3D games, applied cryptography), including work with Goldman Sachs, Microsoft Consulting Services, Acorn, Research Machines, and IBM. He founded the Award for Outstanding Research in Privacy Enhancing Technologies, is a fellow of the British Computer Society, and a member of the advisory bodies of several civil society associations.

8.2.1.3 The HBP panel

- Medical Informatics Platform: Thomas Heinis (Postdoctoral research Fellow in the Data-Intensive Applications and Systems Lab at the École Polytechnique Fédérale de Lausanne)
- Society and Ethics Program: Nikolas Rose (Professor of Sociology and Head of the Department of Social Science, Health and Medicine at King's College London)

8.2.1.4 The audience

- HBP members
- Researchers from the Medical Informatics Platform and the Society and Ethics Program
- Others from HBP: Ethics committees members, industry relations, communication team
- Stakeholders
- Researchers and experts coming from different backgrounds, such as bioethics, neuroscience, social science, cognitive science, the private sector, etc.

8.2.1.5 Webinar timetable

Time (CET)	Presentations
10:00 - 10:05	<i>Introduction by Chairman</i>
10:05 - 10:20	Dennis-Kenji Kipker
10:20 - 10:30	<i>Q&As by HBP panel</i>
10:30 - 10:45	Dirk Lanzerath
10:45 - 10:55	<i>Q&As by HBP panel</i>
10:55 - 11.10	Paul Quinn
11:10 - 11:20	<i>Q&As by HBP panel</i>



11:20 - 11:35	Caspar Bowden
11:35 - 11:45	<i>Q&As by HBP panel</i>
11:45 - 12:00	<i>Closing comments by HBP panel</i>

8.2.1.6 Background

This webinar is the first stakeholder activity arranged by the Danish Board of Technology in collaboration with King's College London - both part of Subproject 12, "the Ethics and Society" Program of HBP (<https://www.humanbrainproject.eu/ethics-and-society>). The task of the Danish Board of Technology in HBP is to promote engagement of outside experts and stakeholders and create dialog with HBP researchers.

In the Human Brain Project (HBP) it is the Subproject 8 that will deliver the Medical Informatics Platform (MIP), where they will transform medical records into research data, extract knowledge and build models of brain diseases. Inference from these models will forward the field of medicine from descriptive symptomatology toward diagnostic, predictive and prescriptive personalised medicine. Thus, it is the work done in their Subproject which is at focus in this twofold webinar on 1) data federation and protection and 2) disease signatures and personalised medicine.

In the following we will provide you with descriptions of the goals and methods for obtaining those that have been provided by the Subproject 8. The methods described below will focus on data federation and protection, as it is the theme of this webinar.

8.2.1.7 The overall goal of the Medical Informatics Subproject in HBP

The first goal of the Medical Informatics Subproject is to federate hospital and other clinical data on all brain diseases across multiple levels of biology for data mining and rule-based cluster analysis. The ultimate goal is to derive unique objective biological signatures of brain disease, which will be used to improve diagnoses, establish more accurate prognoses, and develop new drug discovery pathways.

Clinical data, tools for data mining, and rule-based clustering will be provided through a Medical Informatics Platform. The Medical Informatics Platform will become a service for a) clinicians for objective diagnoses and treatment of brain disease, b) neuroscientists to apply, test and input new models and methods, and c) pharmaceutical and biotech companies for disease target discovery. Unique biological signatures of a disease will also serve as prescriptions to configure brain models of disease for simulation-based predictions across inaccessible levels of biology, studying the causal chain of events in brain diseases, and performing drug simulations.

The MIP will provide end-to-end solutions from the data to advanced analytical tools. Researchers will be able to investigate questions requiring data correlations, distributions and interactions in the context of disease processes and epidemiological factors. Using analytical tools provided, researchers will be able to decipher the relationship between biological variables and clinical phenotypes. Simultaneously and continuously, as data accrue with time and recruitment of new hospitals and data generators, data mining tools will explore all the data to detect recurrent patterns identifying in this way biological signatures of disease. The biological signatures of disease will form the basis for a new disease space that neuroscientists and clinicians can explore.



8.2.1.8 Methods for data federation and protection

The MIP will federate data from multiple sources to allow researchers to interrogate them in a unified, transparent and efficient way. The privacy of data will be strictly maintained through rigorous anonymisation and access rules.

The MIP will build on public and research databases, and hospital data, federated by novel data management and query techniques. This federation software and hardware will allow researchers to query and analyse a very large volume of data without moving them from local servers and without compromising data privacy.

The anonymisation of the local data uses a two-pronged approach to accomplish its task:

- Personal identifiers shall be removed when exporting data from hospital information systems, i.e. even before the MIP accesses the data for the first time

The MIP shall only allow aggregate queries to run and shall filter all results (ensuring they do not contain personal patient information) before returning them to users of the Platform.

Hospital data are often stored locally in disparate sources, of various technologies and with different schemata. Each participating hospital will be requested to build a local data warehouse. This will allow data to be accessed and interrogated by the MIP in the most efficient and secure manner, fully maintaining its accuracy, consistency and privacy. An Extract/Transform/Load (ETL) process will be used for populating local data warehouses (called Local Data Store Mirrors).

