Machine Learning and Big Data for Neuro-Diagnostics: Opportunities and Challenges for Clinical Translation

A briefing report for the Human Brain Project

Christine Aicardi

Saheli Datta

Tara Mahfoud

Nikolas Rose

The Foresight Lab

Department of Global Health & Social Medicine

King's College London

Table of Contents

1. The scope of this report 3
2. The Role of Technology in Diagnosis of Brain Disorders 4
2.1 Revising Diagnostic Categories 4
2.2 Making sense of the data avalanche 5
2.3 Artificial Intelligence and Diagnosis 6
3. The Search for Brain Signatures in the HBP 7
3.1 Neurodegenerative disorders 8
3.2 Epilepsy: F-TRACT and The Virtual Epileptic Patient 9
3.3 Psychiatric Disorders 9
4. Clinical Translation and the Challenge of Trust 10
4.1 Trust in patient-clinician relationships 11
4.2 Clinician-driven innovation vs. researcher-driven innovation 12
4.3 Trust and Epistemological Conflict between Clinicians and Researchers 13
5. Challenges to commercialisation 14
6. Conclusion: Challenges along the pathways to the clinic 16
6.1 Transparency is crucial 16
6.2 Involvement of clinicians upstream is crucial 16
7. Acknowledgements 17
8. References 17

Machine Learning and Big Data for Neuro-Diagnostics: Opportunities and Challenges for Clinical Translation

1. The scope of this report

In this Report, The Foresight Lab of the Human Brain Project (HBP) will examine some developments in neurodiagnostics that make use of machine learning and other algorithms, with a particular focus on the potentials and challenges for clinical translation. This is part of our commitment, in the current phase of the HBP,[1] to undertake research on the contributions of the HBP to the state of the art in relation to psychiatric and neurological diagnostics, with a focus on the role of brain signatures and their clinical, personal, social and ethical implications. These issues have to be understood in the context of more general questions concerning the social and ethical implications of brain research for psychiatric and neurological clinical practice, and, more widely, in relation to identifying biomarkers that will allow more accurate and individualised diagnosis and treatment, and potentially, in the future, could enable preventive intervention for those at risk of developing disorders.[2]

These are large issues, and this report focuses on concerns specifically relevant to the HBP.[3] In the HBP, work is underway within the Medical Informatics Platform (MIP) and throughout various work-packages in other sub-projects to analyse large clinical and research datasets, using algorithms and machine learning, to search for patterns in data that can individuate the neurobiological correlates of a disorder in ways that could be used to aid diagnosis, to target treatments and hence to improve prognosis. During the early phases of the HBP, although the focus was particularly on the dementias, the hope was that the collection and analysis of patterns in large volumes of existing clinical data held in hospitals would be possible to identify 'brain signatures' for other mental disorders – a hope that was shared with other large international research initiatives at the time.[4] However for a range of reasons which we discuss later in this report, it proved difficult to access and curate the volume of clinical data that was required for analysis. Hence, in the current phase of this work, the focus has been narrowed to disorders where there are clear neurobiological correlates, and where there is extensive data potentially available for analysis.[5]

The empirical research upon which this report is based includes analysis of published and grey literature, fieldwork, attendance at relevant workshops and conferences, in addition to stakeholder interviews (in the UK and Europe). While we do refer to some wider issues, we shall largely focus on two areas in this report: dementias and epilepsy, which although perhaps the least problematic of domains, nonetheless illustrate challenges that will be all the more significant in domains where neurobiological evidence of brain anomalies underlying disorders is more ambiguous. As the ultimate aim of development of diagnostic algorithms is for their use in the diagnosis and treatment of patients, we shall focus particularly on the possibilities and challenges of clinical translation. We draw attention to the challenges faced in relating probabilistic

1 The current phase of the HBP is that funded under the second Specific Grant Agreement with the European Commission, referred to here as SGA2.
2 We were asked to consider these wider issues by reviewers of our previous reports.
3 For the broader discussion of the role of brain research and neurobiology for psychiatry and mental health, we refer the reader to Professor Nikolas Rose's book *Our Psychiatric Future: The Politics of Mental Health*, published in October 2018.
4 We discuss these in our first Foresight Report: Aicardi, Christine, Michael Reinsborough, and Nikolas Rose. Foresight report on future medicine: A Report from the HBP Foresight Lab (2015), https://kclpure.kcl.ac.uk/portal/files/86508529/KCLForesightLab_2015_Future_Medicine.pdf
5 Data is also being collected on traumatic brain injury, and in addition the interest in wider issues in mental disorder remains, and a new partner has been recruited to the HBP who will contribute large data sets on mental disorder, however this work is not yet integrated into the MIP.

predictions derived from such algorithms to individualised clinical interventions, and we highlight the importance of trust in the relationships that enable clinical translation of technologies – trust between researchers, clinicians, patients, and regulators.

2. The Role of Technology in Diagnosis of Brain Disorders

2.1 Revising Diagnostic Categories

Much discussion of diagnosis has focussed on its craft like nature, and its dependence on the skilled 'gaze' of the clinician (Feinstein, 1973; Mattingly and Fleming, 1994). However, technology has long played a key part in diagnosis, at least since the invention of the stethoscope in the eighteenth century. Laboratory science has played an increasing role in developing new tools and techniques for diagnostic practices, and the introduction of microscopes, blood chemistry, EEG, imaging technologies such as CT, PET and MRI have sometimes led to the re-classification of diseases based on their underlying pathophysiology.

A belief in the necessity of such reclassification has emerged in psychiatry in the last decade. This was because, despite a large research effort over several decades, involving genetic, scanning and other advanced neurotechnologies, with the exception of some forms of dementia, it proved impossible to identify neurological 'biomarkers' for any of the current diagnostic categories used in clinical practice, for example those embodied in the successive editions since 1980 of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (Rose, 2013). Many psychiatric researchers, notably at the US National Institutes of Mental Health, argued that we should no longer base our research on such diagnostic categories, but develop new approaches that would diagnose disorders on the basis of their biology. They argued that this would almost certainly radically revise our diagnostic categories, revealing similarities in aetiology and neurobiology between disorders that had been kept distinct, and show that broad diagnostic categories such as major depression actually encompass a number of conditions with distinct aetiologies linked to distinct neurobiological bases which should be the focus of diagnosis and treatment.

This approach is now often referred to as 'precision medicine'[6] because it "aims to identify somatic, cognitive, affective, motor and social behaviour domains defined by associated, potentially common, aetiological neural mechanisms. This is in contrast to existing diagnostic criteria which are usually based on patient report, observation and duration of symptoms and do not incorporate biological or neuropsychological markers (Kapur et al., 2012). The hope is that the "aetiology-based research cutting across diagnostic boundaries may be a critical step towards overcoming what some view as "therapeutic stagnation in psychiatry" and …may offer a new and rational path for the development of a stratified psychiatry" (Schumann et al, 2010: 7). The launch of the U.S. National Institute of Mental Health's "precision medicine for psychiatry" project *Research Domain Criteria* (RDoC) in 2010 and the development of such a research strategy supported by the EC H2020-funded *Roadmap for Mental Health Research* (ROAMER, 2013) energised such approaches to psychiatry that aimed to link neural, cognitive and behavioral dimensions. However, at the time of writing this report, this approach has not succeeded in identifying clinically useful neurobiological or genetic biomarkers for diagnostic precision or treatment choice in the area of mental health. Hence the hope that large data sets that contain information from genetic tests, brain scans and other physiological markers, together with data on clinical presentation,

6 It was previously referred to as 'personalised medicine'; later, when it became evident that available technologies would identify, not individuals, but groups with particular biological characteristics, some researchers and organizations chose to use the term 'stratified medicine'. Both of these terms remain in use, leading to some confusion.

symptomatology, treatment and prognosis, when analysed using machine learning techniques, may reveal previously hidden relations between neurobiology, symptomatology and treatment success.

2.2 Making sense of the data avalanche

It is widely recognized that it is challenging to interpret the large quantities of data that are now available in many medical specialisms, including psychiatry and neurology. For example, while neurologists such as Kurt Goldstein (1878-1965) had access to the brains of patients' only after death, and thus diagnosed living patients using case records and visual technologies such as films, now neurologists have access to the living brain via a range of imaging technologies which they use in the diagnostic process - CT, PET, MRI, fMRI, EEG, etc. (Rose and Gainty, 2019). Nonetheless, these images do not speak for themselves and, as in previous diagnostic practices, require the intervention of the trained eye of the expert (see Mahfoud, 2014 for a more detailed overview of these studies). For example, Barry F. Saunders studied the craft practices involved in "learning to read" radiological images such as CT scans. He outlined the complex institutional and hierarchical context within which doctors are trained, with special attention given to tacit practices - or practices that are taught outside of formalised education, that require apprenticeship (Saunders 2008).

Indeed, neurologists we have interviewed for this report highlighted how learning to see brain images was never something they were formally taught as part of their medical degrees, but was a kind of tacit knowledge (Collins, 2010) acquired though their apprenticeship. Anthropologist Andreas Roepstorff uses the term "skilled vision" for this kind of tacit knowledge (2007). While the practices of interpretation are now part of formal education of medical practitioners, they must still be integrated into the 'expert gaze' of neuroscience researchers as well as clinicians. The skills developed by neuroscientists to be able to interpret fMRI brain scans involve both formal and informal education - one is required to know in order to see. Practitioners are not required to understand the mathematics involved in producing these images in order to 'read' and analyse brain scans, but, as we shall see, they do need to have confidence in the process that has led from the initial data to the images that they have to interpret and utilise.

Despite the recognition of the 'craft skills' of interpretation required across clinical medicine, there have long been attempts to codify the diagnostic process, for example in the interpretation of medical imaging, and to formalise it in standard and computerised programmes, with limited success (Doi, 2007). However in an increasing number of areas, machine learning is now being used to analyse case records, clinical and physiological data, test results and images from scans and to link these to diagnosis and prognosis – in some cases producing results that are more accurate and reliable than those of even the most skilled diagnostician (EMERJ, 2019; Nuffield Council on Bioethics, 2018; Future Advocacy, 2018). This has become an area of considerable commercial investment,[7] and it is in this domain that we can locate the current work being undertaken in the MIP of the HBP.

In assessing the social and ethical implications of these developments, it is important to recognise that technological artefacts, whether simple devices such as stethoscopes or blood pressure monitors, or highly sophisticated apparatuses such as scanners, are not merely extensions of the naked eye. The images that they produce not only depend upon very sophisticated technology – magnetic resonance imaging uses magnetic field gradients that act upon certain atoms to generate detectable radio waves that are then processed using sophisticated algorithms to generate data on the distribution of water and fat in the body that is then further processed to generate images of the organs - but also embody many assumptions. For example, the hypothesis

7 In the United States, this has been spurred on by evidence on the cost of settlements for medical malpractice as a result of diagnostic errors. See the analysis by researchers at Johns Hopkins reported here: https://www.hopkinsmedicine.org/news/media/releases/diagnostic_errors_more_common_costly_and_harmful_than_treatment_mistakes

built into fMRI is that an increase in blood flow is a marker of an increase in brain activity; this is embedded in the images produced by the device itself, and despite longstanding critical evaluation, this is seldom questioned in the practices that utilise fMRI for research or diagnosis (Logothetis, 2008). However, the standard fMRI paradigms have become controversial with the increasing recognition that measures of changes in blood oxygenation levels neglect the key role of highly distributed neural activity – known as the 'resting state' - that is necessary for task performance, but is usually 'subtracted' from fMRI outputs as a result of the algorithms that are used to create the images (Gusnard and Raichle, 2001; Raichle, 2011; Callard and Margulies, 2014).

Technology is thus far more than an aid to vision: it renders some things visible at the expense of others, yet frequently does so in ways that are 'black boxed' and not known or fully understood by those who use the results. These issues are further complicated when large volumes of data produced by these technologies in many different clinics and research projects, using different research protocols, are linked and then analysed using machine learning to generate algorithms – once more not known or fully understood by potential users - that produce a medical diagnosis. It is to this question that we now turn.

2.3 Artificial Intelligence and Diagnosis

One problem faced by those seeking to interpret the range of information potentially available for diagnosis, is that the masses of data from different sources now available - fMRI, MRI, PET, CT, EEG, MEG… - are extremely difficult for human beings to integrate and analyse together. It is this problem that machine learning promises to solve. The hope is that machine learning tools can be used to analyse these large amounts of data from different sources, and further, that these computer-driven methods will be more objective than person-driven analyses, because they do not depend on the interpretive skills of different clinicians. Thus many argue that the digitization of data from patients' clinical records and medical images combined with advanced data analytics can enable AI technologies such as machine learning and machine vision to distinguish between different potential diagnoses in a clinically meaningful way and thus can enable clinicians to target specific treatments (Luo et al., 2016).

However, many of these discussions have concluded that these 'new' technologies can only assist, not replace, expert opinion (Stone et al., 2016). Such use of technology is often referred to by practitioners as 'augmented intelligence'. Data-driven technologies aimed at assisting healthcare practitioners in diagnostic processes have been approved by the US Food and Drug Administration (FDA) in recent years (Future Advocacy, 2018). For example, the smartphone application 'Viz.AI' analyses CT images of the brains of patients admitted to hospitals with symptoms of stroke, identifies vessel blockages through these images, and sends this analysis via text to neurovascular specialists. This software was approved by the FDA as a "clinical decision support software" based on evidence submitted by the developers which demonstrated through a clinical trial that the software application more quickly identified the vessel blockages.[8] There are several other software applications under development (i.e. not yet approved by the FDA), such as the collaboration between DeepMind Health and Moorfields Eye Hospital, London which uses neural networks to diagnose Age-related Macular Degeneration (AMD) through the analysis of Optical Coherence Tomography (OCT) (De Fauw et al, 2018).