All query results coming from a Local Data Store Mirror are filtered for any personal identifiers left (e.g. headers of imaging data) when they leave the Local Data Store Mirror and enter the Data Federation system. Data security and privacy follow principles established by data governance bodies, which include SP8 team and local data providers.

HBP's *in situ* query technology will make it possible to query data stored on local hospital servers without moving it to a central location. When implementing the technology, the HBP will adopt a pragmatic strategy that guarantees respect for differing local regulations, the privacy policies of hospitals and any additional restrictions on use requested by individual patients (e.g. restrictions on the use of specific categories of data). To meet this need, they will develop novel configurable security mechanisms both at the Platform level and the general data release level. At the Platform level, they will develop novel access control techniques, to be deployed on local servers, giving data owners (hospitals, research institutions etc.) fine-grained control over the configuration of security permissions (general policies, policies for individual patients, policies for individual items of data) and ensuring that authorised personnel maintain the ability to access individual data in an emergency (e.g. when a researcher finds an undetected tumour in neuroimaging data). They will also develop additional statistical tools that make it possible to quantify how much information has been disclosed to which researchers and in what form.

8.2.1.9 Data

The data provided by hospitals include raw data, such as imaging data (MRI, PET), CSV/XLS files, raw text, output of proprietary medical systems, as well as relational databases. The data traverse a series of sub-processes. They are anonymised, normalised and integrated. Finally, the data and the information extracted from them are released (i.e., made available for *in-situ* query) as Variables associated with Provenance descriptors.

8.2.1.10 Variables

Variables are metadata that provide scientific descriptions of data. Variables are extracted by a mixture of number crunching, image processing and text mining, and are mapped to a

common standard. Variable extraction data processing happens at the Data Store Mirror level in each data provider site, locally in most cases.

Below are the Variable descriptions, along five dimensions:

- Scale - genetic, molecular, cellular, circuits, systems.
- Time - data acquisition at single point/multiple points per subject.
- Space - centred on the brain (based on brain regions).
- Pathology - clinical labels of brain pathology.
- Demographics - age, gender.

8.2.1.11 Provenance

For each of the Variables the Provenance describes precisely the materials and methods (exact algorithms and their versions) used to create them:

- Materials: brain imaging - magnetic resonance imaging/computer tomography/PET, neurophysiology, etc.
- Methods: voxel-based morphometry - volume, surface-based morphometry - cortical folding, power spectra, SI units for biochemical assays, etc.

8.2.1.12 Ontologies

Ontologies are controlled vocabularies describing the Variables and relations among them with a meaningful grammar in the specific domain of interest (clinical phenotype, medical terms, brain regions, genes and proteins). In the first instance, we will use established ontologies for each domain, which will be updated as required. Examples of Ontologies: ICD-10/9 disease classification from the World Health Organisation adopted by hospitals, OMIM, DiseasesDB, eMedicine.

8.2.1.13 Data Integration and Federation

Data Integration is the process of merging data from different sources. Data Federation is a type of Data Integration. This is a process and virtual database that allows queries to heterogeneous and fragmented databases with a “no copy” and “no move” policy in relation to original source data. The virtual database contains only the Variables and the Provenance describing the original data in the local databases at each site in the network of MIP-associated infrastructures.

8.2.1.14 Ethical and regulatory issues

Subproject 8's work will lead to new techniques for the early diagnosis of neurological and psychiatric disease, and for personalised treatment (predictions of response to treatment; reduction of adverse drug reactions etc.), which the Medical Informatics Platform will make available to the research community and clinicians. The processes of data federation and the new techniques of data mining for disease signatures to guide clinical interventions raise immediate and longer-term legal, ethical and social issues. Issues of immediate concern include questions related to the use of anonymised data, data protection and informed consent. Many other questions arise from the wish to personalise diagnosis and treatment in neurological and psychiatric diseases.



8.2.2 Webinar on disease signatures and personalised medicine: social, ethical and legal issues - Programme and background material

8.2.2.1 Aims of the webinars

The webinars are a forum where HBP stakeholders and relevant experts can have a constructive dialogue on ideas, considerations and concerns arising from HBP research, and technologies.

The objective of this webinar is to discuss the ethical and social issues that may arise from the identification of disease signatures and their use to personalise medicine. This may also involve a wider consideration of the social implications related to the future of medicine. This webinar will strengthen HBP researchers' understanding of the broader social, political, ethical, legal implications of their work.

During the webinar there will be four presentations, each followed by questions raised by a HBP panel (see 8.2.1.5).

8.2.2.2 The speakers

Emilio Mordini is a Psychiatrist and Philosopher. He currently chairs *Responsible Technology*, a French consultancy devoted to responsible research and innovation. He has been formerly trained as a psychoanalyst and taught bioethics in the Medical School of the University "La Sapienza" of Rome (1994-2005). He served as a scientific secretary of the Bioethical Commission of the Italian National Research Council (1996-2004), and as a director of the Centre for Science, Society and Citizenship (2002-2013). He is a member of the management committees of the *COST Action IC1106 Integrating Biometrics and Forensics for the Digital Age*, and the *COST Action IC1206 De-identification for privacy protection in multimedia content*. He serves in a number of scientific and editorial committees. He has extensively published in academic peer reviewed publications, and edited 14 books. Emilio's research interests have focused on the notion of social unconscious - which derives from scholars such as Castoriadis, Hopper, Weinberg, and Taylor - and applied it to ethical and societal aspects of emerging technology and science.

Doug Brown is the Director of Research and Development at the Alzheimer's Society. Increasing investment in research is a key goal of the Alzheimer's Society which aims to invest more than £10 million a year in dementia research by 2017, and has committed to spending at least £100m over the next decade. The research programme will continue to provide a vehicle for funding across the spectrum of dementia research, into the cause, cure, care and prevention of dementia, and will focus on translating and implementing research findings for ultimate benefit of people affected by dementia. Doug is also a Trustee of the Association of Medical Research Charities.

Barbara Prainsack is Professor of Sociology at the Department of Social Science, Health & Medicine at King's College London. Her work addresses the regulatory, social, and ethical dimensions of bioscience and biomedicine. Recent publications include: *Genetics as Social Practice* (Ashgate, 2014; ed. with Silke Schicktanz and Gabriele Werner-Felmayer), and *Solidarity: Reflections on an Emerging Concept in Bioethics* (Nuffield Council on Bioethics, 2011; with Alena Buyx). From 2011-2013 Barbara led the European Science Foundation's Forward Look on 'Personalisation for the European Citizen.' For more information:

<http://www.kcl.ac.uk/sspp/departments/sshm/people/academic/barbaraprainsack.aspx>

Gregor Wolbring is Associate Professor in Department of Community Health Sciences, Stream Community Rehabilitation and Disability Studies, Faculty of Medicine, Calgary, Canada. He holds a diploma in Biochemistry at the University of Tübingen and at the



University College London, UK. Parallel to his biochemistry work he did policy work around disability issues and around emerging technologies. That culminated in a faculty position covering these topics in 2008. He has interest in ability governance, Social, ethical, legal, economic, cultural and governance issues of new, emerging and converging sciences and technologies (S&T), bioethics etc. For more information: http://www.crd.s.org/research/faculty/Gregor_Wolbring2.shtml

8.2.2.3 The HBP panel (asking questions following each presentation)

- Medical Informatics Platform: Ferath Kherif (Deputy Director, Laboratoire de Recherche en Neuroimagerie, CHUV/UNIL)
- Society and Ethics Program: Nikolas Rose (Professor of Sociology and Head of the Department of Social Science, Health and Medicine at King's College London)

8.2.2.4 The audience

- HBP members
- Researchers from the Medical Informatics Platform and the Society and Ethics Program
- Others from HBP: Ethics committee members, industry relations, and the communication team
- Stakeholders
- Researchers and experts coming from different backgrounds: bioethics, neuroscience, social science, cognitive science, and the private sector.

8.2.2.5 Webinar timetable

Time (CET)	Presentations
13:00 - 13:05	Introduction by Chairman
13:05 - 13:20	Emilio Mordini
13:20 - 13:30	Q&As by HBP panel
13:30 - 13:45	Doug Brown
13:45 - 13:55	Q&As by HBP panel
13:55 - 14.10	Barbara Prainsack
14:10 - 14:20	Q&As by HBP panel
14:20 - 14:35	Gregor Wolbring
14:35 - 14:45	Q&As by HBP panel
14:45 - 15:00	Closing comments by HBP panel

8.2.2.6 Background

This webinar is the first stakeholder activity arranged by the Danish Board of Technology in collaboration with King's College London - both part of Subproject 12, which is the Ethics and Society Program of HBP (<https://www.humanbrainproject.eu/ethics-and-society>). The



task of the Danish Board of Technology in HBP is to promote engagement of outside experts and stakeholders and create dialog with HBP researchers.

HBP Subproject 8 will deliver the Medical Informatics Platform (MIP), which will be used to transform medical records into research data, extract knowledge and build models of brain diseases. Inference from these models will advance the field of medicine from descriptive symptomatology toward diagnostic, predictive and prescriptive personalised medicine. Thus, SP8's work is the focus of this twofold webinar on 1) data federation and protection and 2) disease signatures and personalised medicine.

8.2.2.7 The overall goal of the HBP Medical Informatics Subproject

The first goal of the Medical Informatics Subproject is to federate hospital and other clinical data on all brain diseases and across multiple levels of biology for data mining and rule-based cluster analysis. These data will be used to derive unique objective biological signatures of brain disease, which will in turn be used for diagnosis, more accurate prognosis and new types of drug discovery for the development of new medicines.

Clinical data, tools for data-mining and rule-based clustering will be provided through a Medical Informatics Platform. The Medical Informatics Platform will become a service for a) clinicians for objective diagnoses and treatment of brain disease; b) neuroscientists to apply, test and input new models and methods; and c) pharmaceutical and biotech companies for disease target discovery. Unique biological signatures of a disease will also serve as prescriptions to configure brain models of disease for simulation-based predictions across inaccessible levels of biology, studying the causal chain of events in brain diseases, and performing drug simulations.