Despite the problems that many have identified with current diagnostic categories, much psychiatric research in this area still begins from or utilises such diagnoses. Thus a collaboration between IBM and the University of Alberta uses neural networks to diagnose schizophrenia through an analysis of fMRI scans while patients undertook an audio-based exercise. The researchers claim to have identified "combinations of statistical features extracted from the data that can serve as reliable statistical (bio)markers of the disease, capable of accurately discriminating between schizophrenic patients and controls" (Gheiratmand et. al. 2017). These 'bio-markers' included an

8 https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm596575.htm

"abnormal" increase of connectivity between the thalamus and the primary motor/primary sensor cortex as well as "hyperconnectivity" in the fronto-parietal network. While the model was relatively successful at predicting a clinical diagnosis of schizophrenia (at above 70%), this research has not yet undergone the clinical trials needed for software regulation and approval. Further, it is being undertaken at a time when the very categories such as schizophrenia are contested by many experts in the field. There is a widespread recognition that, across the whole spectrum of mental disorders, and particularly in relation to psychoses, similar symptomatology may result from very different neurobiological pathways. There is a certain unhelpful circularity in seeking brain based biomarkers that correlate with symptom based diagnoses that are themselves contested and considered to lump together a variety of conditions that are developmentally, neurobiologically and prognostically distinct (Murray, 2016).

Some six years ago, an Editorial in *Nature Biotechnology* entitled "What happened to personalized medicine?" reflected on the slow progress and unrealized hopes of those who predicted a revolution in medical diagnosis and treatment targeting based on biomarkers. The barriers they identified were less biological than social: a need "to broaden the concept of personalized medicine from the genetically reductionist version to one that includes other types of markers"; a need for more long term studies "linking specimens, sequence and other biomarker information to clinical outcomes", a need for patients to be encouraged to share their data for research purposes, and a need to educate physicians "about the new diagnostics and how to integrate them with existing clinical information" which will require not only better education but also "the development of robust point-of-care devices and data-sharing technology and the establishment of trusted sources (e.g., medical association position statements on tests or the National Institutes of Health's genetic testing registry) (Nature Biotechnology, 2012). Indeed the issue of trust is crucial, for if clinicians and patients do not have legitimate trust in the accuracy, validity and utility of biomarkers, and that certainly includes brain based biomarkers, whatever the hopes of those who develop them, they will not 'translate' into clinical practice.

However, before we consider this issue of translation, and the key role of trust, we will return directly to the work of the MIP of the HBP, for, as we have said, for the above reasons as well as others, the MIP has restricted its current focus to conditions where there is little or no dispute about the links between symptomatology and brain based anomalies. As we have said, we will focus on epilepsy and dementia.

3. The Search for Brain Signatures in the HBP

The initial aim of the HBP's Medical Informatics Platform (MIP) – which formed Sub-Project 8 (SP8) in HBP research up to the end of March 2020 - was to be "a collaborative open source platform … that allows researchers worldwide to share medical data, enabling the use of machine-learning tools for brain-related diseases, while strictly preserving patient confidentiality."[9] With this goal in mind, during the current phase of research,[10] the MIP prioritised the need to generate scientific impact from brain disease research. This involved consolidating and accelerating the readiness of the platform for end users, and giving careful consideration to the range of disorders that could realistically be addressed by the MIP.[11]

9 https://www.humanbrainproject.eu/en/about/project-structure/subprojects/#SP8, 26/11/2018

10 This refers to work in the period of the Second Grant Agreement (SGA2) to fund the HBP from 1 April 2018 to 31 March 2020.

11 In SGA3, there is a proposal to develop what has been termed MIP Central, which will contain data that is not public, and which requires special treatment to ensure privacy that cannot be carried out in the existing federated structure – for example the 'defacing' of data from patients undergoing intracerebral recording which can be used to reconstruct identifiers of specific individuals. The MIP Central will also seek to develop specific encryption technologies, using a unique identifier, that can connect all data from a given patient, but without it being possible for that to be traced back to that specific individual. As this proposal is currently at an early stage, we shall not discuss it in this report.

Following extensive discussion and refinement of the technical infrastructure design, in part to address the privacy and confidentiality concerns raised in a previous report by the Foresight Lab[12] and the subject of the first Opinion by the Ethics and Society Division of the HPB[13], the MIP uses a data collection architecture employing MIP-Local terminals deployed in hospitals and clinical research settings, and federating them in ways that aim to both preserve the anonymity of the patients to whom the data refers, and to enable researchers to explore the aggregated data to identify potential correlations between brain related data and clinically relevant symptomatology. Developments proposed in the current phase of work include ongoing development and refinement of the interfaces between data providers and researchers, with new end-users of the interface between the MIP-Federate layer and the MIP-Local terminals deployed in hospitals and clinical research setting. It also entails the development of new data alignment and mapping tools to connect the data from hospitals, and from a range of cohorts collected in the course of clinical trials or other research. It also requires the integration and extensive testing of data analytics algorithms and methods into the MIP – which is being complicated by more stringent data anonymisation requirements to conform with the General Data Protection Regulation (GDPR).

3.1 Neurodegenerative disorders

Neurodegenerative disorders (Alzheimer's, Parkinson's) have so far been the focus of the data analysis in SP8. A core dimension of the MIP is the development of data analytics tools that will be accessible through MIP-Local terminals. As of yet, the data analysis tools in SP8 have been developed using publicly accessible data-bases, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Parkinson's Progression Markers Initiative (PPMI). This is due to problems inhibiting data sharing between hospitals and researchers within SP8. While the hope had been that clinicians in hospitals would readily share their data with the MIP, in reality it became clear that incentives for clinicians to share medical data are not high enough to perform the labour necessary to make the data findable, accessible, interoperable and reusable (FAIR). The machine learning tasks currently performed include: semi-supervised classification, re-description mining, option predictive clustering, feature ranking, and equation discovery. All these methods require standardised and well-distributed data-sets to explore associations between features in the data, and the results of various clinical and imaging tests.

For example, research in the Department of Knowledge Technologies at Jožef Stefan Institute in Ljubljana, Slovenia found associations between gene variants and cognitive impairments in the ADNI databases, as well as a high correlation of PAPP-A (pregnancy associated plasma protein) and Alzheimer's Disease, which had not been reported elsewhere and was identified as an area for further study (Mihelčić et. al. 2017). This research has served as a 'proof of principle' within the HBP. However, researchers are aware that to move beyond 'proof of principle' more clinical data from hospitals would be required. Towards this end, the Neurodegenerative Virtual Brain (TVD-NDD) joined the HBP as a partnering project (WP8.7) in 2018 and is run by The Virtual Brain consortium (TVB),[14] which aims at bringing to the MIP additional data processing and integration facilities. Using EEG, MEG and BOLD data, the role of TVD-NDD is to develop multiscale simulations that link these data to whole-brain atlases, as well as neuron and population-level simulations in order to "identify key mechanisms that predict neurodegenerative disease (NDD) progression". This aspect of the HBP's MIP builds on The Virtual Brain platform which was originally developed as a generic modelling platform, later developed for epilepsy modelling, and which is described in more detail below.