The MIP will provide end-to-end solutions from the data to advanced analytical tools. Researchers will be able to investigate questions requiring data correlations, distributions and interactions in the context of disease processes and epidemiological factors. Using the analytical tools provided, researchers will be able to decipher the relationship between biological variables and clinical phenotypes. Simultaneously and continuously, as data accrue with time and recruitment of new hospitals and data generators, data mining tools will explore all the data to detect recurrent patterns identifying in this way biological signatures of disease. The biological signatures of disease will form the basis for a new disease space that neuroscientists and clinicians can explore.

8.2.2.8 Methods for disease signatures and personalised medicine

Machine learning and data mining tools will provide a comprehensive classification of brain diseases based on biological signatures, i.e., parameterised combinations of biological features and markers. The biological signature of brain diseases will form the basis for a new multi-dimensional brain disease space, facilitating scientific investigation and permitting personalised medicine.

To secure anonymisation when MIP release data to contributors of the HBP, the software used in the HBP will preserve researchers' ability to select data based on patients' personal attributes (gender, age, country/region of origin, country/region of residency, ethnic origin, clinical diagnosis etc.) while providing mathematical proof that it is impossible to identify the presence of a specific individual in a given data set.

The HBP seeks to provide researchers in medicine and pharmacology with the tools they need to accelerate research into the causes, diagnosis and treatment of neurological and psychiatric disease. The HBP has four long-term aims:

- 1) To identify differential disease signatures from clinical data made available through the Medical Informatics Platform (Subproject 8) and to develop new nosological



classifications based on predisposing factors and biological dysfunctions rather than symptoms and syndromes.

- 2) To use biological signatures of disease as a source of insights into disease processes, testing specific hypotheses of disease simulation through modelling and simulation.
- 3) To use disease models to identify potential drug targets and other possible treatment strategies and to predict desirable and adverse effects.
- 4) To develop strategies for personalised medicine, allowing the development of treatments adapted to the specific condition of individual or specific subgroups of sensitive or vulnerable patients.

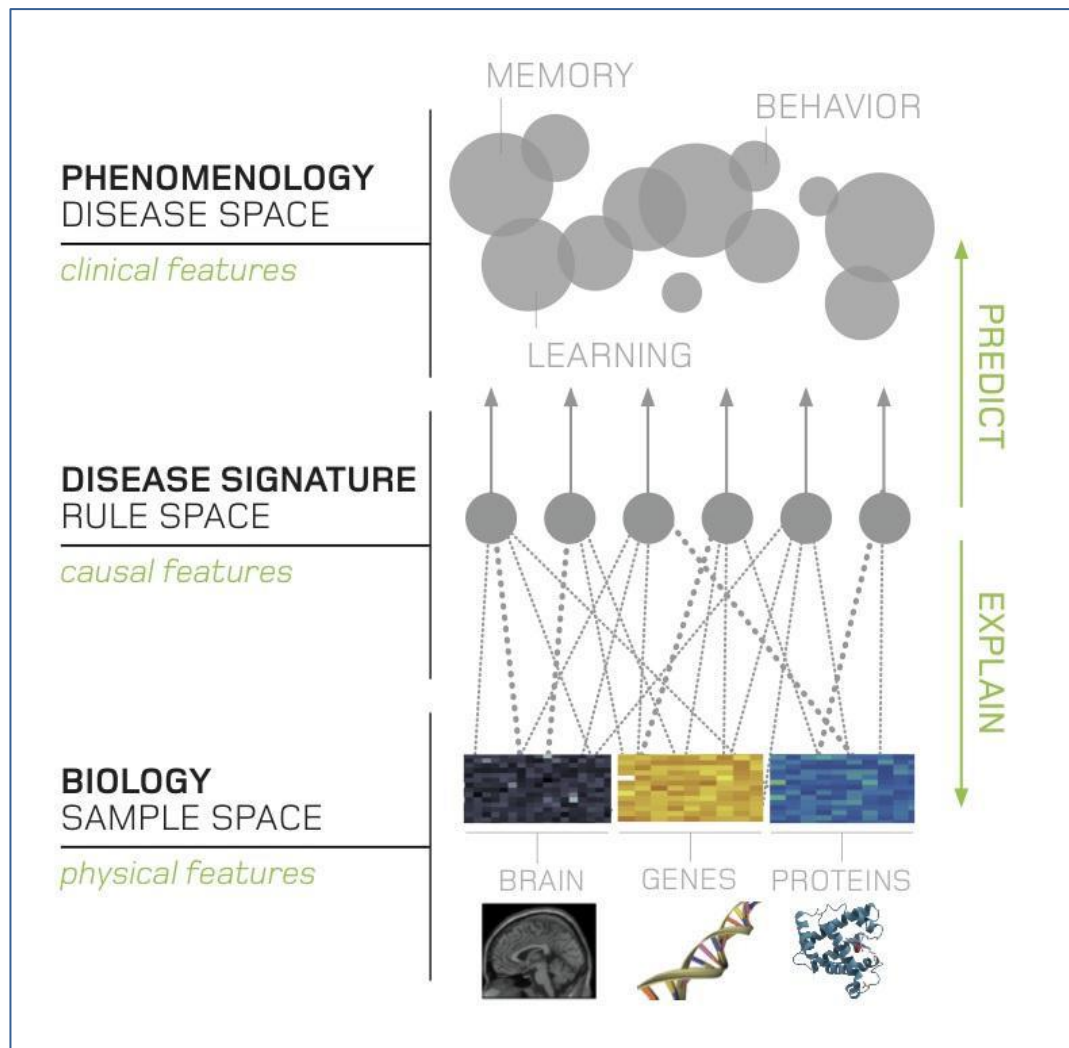
8.2.2.9 Biological Signatures of Diseases

The Biological Signatures of Diseases are deterministic mathematical constructs that aim to describe variability both at the phenomenological level (clinical features with symptoms and syndromes) and at the biological level (genetic, proteomic, etc.). Biological signatures of disease account for the fact that a symptom of brain dysfunction can be due to many biological causes (one-to-many symptom mapping) and that a biological cause can present with many symptoms (many-to-one symptom mapping). In reality, the situation is often one of many mappings between symptoms and biological causes. With advanced computing power and data mining, nearly exhaustive searches of a data space can be performed to identify sets of rules that describe homogeneous populations, to explain their biological data and to predict the pattern of symptoms.

Biological Signatures of Diseases are the results of a continuous dynamic data mining process of clinical data in local data sources. These results are aggregated to generate a single multi-modal description of the disease space.



Illustration of the biological signature of brain diseases/Continuous data mining process:



Example of how the disease signatures could be used by a clinician in neurology, Beth:

- Preconditions:
 - The biological signatures of diseases produced by the data mining algorithms are available at the MIP Web Portal.
 - The Variables that describe each disease signature cluster have been released and are available at the MIP Web Portal.
- Success scenario:
 - Beth is interested in taking forward personalised diagnostics using the biological signatures of the disease.
 - She selects the Biological Signatures of Diseases service and uses the interface to classify her own patient by comparing his clinical and biological characteristics with the whole range of provided biological signatures of diseases using an optimal matching algorithm.
 - She does this by selecting Variables of interest - e.g. demographic data, blood cholesterol, neuropsychological scores, genetic burden, etc.
 - She enters values for those Variables.



- She retrieves a list of disease signatures ordered according to the best match. The distribution of values of the other unselected Variables is also displayed along with their uncertainty - e.g. genotype, clinical scores, cardio-vascular risk factors.
- She also retrieves a 3D brain map with highlighted anatomical regions affected by the particular disease corresponding to the optimally matched disease signature. She can compare the map with the anatomy pattern of her own patients.
- Depending on how well the disease signature cluster matches her criteria, Beth can add new Variables to determine the stability of her classification in relation to the number of criteria or Variables used.
- She can compare the derived disease signature cluster to conventional clinical classification - e.g. ICD-10, DSM V classification.
- If needed, she can review her patients (data) to verify the derived disease signature cluster by similarity and by differences with other patients.

8.2.2.10 Ethical issues

Subproject 8's work will lead to new techniques for the early diagnosis of neurological and psychiatric disease, and for personalised treatment (predictions of response to treatment; reduction of adverse drug reactions etc.). The Medical Informatics Platform will make these tools available to the research community, and ultimately, to clinicians. The processes of data federation and the new techniques of data mining for disease signatures to guide clinical interventions raise immediate and longer-term legal, ethical and social issues. Issues of immediate concern include questions related to the use of anonymised data, data protection and informed consent. Many other questions arise from the wish to personalise diagnosis and treatment in neurological and psychiatric diseases. Issues of immediate concern include questions related to the use of anonymised data. Equally important are longer-term issues of therapeutic equity.

See this page for the ethical considerations presented at the HBP website:
https://www.humanbrainproject.eu/en_GB/faq/ethics.

9. References

¹ The HBP-PS Consortium, Lausanne, “The Human Brain Project. A Report to the European Commission,” April 2012.
https://www.humanbrainproject.eu/documents/10180/17648/TheHBReport_LR.pdf/18e5747e-10af-4bec-9806-d03ae57655

² http://ec.europa.eu/research/participants/portal/doc/call/h2020/h2020-fetflag-2014/1595110-6pilots-hbp-publicreport_en.pdf

³ Future and Emerging Technologies (FET),
http://cordis.europa.eu/fp7/ict/programme/fet/flagship/home_en.html

⁴ We use here the term ‘responsible research and innovation’ as the latest in a series of terms describing the evolving tradition of technology assessment (TA), or assessment of ethical, legal, and social aspects (ELSA) of research.

⁵ Collinridge, David (1980) *The Social Control of Technology* (Michigan: Francis Pinter); see also Rosenberg, Nathan (1994), *Exploring the Blackbox: Technology, Economics and History*, (Cambridge: Cambridge University Press); Arthur, Brian (1994) *Increasing Returns and Path Dependency in the Economy*, (Ann Arbor, MI: University of Michigan); David, Paul (1997), *Path Dependency and the Quest Historical Economics*, University of Oxford, Discussion Papers in Economic and Social History, No. 20.