12 https://kclpure.kcl.ac.uk/portal/en/publications/foresight-report-on-future-medicine(7a09988a-4596-4505-ac25-76f8e71937a9).html
13 https://sos-ch-dk-2.exo.io/public-website-production/filer_public/42/e2/42e28dca-6d5d-4513-9771-88ab71fc3ce1/data_protection.pdf
14 https://www.thevirtualbrain.org/tvb/zwei/home

3.2 Epilepsy: F-TRACT and The Virtual Epileptic Patient

The Virtual Epileptic Patient (VEP) merges together various different individual patient data with generic brain models, blurring the boundaries between data and model. The constraints, or limitations of the models, are set using a patient's structural data from MRI and connectivity data from Diffusion-weighted MRI (dMRI). These data undergo a series of processing steps using open-access software. The non-personalised models and tools include: 1) a "parcellation template" which divides the brain into different functional regions, and 2) the Epileptor model - a population neuron model - which simulates features of seizure dynamics. The network models are 'fitted' to the functional data obtained from an individual patient's SEEG recordings.

VEP is described by its developers as a 'Personalised Brain Model' (Jirsa et al., 2017). The idea of a personalised brain model is something of an oxymoron because a model by its very definition is generic. The justification for the term here refers to the use of individual patient data as constraints and parameters, and utilising them for fitting and validating the models. A new model is built for each individual patient, although certain features and functions remain generic.

VEP is currently part of a clinical trial across several clinics in France investigating the effectiveness of computational neuroscience tools in determining the placement of SEEG electrodes, in the localization of patients' 'epileptogenic zone' to be removed during surgery, and in simulating surgical procedures and outcomes in terms of compromises to brain function[15]. The suggestion is that this tool can become a routine part of the pre-surgical analysis of epilepsy patients. Despite the promising results of these modelling techniques, some clinicians we interviewed for this report expressed concerns about the introduction of these 'black-boxed' models into clinical procedures. We explore these in Section 4.3.

In addition to the VEP, a new data-rich partnering project (WP8.8) that joined the HBP in 2018 has the objective to provide access through the MIP to a multicentre database of human intracranial responses to direct cortical stimulations (CCEP) performed during stereoelectroencephalography (SEEG) explorations of epileptic patients who are candidates for resective epilepsy surgery. The F-TRACT database[16], constituted through working with 25 hospitals, has as its aim is to work towards the standardisation of the SEEG exploration of epileptic patients and the development of a human atlas of cortico-cortical connections.

3.3 Psychiatric Disorders

With the aim of expanding the scope of the MIP, new partners have joined to pursue research related to psychiatric disorders such as schizophrenia, depression, and addiction. Given that such research requires large quantities of clinical data required to develop and test data analytics tools, a Call of Expressions of Interest was launched in the Spring 2018 to find interested consortiums that already possess such large clinical data sets.

Data-intensive algorithmics has been facing difficulties in the domain of mental health research due to a number of complex factors, but it is increasingly gaining traction (Schumann et al., 2010; Kendler et al., 2018; Karatheodoris Davis, 2019). A key data-rich Partnering Project (WP8.10) resulting from this call is comprised of 4 research centres (3 in Germany, 1 in the UK), and is working with datasets related to schizophrenia, depression, and addiction. The German centres are part of the Psychiatric Imaging Network Germany (PING), which serves to coordinate, standardize and optimize data acquisition and storage of the neuroimaging studies of 9 participating clusters[17]. The UK centre, King's College London, leads IMAGEN, a European research project

15 http://ins.univ-amu.fr/science/epinov-improving-epilepsy-surgery-management-and-prognosis-using-virtual-brain-technology/
16 https://f-tract.eu/
17 http://www.ping.rwth-aachen.de/

"examining how biological, psychological, and environmental factors during adolescence may influence brain development and mental health. Using brain imaging and genetics, the project will help develop prevention strategies and improved therapies for mental health disorders in the future.[18]" Between them, these 4 centres pursue a large diversity of studies and collect clinical datasets, spanning 3 dimensions usually studied independently:

1) Behavioural and cognitive data: Numerous questionnaires and tests are applied to acquire the relevant behavioural measures. These measures are overlapping across different studies. Some of the major tests include: Positive and Negative Syndrome Scale (PANSS), Beck's depression Inventar (BDI), Global Assessment of Functioning (GAF).
2) Structural and activity imaging of the brain
3) Genetic and biological data (collected through targeted blood tests)

Each of the 4 centres have a MIP-Local terminal, and will be using, and testing, different data analytical tools and methods provided by the MIP, such as Sparse Canonical Clustering Analysis (SCCA). Running multivariate analysis across the 3 dimensions in conjunction aims at making new clusters emerge that will highlight patterns not found in any single dimension. Such a cluster constitutes a pathophysiological model of disease that can then be analysed and tested to assess its relevance. Further, the 2 lead centres (King's College London and Aachen University) are cooperating with the MIP on its mission of interconnection and integration through MIP-Federate, by helping develop text mining and design data structures, curation, etc. As this work is at an early stage, we have not explored it in detail in this report, however the more general issues that we raise below are certain to affect the use of conclusions drawn from algorithmic data mining that we discuss below.

4. Clinical Translation and the Challenge of Trust

At the 2018 Summit meeting of the HBP, the chair of the HBP Clinical Advisory Board urged researchers to remember that when developing even the most wonderful of medical tools, it is mandatory, always, to develop the human side of these tools. The human side of data-intensive algorithmics for brain disorders comprises a triangle of trust relationships: between patients and researchers, between patients and clinicians, and between clinicians and researchers. The Foresight Lab has explored the former during the Ramp-Up Phase of the HBP, in its Foresight Report on Future Medicine[19] and in its contribution to the HBP Ethics and Society Opinion on Data Protection and Privacy.[20] One of our central recommendations was that patients, patient representatives and patient bodies should be involved at every stage of the research, and actively engaged in the process – patients should be regarded as collaborators in the co-production of results, not merely as data points. Here we explore similar issues concerning questions of trust between patients and clinicians and between clinicians and researchers as they may impact the use of diagnostic technologies based on machine learning techniques of analysis of large quantities of patient data gathered in hospitals, in research projects and elsewhere.

Such data-driven medical research necessitates collecting and accessing increasingly detailed data that stand as proxy for increasingly large cohorts of patients. A consequence is the frequent lack of direct relations between patients and researchers, with data collection and access largely intermediated by hospitals and other healthcare institutions. Gathering large quantities of data anew from large numbers of patients is a costly and time-consuming undertaking, and hence there is a strong incentive to re-use existing data from previous research subjects or from patients'

18 https://imagen-europe.com/
19 https://kclpure.kcl.ac.uk/portal/en/publications/foresight-report-on-future-medicine(7a09988a-4596-4505-ac25-76f8e71937a9).html
20 https://sos-ch-dk-2.exo.io/public-website-production/filer_public/42/e2/42e28dca-6d5d-4513-9771-88ab71fc3ce1/data_protection.pdf

medical records in hospitals and clinics. Data anonymisation is often used, as it is in the Human Brain Project, in order to avoid having to seek new consent from patients or their relatives in order to re-use these data for a new purpose. However it can lead to a double bind: excess anonymisation can void the data of valuable research information and yield only noise, but insufficient anonymisation has the potential for re-identification. More significantly for our current discussion, from the patients and public perspective, the lack of direct human relation with researchers, the lack of control and oversight on how and when patients' data may be used and re-used, and the opacity of certain research initiatives, can lead to lack of confidence in the veracity and clinical utility of the resulting diagnostic technologies, as well as to perceived breaches of trust.