⁶ Engineering and Physical Science Research Council (EPSRC-UK). This formulation was influenced by work published in Stilgoe, J., Owen, R. & Macnaghten, P. [Developing a framework of responsible innovation](#). *Research Policy*. 2013;42:1568-1580, which in turn drew from a long history of science and technology policy (STP) scholars experimenting with and defining methods of technology assessment.

⁷ European Cooperation on Science and Technology (COST), Cost 22 Action 2007

⁸ Wilkinson, Angela “Scenario Practices: In Search of Theory,” *Journal of Future Studies*, February 2009 13(3): 107-114

⁹ Before anything develops that might be seen as a capacity of a community (for example the disability community is a key constituency with whom the HBP must engage) to respond to various situations, members of the community must know that new brain research is occurring and that this is an issue for themselves and their community to consider, perhaps in view of making changes or preparations. Largely the work in the Ramp-Up Phase of the Human Brain Project can only be seen as this initial conception of capacity building. Genuine capacity building in any community is a considerably more elaborate and resource intensive process than provided for by initial foresight work.

¹⁰ Government Office for Science (UK), *Drugs Futures 2025*, London: Government Office for Science; Dwarakanath, T.R.; Wender, B.A.; Seager, T.; Fraser, M.P., “Towards anticipatory life cycle assessment of photovoltaics,” Photovoltaic Specialists Conference (PVSC), 2013 IEEE 39th, vol., no., pp.2392,2393, 16-21 June 2013 doi: 10.1109/PVSC.2013.6744956. For other projects please see the website of European Foresight at <http://forera.jrc.ec.europa.eu/projects.html>

¹¹ For more detailed description of various methods see E.U. research project http://forlearn.jrc.ec.europa.eu/guide/4_methodology/methods.htm. Many EU nations also have foresight projects. In the UK see <https://www.gov.uk/government/collections/foresight-projects>.

¹² The work of the DBT is formally part of Work Package 12.3 in the Ethics and Society work program of the HBP.



¹³ See for example, T. G. Beach, S. E. Monsell, et al., "Accuracy of the Clinical Diagnosis of Alzheimer Disease at National Institute on Aging Alzheimer's Disease Centers, 2005-2010," *Journal of neuropathology and experimental neurology*, vol. 71, p. 266, 2012.; S. Kloppel, C. M. Stonnington, et al., "Accuracy of dementia diagnosis: a direct comparison between radiologists and a computerized method," *Brain*, vol. 131, pp. 2969-74, Nov 2008.

¹⁴ Nuffield Council on Bioethics (2010) *Medical profiling and online medicine: the ethics of 'personalised healthcare' in a consumer age*. London: Nuffield Council on Bioethics; European Science Foundation (ESF) (2012) *Personalised Medicine for the European Citizen - towards more precise medicine for the diagnosis, treatment and prevention of disease*. Strasbourg: ESF; National Academy of Sciences (NAS) (2011) *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*. Washington, DC: NAS.

¹⁵ The sequence of publications on this issue by Steven Hyman, as Director and then former Director of the National Institute for Mental Health in the US, give a clear indication of these hopes and disappointments: Hyman, S. E. (2003). 'Diagnosing disorders', *Scientific American*, 289(3), 96-103; Hyman, S. E. (2008). 'A glimmer of light for neuropsychiatric disorders', *Nature*, 455(7215), 890-893; Hyman, S. E. (2014). 'Revitalizing Psychiatric Therapeutics', *Neuropsychopharmacology*, 39(1), 220-229. Hyman, S. E. (2012). 'Revolution Stalled', *Science Translational Medicine*, 4 (155), 155cm11. On the failure to find biomarkers to underpin diagnoses in DSM 5, see the views of Hyman's successor as Director of NIMH, Thomas Insel, at <http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml>.

¹⁶ Notably the RDoC approach proposed by the NIMH: <http://www.nimh.nih.gov/research-priorities/rdoc/nimh-research-domain-criteria-rdoc.shtml>.

¹⁷ J. W. Smoller, N. Craddock, K. Kendler et al., "Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis," *The Lancet*, vol. 381, no. 9875, pp. 1371-1379, 2013.

¹⁸ The term *diseasome* appears to have been coined by Albert-László Barabási, see Barabasi, A.-L., Gulbahce, N., & Loscalzo, J. (2011). 'Network medicine: a network-based approach to human disease', *Nat Rev Genet*, 12(1), 56-68 and http://www.nytimes.com/interactive/2008/05/05/science/20080506_DISEASE.html?_r=0

¹⁹ For the current situation, please see the relevant webpage of the European Commission: <http://ec.europa.eu/justice/data-protection/>

²⁰ As we shall see, in the light of the discussions during the course of preparing this foresight report, the MIP now intends to seek informed consent wherever possible, and to ensure that all its procedures are fully supported by local research ethics committees before the use of any data.

²¹ In a rather different area, we have seen such controversies over the provision of very expensive drugs for individuals with last stage cancer, where response is uncertain, and the gains in life expectancy are small.

²² Singh, I., & Rose, N. (2009). 'Biomarkers in psychiatry', *Nature*, 460(7252), 202-207.

²³ A further possibility suggested by one of our reviewers would be where an individual's identity was inferred by a 'harmful intruder' as suggested in the widely publicized work of Yaniv Erlich at the Whitehead Institute, see <http://wi.mit.edu/news/archive/2013/scientists-expose-new-vulnerabilities-security-personal-genetic-information>

²⁴ We should point out that the *fear* of genetic discrimination on the basis of such data are currently more important than the *reality*, since in most jurisdictions there is a moratorium that prevents or restricts discrimination by insurance companies, employers or others on the basis of genetic information. Nonetheless, public concerns about genetic discrimination persist in many jurisdictions: see, for example Geelen, E., Horstman, K., Marcelis, C. L. M., Doevendans, P. A., & Van Hoyweghen, I. (2012). 'Unravelling fears of genetic discrimination: an exploratory study of

Dutch HCM families in an era of genetic non-discrimination acts', *Eur J Hum Genet*, 20(10), 1018-1023.

²⁵ One of our reviewers pointed out that although hospitals currently hold a great deal of 'personal data' in electronic form and hence are liable to breaches of data security, these are actually very rare, and most citizens regard their medical providers as highly trustworthy. However our Foresight works shows that new concerns about 'big data' security arise when there are proposals for these data to be networked and aggregated.

²⁶ It should be pointed out that the technical infrastructure used by the MIP has been designed to be relatively cheap - at about 20-50,000 Swiss Francs per participating hospital, although this sum does not include the costs of training and maintenance.

²⁷ Whether or not such consent issues could be a 'fatal hurdle' depends in part on whether the de-identification and data aggregation technologies deployed by the MIP are deemed to produce full de-identification and hence not subject to any revision of the European Union's data protection regulations.

²⁸ Available at URL: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

²⁹ Activity led by WP 8.1 in HBP SP8 - Medical Informatics Platform.

³⁰ Activity led by WP 8.4 in HBP SP8 - Medical Informatics Platform.

³¹ For more information on Huntington Disease patients' engagement, see the webpages of The International Huntington Association (<http://www.huntington-assoc.com/>) and of Enroll-HD, a Prospective Registry Study in a Global HD Cohort (<http://www.enroll-hd.org/html/about?enrollsid=e376d1ad02458d0cb82222fc4afad80b>).

³² Acquisti, Alessandro, Brandimarte, Laura, and Loewenstein, George, "Privacy and human behavior in the age of information," *Science* 30 January 2015, Vol. 347, Issue 6221, pp. 509-514, DOI: 10.1126/science.aaa1465.

³³ It is, of course, not accidental that this concern is most evident in Germany, given the horrific consequences of persecution and elimination of individuals thought to carry genetic disabilities in the 1930-145 period.

³⁴ Nuffield Council on Bioethics, "The collection, linking and use of data in biomedical research and health care: ethical issues," February 2015 (<http://nuffieldbioethics.org/project/biological-health-data/>).

³⁵ This is not a problem specific to the Medical Informatics Platform of the HBP, but more generally of data initiatives involving the re-use of data in novel contexts, as is highlighted in the Nuffield Council on Bioethics report of February 2015.

³⁶ For a landmark positioning paper on Enterprise Security Information Architecture, see whitepaper by Gartner: Kreizman, Greg, and Robertson, Bruce, "Incorporating Security Into the Enterprise Architecture Process," 24 January 2006, available at URL: <https://www.gartner.com/doc/488575>.

³⁷ We are grateful to David Ford (Professor of Health Informatics and leader of the Health Informatics Group at Swansea University College of Medicine, Wales, UK) for sharing with us his experience and expertise as joint lead of the Health Information Research Unit for Wales (HIRU), which main product is the SAIL Databank, an internationally recognised data linkage resource formed from a wide variety of routinely collected data from across Wales. For more information, look at the SAIL Databank webpages (<http://www.saildatabank.com/>). See also: Jones, Kerina H., McNerney, Cynthia L., and Ford, David V., "Involving consumers in the work of a data linkage research unit," *International Journal of Consumer Studies* Vol. 38 (2014), pp. 45-51; Jones, Kerina H., et al., "A case study of the Secure Anonymous Information Linkage (SAIL) Gateway: A privacy-protecting remote access system for health-related research and evaluation," *Journal of Biomedical Informatics* Vol. 50 (2014), pp. 196-204.



³⁸ INVOLVE (2014) Guidance on the use of social media to actively involve people in research. Eastleigh: INVOLVE. For more information, go to the INVOLVE website (<http://www.invo.org.uk/>).

³⁹ We have not discussed here the conventional concerns about how to handle incidental findings. These are unlikely to be an issue for the aggregated data utilised in the MIP. However all researcher and research groups in the HBP are familiar with and committed to the well-established protocols for the communication of incidental findings, and these are included in all procedures for informed consent; should potential subjects not consent to the procedures for feeding back clinically relevant incidental findings, either to them directly or via their medical practitioner, they will not be included in the study.