4.1 Trust in patient-clinician relationships

The 2012 Health Survey by the Pew Research Centre and California Healthcare Foundation of 3014 adults living in the United States found that 70% of study respondents "got information, care or support from a doctor or other health care professional, 60% got information or support from friends and family and 24% got information from others who have the same health condition - also known as peer-to-peer support" (Fox and Duggan, 2013). It also found that "Clinicians are the central resource for information or support during serious episodes". This not only attests to the high degree of trust between patients and clinicians when it comes to health decisions but to the fact that "the care and conversation take place mostly offline", which adds a physical face-to-face dimension to the ways in which trust between clinicians and patients are nurtured and maintained (Fox and Duggan, 2013).

As to "Why do individuals trust their doctors the most?" a similarly large survey by PricewaterhouseCoopers (2012) found that "Human relationships" were instrumental in trust-building and trust-maintaining processes. In the adoption of clinical technologies, this is key as Kathryn Armstrong, senior producer of web communications at Lehigh Valley Health Network, USA noted:

> "You want to trust and connect with the people providing you the care. It's easier to trust a person than an organization ... while medical technology companies will disseminate information via their product sites, very few have actually engaged with patients due to regulatory concerns. Healthcare providers have the ability to form human relationships and connections with their patients, which ultimately leads to increased trust" (ibid).

In 2012, PricewaterhouseCoopers' (PwC) conclusion was that "as building these relationships [of trust with patients] becomes increasingly important to establishing trust and credibility with consumers, healthcare companies will need to reconsider their approach to these relationships." According to Ferdinand Velasco, Texas Health's chief medical information officer (op. cit 27),

> There is a lot of patient data—clinical and soon genomics as well. But what is really happening in our patients' lives is missing to us and their record— what's happening in their lives is happening in the social space. ...If we understand the life factors that impact when and who they select for care and what challenges they face after receiving care, there is a lot of potential for merging analytics with the clinical side and improving care.

The lack of access to various facets of human suffering that lies underneath the bare data is particularly acute for organisations and institutions engaged in translational research where researchers traditionally have no direct interaction with patients except during clinical trials. Yet, despite this lack of engagement with patients and clinicians in the upstream development stages, the expectation of scientific innovation is that once the technology is developed it will be readily adopted by both clinicians and patients at the researchers' word of its power to improve their lives (Datta, 2018). Many scholars suggest that public engagement 'downstream' – that is to say after a technology is developed - "'is largely ineffective in rebuilding public trust" (Wynne, 2006, p. 217). Technology driven research has widened this divide between the context of research and the clinical situation (Epstein, 1996; Smith et al., 2017), creating a widening "chasm in the

understanding of end-users between provider's imagination of a future user and users' lived experiences" (Datta, 2018, p. 354).

Some technology providers have attempted to go beyond an imagined picture of the end user (Hyysalo & Johnson, 2015) to incorporate evidence on user experiences. Thus the PwC report suggests that technology designers should:

> Invest in monitoring targeted conversations and integrating data into product decisions across research and development, drug safety, product complaints, sales and marketing, market research, and other business operations. Treat social media as another source of business intelligence that can provide insights at the aggregate level (e.g. how is your product working, is there an untapped market, and what improvements can be made?). (PwC, 2012, p. 32).

While, in one sense, this advice recognises that translation will only happen if a technology meets user's needs and is designed in a way that they trust, it is unfortunate if the only considerations that lead to such user engagement are the dynamics of profit maximisation.

4.2 Clinician-driven innovation vs. researcher-driven innovation

Clinician-driven innovation or *medical innovation* is innovation where the "main goal of innovative care is to improve an individual patient's condition" (International Society for Stem Cell Research, 2008, p. 15). However, such individualised knowledge of patient's responses to medical innovation is often considered anecdotal, non-generalisable and thus inconclusive for proving safety or efficacy. This is distinguished from clinical research or researcher-driven *scientific innovation* where the "aim is to produce generalisable knowledge" based on the conclusiveness of randomised clinical trial data (ibid). Patient acceptance of innovative medical technologies, while it may be informed by personal experience or the experience of other patients, ultimately relies on the relationship of trust between clinicians and patients (Lindvall & Hyun, 2009).

Patient acceptance of scientific innovation thus reflects in large part the trust that the administering clinician has in the use of an innovative technology in clinical settings. This suggests that researchers need to build trust with clinicians for technology adoption and as numerous studies show, including clinician experiences meaningfully in upstream technology development processes - rather than as part of a box-ticking exercise in post-development processes - is key to meaningful technology adoption and adaption (Wilsdon & Willis, 2004; Wong et al, 2008). For instance, a study of the development of an 'electronic clinical handover system' at the Department of General Internal Medicine (DGIM), Royal Hobart Hospital (RHH) found that:

> Although there are clearly practical difficulties in addressing and responding to the heterogeneous requirements expressed by different users, marginalizing these views ultimately is to the detriment of the systems built. …[meaningful engagement] with clinicians in the development of a sustainable system …by drawing attention to the importance of users [i.e. clinicians] and by outlining the practical experience of dealing with the diversity of requirements and views expressed, within the domain of eHealth [was important] for a user-centred systems approach… (Wong et al, 2008).

In the domain of psychiatry research and practice, Sullivan et al. (2005) argue for a "bottom-up" approach in which services researchers assist frontline clinicians in testing interventions that clinicians themselves have devised …and result in interventions that are more likely to be sustained over time" (see also Board, 2010; Borkovec, 2004).

4.3 Trust and Epistemological Conflict between Clinicians and Researchers

While the HBP Medical Informatics Platform is not currently intended for use in clinical settings, there are work packages within SP8 that involve collaborations between researchers and clinicians. Interviews with members of these work packages have highlighted the importance of close collaborations between clinicians, or "domain experts", and data analysts not just for the sharing of medical data, but more importantly for defining the research questions pursued. Computer scientists in SP8 stated that research questions should be defined by clinicians, and only then can machine learning tools or methods be selected and decisions made about which types of data are needed. The current relationship between clinicians and computational neuroscientists in some work packages of the HBP is that the clinic sends researchers anonymised patient data, researchers perform analyses using computational models, and send back the results of the analyses. Trust emerged in interviews as an important element in the relationship between clinicians and researchers, particularly around data bias, model transparency and different epistemological traditions in neurology, the neurosciences, and computer sciences.

With regards to model transparency, researchers found it concerning that clinicians do not understand the modelling frameworks used - specifically the assumptions of the statistical, machine learning and other data analysis methods. Clinicians, on the other hand, found it concerning that modellers do not understand the biology and physiology of the conditions being explored, which they deemed necessary to develop clinically useful models. Computational modellers suggested that future educational programmes for physicians needs to include training in computational methods in order to adapt to changing clinical contexts where machine learning and other computational tools are likely to become more common-place. While modellers acknowledged that "domain-specific" training would be desirable to collaborate with clinicians, the generalist traditions of computer science require that tools are developed for general-purpose and are then adapted to specific use-cases. While some modellers do develop expertise in specific fields (such as neuroscience, or oncology), it is more common for computer scientists and engineers to move across biological domains of expertise.