⁴⁰ A variety of other approaches are also under exploration: recurrent least square support vector machines, topological graph analysis, Schwartzian space based analyses (eg, Hypercube), among others. This is one of the most active areas of research in the MIP at present. For the purpose of this report we have selected one early proof of principle study for description to give a sense of method and to contextualize participant remarks, questions, and concerns.

⁴¹ Tal Galili, Alexis Mitelpunkt, Netta Shachar, Mira Marcus-Kalish and Yoav Benjamini. “Categorize, Cluster and Classify: a 3-C strategy for scientific discovery in the Medical Informatics platform of the Human Brain Project” in Džeroski, S., Panov, P., Kocev, D., Todorovski, L. (Eds.) *Discovery Science, 17th International Conference, DS 2014, Bled, Slovenia, October 8-10, 2014, Proceedings* Springer Press 2014.

⁴² For example, from a genetic point of view, Single Nucleotide Polymorphisms (SNPs) are associated with mental disorder but are often pleiotropic (having more than one effect). J. W. Smoller, N. Craddock, K. Kendler et al., “Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis,” *The Lancet*, vol. 381, no. 9875, pp. 1371-1379, 2013.

⁴³ While this view may not be widely accepted, the UK Alzheimer’s Society does recommend a variety of talking therapies for those diagnosed with this condition: http://www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=92

⁴⁴ Rose, Nikolas *The Politics of Life Itself* Princeton University Press 2007.

⁴⁵ See, for example, the public dialogue on stratified medicine conducted by OPM for Sciencewise in the UK in 2014: <http://www.sciencewise-erc.org.uk/cms/assets/Uploads/Stratified-medicine-a-public-dialogue-Final-report.pdf>

⁴⁶ Campbell, T. *Dyslexia: The Government of Reading*. London: Palgrave Macmillan 2013.

⁴⁷ Wolbring, Gregor “Disability Rights Approach to Bioethics” *Journal of Disability Policy Studies* Winter 2003 vol. 14 no. 3 174-180 doi: 10.1177/10442073030140030701. More generally see Wolbring “The Politics of Ableism” *Development* 51, 252-258 (June 2008) doi:10.1057/dev.2008.17

⁴⁸ Swan, Melanie, “Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking” *Int. J. Environ. Res. Public Health* 2009, 6(2), 492-525; doi:10.3390/ijerph6020492

See also Wolbring, Gregor, Lucy Diep, Sophya Yumakulov, Natalie Ball, Verlyn Leopatra and Dean Yergens, “Emerging Therapeutic Enhancement Enabling Health Technologies and Their Discourses: What Is Discussed within the Health Domain?” *Healthcare*, 2013 1, 20-52; doi:10.3390/healthcare1010020

⁴⁹ A more long term concern of the project might occur if new medical innovations enable improvements to mental abilities. This is already an identifiable debate within student communities where drugs designed to address attention deficit disorder are now commonly used to improve academic performance. Ability creep was defined by one of our webinar participants as being when enhanced abilities become new norms or expectations. This process might decrease the security (employment, identity, resource access, etc.) of people with different, disabled, or typical (unenhanced) abilities.



⁵⁰ Morris, Z. S., Wooding, S. & Grant, J. “The answer is 17 years, what is the question: understanding time lags in translational research” *The Journal of the Royal Society of Medicine*. 104, 12, p. 510-520 (Dec 2011).

⁵¹ See, for example, the work of the UK’s Technology Strategy Board (now called Innovate UK) on the ‘roadmap’ for the implementation of personalized or stratified medicine: <https://connect.innovateuk.org/documents/2843120/3724280/Stratified+Medicines+Roadmap.pdf/fbb39848-282e-4619-a960-51e3a16ab893>

⁵² European Science Foundation, *ESF Forward Looks: Personalised Medicine for the European Citizen: Towards more precise medicine for the diagnosis, treatment and prevention of disease* ESF: Strasbourg 2012.

⁵³ Lunshof, Church and Prainsack have argued that access, control, and ownership of data entail slightly different relations of a person or entity to data. Access allows a person to know what information is held in relation to themselves and may create greater research transparency. Control allows a person to withdraw that information from circulation. Ownership is a more substantial set of rights which entitles the person to a proprietary relationship with the information: Lunshof, J.E., Church, G., and Prainsack, B. (2014). Raw personal data: providing access. *Science* 343/6169: 373-374.

⁵⁴ Prainsack B (2014). Personhood and solidarity: What kind of personalized medicine do we want? *Personalized Medicine* 11/7: 651-657.

⁵⁵ Joseph Schwartz, “The soul of soulless conditions? Accounting for genetic fundamentalism” *Radical Philosophy* 86:2-5 (Nov/Dec 1997).

⁵⁶ It should be pointed out that the technical infrastructure used by the MIP has been designed to be relatively cheap - at about 20-50,000 Swiss Francs per participating hospital, although this sum does not include the costs of training and maintenance.