Shorter-term technical solutions are being explored to circumvent these problems. For example, some members of SP8 have argued that machine learning tools need to be interpretable to clinicians. This means that the clinician needs to be able to trace back the way an algorithm has come to a certain conclusion, thus rendering the decision-making 'transparent'. Indeed, transparency is one of the key ethical principles in discussions around accountability in AI (Mittelstadt et. al. 2016). An association of researchers in Microsoft, Google, and others have, for example, proposed the principles of "Fairness, Accountability, and Transparency in Machine Learning" to address the "potentially discriminatory impact of machine learning" as well as the "dangers of inadvertently encoding bias into automated decisions" (FATML, 2018). Interviews with computer scientists in the HBP suggested that, to adhere to the principle of transparency, supervised and semi-supervised learning algorithms were preferable to unsupervised learning algorithms. This is because the supervised and semi-supervised classification algorithms being developed in the HBP can be represented as decision-trees, which are more interpretable to collaborating clinicians than the 'black box' of unsupervised learning.

However, even if algorithms are less opaque, there remain other epistemological obstacles – that is to say, clashes of epistemologies - to collaborations between clinicians and researchers. Clinicians interviewed for this report talked about the importance of clinician-patient interaction for diagnosis. For example, in the diagnosis and treatment of epilepsy, clinicians carry the patients through from diagnosis to pre-surgical screening, surgery, and post-surgical rehabilitation. This is seen by some as an already personalised treatment since each patient is unique in terms of the symptoms exhibited and surgical treatment needed. This 'holistic' treatment of the individual - common in neurology - is seen by computational neuroscientists as "subjective" and "biased", elements that need to be removed or reduced in clinical settings which were described as "low

validity environments" by computational neuroscientists. Psychologist Daniel Kahneman's work on decision-making was used to support this idea: "to maximize predictive accuracy… decisions should be left to algorithms in low validity environments" (2013).

The 'tacit knowledge' of the clinicians is crucial, for example concerning the decision on where to place SEEG electrodes for an epileptic patient, and the consequent analysis of the various data sets from the patients to decide which part of the brain to remove during surgery. Researchers consider such tacit or experiential knowledge as an implicit model because it involves a series of steps that result in a decision, or solution. However, they expressed concern that these steps are not made explicit. On the other hand, clinicians defended this 'intuition' as the result of years of diagnostic experience and surgical interventions that become internalised in their judgement processes. Clinicians insist that just because these were not fully articulated, or perhaps even capable of full articulation, does not mean they were not valid.

The epistemological conflict between the 'subjective' knowledge of the physicians and the so-called 'objective' knowledge of the data analysis finds its way into discussions around data bias. Data scientists need what they call "good data distribution", standardised data, and comprehensive meta-data. An HBP computer scientist said:

> There are different cultures of how clinicians diagnose and treat – different clinicians with different expertise and knowledge can label patients differently. And many scores – like the Montreal Cognitive Assessment (MoCA) – are arbitrary, on a scale of one to five or something. This kind of noise can be compensated for if you have large amounts of physicians involved who can make sure the data is distributed well and can check the quality of measurements, missing data.

Interpretable or not, clinicians interviewed for the report suggested that if the computational models recommended a different diagnosis or course of treatment than that recommended by clinicians, the clinician's - rather than the researchers' analyses - would be pursued. However, since the use of machine learning tools in clinical settings is still undergoing clinical trials, it remains to be seen what the clinical implications of the use of these tools will be, and how clinicians will respond to the introduction of these new techniques. Furthermore, because of these epistemological differences amongst clinicians themselves, and between clinicians and researchers, HBP scientists have said that developing their models for clinical use will require the involvement of industry partners to translate or develop these models into marketable and user-friendly software – of course, after undergoing clinical trials for EMA or FDA approval. The next section discusses the difficulties HBP is likely to face in commercialising the data analytics tools currently in development.

5. Challenges to commercialisation

Beyond the translational complexities already discussed, there are many challenges in the final step, that is to say how these emerging computational diagnostics models developed within the HBP, once embodied in 'software' will reach the bedside. Within the HBP, it is likely that this final step - in fact a series of complex steps - will be undertaken by commercial companies skilled in the regulation and marketing of medical devices. The pathway to market involves complying with complex regulatory standards to obtain market authorisation.[21] These require the investment of

21 European Medical Devices Directive (MDD 93/42/EEC) till April 2020 and the EU Medical Device Regulation (MDR 2017/745) thereafter. Annex VIII, Chapter III, Rule 11 of the EU MDR classifies software as a medical device (SaMD) as "…software used for diagnosis or therapeutic purposes" under Class IIa medical device "UNLESS: Failure could cause a serious deterioration of health or surgical intervention – then it is Class IIb (e.g., software driving monitoring of a respiratory or circulatory system)". Other SaMD relevant sections are Annex 1 (General Safety and Performance Requirements), Chapter 1, Sections 15 (page L 117/99), 17 (for Programmable Electrical Medical System; page L 117/100-101), 22 (use by lay persons). Relevant FDA regulations are the US Quality System Regulation (21 CFR Part 820; specifically 820.30 on design control), an updated SaMD guidance in 2017 (https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-

considerable capital in the generation of clinical evidence, evaluation, verification and validation processes.[22] This work is likely to be undertaken by established medical device corporations, as the venture capital base in the EU is not only notoriously sparse in early-stage investment unlike in the US (Byword & Timmons, 1992; Faria & Barbosa, 2014), but almost non-existent for medical applications which are perceived as highly risky based on uncertain regulatory environments (Ackerly et al, 2009) involving lengthy pathways to market (Fleming, 2015). The EC's H2020 undertook initiatives to boost early-stage financing by "protect[ing] up to €320.14million to help innovative firms to gain access to various types of risk financing".

According to principal investigators of medical devices research projects based in leading Europe research institutions, investors are wary of the stricter regulatory environment within the EU compared to the US, and hence consider this an area that is highly risky for investment. This is consistent with studies of VC financing trends in medical applications in Europe (Ackerly et al, 2009; Fleming, 2015). Thus large corporations in the medical devices industry present the most likely option for Europe-based innovators to unlock both capital and regulatory investment needed to commercialise, although this private sector investment tends to come at the mature stages when there is more certainty of a product's 'commercialisability'. For instance, in the EPINOV project funded by INSERM (France's national funding for research in health and medicine), researchers from Aix Marseille Université collaborate with Hôpitaux de Marseille (and other hospitals and clinics across France). However, the software development at the mature stages based on the Virtual Epileptic Patient models currently under clinical testing, will be done by the private actor Dassault Systèmes. This trend of mature-stage buy-out type of extractive innovation financing - i.e. extracting private value from publicly funded innovation - is consistent with financing trends in the European medical device innovations market (Lehoux et al, 2014 in Lehoux et al, 2016).

The capacity to engage with and successfully establish partnerships with large private medical devices actors will determine which innovation reaches the market, and thus it is not surprising that many of the developments that we have discussed in this report already involve collaboration with industry. Since most large public research institution-based innovators in Europe and elsewhere (are mostly contractually obligated to) rely on their home institution's 'technology transfer centres' (TTCs) (eg. http://www.mttc.org/) to facilitate the entire 'partnering' process particularly in formalising institutional and researchers shared rights to intellectual property generated, the HBP - as an external actor to this process - can play an essential match-making role in pairing eligible private actors to invest in the research it funds. Notably, TTCs backed by large universities often partner with private legal expertise specialised in navigating the regulatory-approval seeking process that are costly but typically offered gratis to affiliated researchers.

Issues of intellectual property add further complexities for commercialisation of HBP-funded research. Given that the HBP-funded component of most research team's innovation is required to be open-access or publicly accessible e.g. as GNU General Public License (GNU-GPL) software, this part of the research will have no intellectual property or 'exchangeable value' to offer private actors in return for investment. Only the non-HBP funded non-publicly accessible component of each research team's innovation - or what is known as closed-source or proprietary software - will thus be of interest to commercia actors. This separation of research from its output in software development driven by research funder's open-access requirement is not new and increasingly the norm, as, for example, in ONIX[23] - "a closed source yet globally distributed SDN [software-defined networking] controller" is an early example (Berde et al, 2014). This is to some extent recognised

gen/documents/document/ucm524904.pdf); 'draft non-binding recommendations' in 2017 to deregulate a new class of SaMD as low-risk "Clinical and Patient Decision Support Software" although medical imaging software continue to be regulated (page 6 in https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm587819.pdf).

22 IMDRF/SaMD WG/N41FINAL:2017. Software as a Medical Device (SaMD): Clinical Evaluation Authoring Group: International Medical Device Regulators Forum. (Updating IMDRF/SaMD WG/N12FINAL:2014).

23 https://www.usenix.org/legacy/event/osdi10/tech/full_papers/Koponen.pdf

in developments such as The Pan-European Venture Capital Fund(s)-of-Funds programme.[24] Of course, the question of whether closed-source policies should be adopted to encourage the private extraction of publicly-funded research, remain a matter of debate. As does the more problematic question of the extent to which profit-based interests of extractive innovation financing, rather than social and medical need, shape which health technologies reach the market.

6. Conclusion: Challenges along the pathways to the clinic

6.1 Transparency is crucial

Transparency and open science are core values of responsibility in scientific research and innovation. These issues have become prominent in debates over algorithmics, with the demand for explainability and accountability of inscrutable systems getting stronger (Vayena et al., 2018). This demand is gaining traction as it is increasingly believed to be a mandatory step towards safety and trustworthiness of AI and machine learning systems, and it is seen as integral to ethically-aligned design principles. Unsupervised learning algorithms, which aim is to discover inherent structures in data without using pre-existing categories, are under special scrutiny for being notoriously inscrutable even to their designers. In this context, the idea of using such unsupervised machine learning to discover 'brain signatures' that could bring about a complete revision of the classification of mental disorders appears more problematic than ever.

6.2 Involvement of clinicians upstream is crucial

Ongoing collaborations between clinicians and modellers in the HBP, especially in the field of epilepsy, provide an opportunity to reflect on the obstacles posed by innovations in data analytics and their clinical applications. The involvement of clinicians in the formulation of research problems and in the start of the design of research projects before they reach clinical trials has been effective in fostering these collaborations.

However, clinicians do not have the required training to critically analyse the results of these new data analysis tools, and because of a lack of understanding of these models and the ways algorithms reach decisions, have expressed a distrust of the results. While unsupervised learning has been successful in diagnosis, especially in relation to the analysis of medical images (Google DeepMind), these are far less interpretable than other machine learning methods. And even if machine learning tools are more interpretable, clinicians interviewed for the report have stated that if the data analysis offers different conclusions than their medical opinion, they will likely trust their own opinion. Research shows that 'low-level' clinical tests are not always trusted across institutions, and some clinics only trust data and tests from their own labs. Thus involvement of clinicians in early stage research is a crucial step towards gaining clinical trust necessary for translation to the clinic more generally.

There is a need to acknowledge the tacit knowledge of clinical reasoning, which has made it difficult in the past to incorporate algorithmic tools. While there is a trend towards the quantification of healthcare in recent years, and the view that the 'subjectivity' of clinicians' assessments are 'biased', it is important to see the value of the experiential and tacit knowledge of clinicians that is so foundational to the relationships of trust between patients and physicians. Technologies that are perceived to add value to clinical reasoning rather than competing with it

24 http://www.eif.org/what_we_do/equity/paneuropean_venture_capital_fund_of_funds/index.html

(e.g. the new generation of AI-integrated or 'smart' computation diagnostics tools) are more likely to win clinician's trust.

7. Acknowledgements

We would like to convey our special thanks to Dr Edison Bicudo (University of Sussex, UK), Professor Alex Faulkner (University of Sussex, UK), and to all who accepted to participate in this research. None of them bear responsibility for our interpretation or use of their comments.

The authors acknowledge support from the European Union's Horizon 2020 Research and Innovation Programme funding for the Human Brain Project (Special Grant Agreement 2 number 785907).

8. References

Ackerly, D. C., Valverde, A. M., Diener, L. W., Dossary, K. L., & Schulman, K. A. (2009). Fueling innovation in medical devices (and beyond): venture capital in health care. *Health Affairs*, *28*(1), w68-w75.

Berde, P., Gerola, M., Hart, J., Higuchi, Y., Kobayashi, M., Koide, T., ... & Parulkar, G. (2014, August). ONOS: towards an open, distributed SDN OS. In *Proceedings of the third workshop on Hot topics in software defined networking* (pp. 1-6). ACM.

Board, C. C. (2010). Collaborating with Consumers, Providers, Systems, and Communities to Enhance Child Mental Health Services Research. *Children's Mental Health Research: The Power of Partnerships*.

Borkovec, T. D. (2004). Research in training clinics and practice research networks: A route to the integration of science and practice. *Clinical Psychology: Science Practice, 11*, 211–215.

Bygrave, W. D., & Timmons, J. (1992). Venture capital at the crossroads.

Collins, H. (2010). Tacit and Explicit Knowledge. University of Chicago Press.

Datta, S. (2018). Emerging dynamics of evidence and trust in online user-to-user engagement: the case of 'unproven' stem cell therapies. *Critical Public Health 28(3),* 352-362.

De Fauw, J., et al. (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine* 24(9): p. 1342-1350.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*. 31(4): p. 198-211.

EMERJ (2019) Artificial Intelligence in Healthcare – A Comprehensive Overview. https://emerj.com/ai-sector-overviews/artificial-intelligence-in-healthcare-a-comprehensive-overview/

Epstein, S. (1996). *Impure science: AIDS, activism, and the politics of knowledge* (Vol. 7). Berkeley: University of California Press.

Faria, A. P., & Barbosa, N. (2014). Does venture capital really foster innovation?. *Economics Letters 122*(2), 129-131.

Feinstein, AR. (1973). An analysis of diagnostic reasoning. I. The domains and disorders of clinical macrobiology. *Yale J Biol Med*. 46(3):212-32.

Fleming, J. J. (2015). The decline of venture capital investment in early-stage life sciences poses a challenge to continued innovation. *Health Affairs 34*(2), 271-276.

Fox, S., & Duggan, M. (2013). Health online 2013. *Pew Research*. Retrieved from http://www.pewinternet.org/2013/01/15/health-online-2013/

Future Advocacy (2018). Ethical, Social, And Political Challenges Of Artificial Intelligence In Health. *A report with the Wellcome Trust*, April. https://wellcome.ac.uk/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf

Gheiratmand, M., I. Rish, G. A. Cecchi, M. R. G. Brown, R. Greiner, P. I. Polosecki, P. Bashivan, A. J. Greenshaw, R. Ramasubbu & S. M. Dursun (2017). Learning stable and predictive network-based patterns of schizophrenia and its clinical symptoms. *Schizophrenia* 3(22).

Gusnard, D. A., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat. Neurosci. Rev.* 2, 685–694.

Horizon 2020. Access to risk finance. Work Programme 2014–2015. European Commission Decision C (2013)8631 of 10 December 2013. http://ec.europa.eu/programmes/horizon2020/en/news/horizon-2020- brief-eu-framework-programme-research-innovation

Hyysalo, S., & Johnson, M. (2015). The user as relational entity. *Information Technology & People 28*(1), 72–89.

ISSCR. (2008). Guidelines for the clinical translation of stem cells. *International Sociey for Stem Cell Research*. Retrieved from http://www.isscr.org/docs/default-source/all-isscr-guidelines/clin-trans-guidelines/isscrglclinicaltrans.pdf?sfvrsn=6

Jirsa, Viktor, T. Proix, D. Perdikis, M.M. Woodman, H. Wang, J. Gonzalez-Martinez, C. Bernard, C. Bénar, M. Guye, P. Chauvel, F. Bartolomei (2017). The Virtual Epileptic Patient: Individualized whole-brain models of epilepsy spread. *NeuroImage* 145(B): 377-388.

Kahneman, D. (2013). Thinking, Fast and Slow. Farrar, Straus and Giroux.

Kapur, S., Phillips, A.G., Insel, T.R., 2012. Why has it taken so long for biological psychiatry to develop clinical tests, and what to do about it. *Molecular Psychiatry 17*, 1174–1179.

Karatheodoris, D. L. (2019). Common knowledge: shared genetics in psychiatry. *Nature Neuroscience* 22, 331–332.

Kendler, K., Gardner, C., Neale, M., Aggen, S., Heath, A., Colodro-Conde, L., . . . Gillespie, N. (2018). Shared and specific genetic risk factors for lifetime major depression, depressive symptoms and neuroticism in three population-based twin samples. *Psychological Medicine,* 1-9.

Lehoux, P., Daudelin, G., Williams-Jones, B., Denis, J. L., & Longo, C. (2014). How do business model and health technology design influence each other? Insights from a longitudinal case study of three academic spin-offs. *Research Policy 43*(6), 1025-1038.

Lehoux, P., Miller, F. A., & Daudelin, G. (2016). How does venture capital operate in medical innovation?. *BMJ innovations*, *2*(3), 111-117.

Lindvall, O., & Hyun, I. (2009). Medical Innovation Versus Stem Cell Tourism. *Science*, *324*(5935), 1664–1665.

Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453(7197): p. 869-878.

Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical Informatics Insights*.

Mahfoud, T. (2014) Extending the mind: a review of ethnographies of neuroscience practice. Frontiers in Human Neuroscience Vol 8(359).

Mattingly, C. and M. H. Fleming (1994). *Clinical Reasoning: Forms of Inquiry in a Therapeutic Practice*. Philadelphia: F.A. David Company.

Mihelčić M,S Ï imić G, Babić Leko M, Lavrač N, DzÏeroski S,S Ï muc T, et al. (2017) Using redescription mining to relate clinical and biological characteristics of cognitively impaired and Alzheimer's disease patients. PLoS ONE 12(10).

Mittelstadt, B. D., P. Allo, M. Taddeo, and S. Wachter (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society* 3 (2).

Murray, R. M. (2017). Mistakes I Have Made in My Research Career, *Schizophrenia Bulletin*, Volume 43, Issue 2: 253–256.

Nature Biotechnology (2012). What happened to personalized medicine? Nature Biotechnology 30 (1).

Nuffield Council on Bioethics (2018). Artificial intelligence (AI) in healthcare and research. Bioethics Briefing Note, May: http://nuffieldbioethics.org/wp-content/uploads/Artificial-Intelligence-AI-in-healthcare-and-research.pdf

PricewaterhouseCoopers. (2012). Social media 'likes' healthcare. *PWC Health Research Institute*. Retrieved from https://www.pwc.com/us/en/health- industries/health-research-institute/publications/pdf/health-care-social-media-report.pdf

Raichle, M. E. (2011). The restless brain. *Brain Connect.* 1, 3–12.

ROAMER—A Roadmap for Mental Health Research in Europe, (accessed 1.05.2013). Available from: ⟨http://www.roa mer-mh.org⟩.

Roepstorff, A. (2007). Navigating the brainscape: When knowing becomes seeing. In Grasseni, C. (ed) Skilled Vision, Berghahn, Oxford: 191-206.

Rose, N. (2013). What is Diagnosis For? Talk given at the Institute of Psychiatry Conference on *DSM-5 and the Future of Diagnosis*. 4 June 2013.

Rose, N. and Gainty, C. (2019) Neurovisions, *Advances in Clinical Neuroscience and Rehabilitation*, 18 (2), 17-18

Saunders, B. (2008). CT Suite: The Work of Diagnosis in the Age of Noninvasive Cutting. Duke University Press.

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., et al. (2010). The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Molecular Psychiatry 15*, 1128–1139.

Smith, R.C., Bossen, C., & Kanstrup, A. M. (2017). Participatory design in an era of participation. *CoDesign, 13*(2), 65–69.

Stone, P., R. Brooks, E. Brynjolfsson, R. Calo, O. Etzioni, G. Hager, J. Hirschberg, S. Kalyanakrishnan, E. Kamar, S. Kraus, K. Leyton-Brown, D. Parkes, W. Press, A. L. Saxenian, J. Shah, M. Tambe, and A. Teller (2016). Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA. http://ai100.stanford.edu/2016-report.

Sullivan, G., Duan, N., Mukherjee, S., Kirchner, J., Perry, D., & Henderson, K. (2005). The role of services researchers in facilitating intervention research. *Psychiatric Services, 56*, 537–542.

Vayena, E., A. Blasimme, and I.G. Cohen (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine* 15 (11).

Wilsdon, J., & Willis, R. (2004). See-through science: Why public engagement needs to move upstream. Demos.

Wong, M. C., Turner, P., & Yee, K. C. (2008). Involving clinicians in the development of an electronic clinical handover system-thinking systems not just technology. *Studies in health technology and informatics*, *136*, 490.

Wynne, B. (2006). Public engagement as a means of restoring public trust in science – Hitting the notes, but missing the music? *Community Genetics, 9*(3), 211–220.