Think Tank

ROUNDTABLE: **UK-Ireland**

The use of existing Big Data to improve healthcare

Proceedings and results from the EIT Health Roundtable Meeting

20 July 2018 at Keble College, University of Oxford, UK

Introduction

Using Big Data for healthcare in the UK and Ireland

This report contains the proceedings and conclusions of a EIT Health Think Tank Roundtable Discussion about the use of Big Data to improve healthcare in the UK and Ireland. At the 20 July 2018 Roundtable in Oxford, UK, leading experts from the UK and Ireland presented best practices and engaged in discussions that have generated many actionable recommendations.

The UK-Ireland Roundtable was one of several such meetings organised around Europe by the EIT Health Think Tank to obtain regional perspectives that further a discussion on pressing healthcare issues. Conclusions and recommendations stemming from these discussions are used to ensure that EIT Health's strategy and activities focus on what matters most. The Think Tank also guides EIT Health Public Affairs efforts aimed at contributing to health policy and healthcare system improvements across Europe.

The EIT Health Think Tank is a forum of experts and thought leaders cooperating to shape the future of healthcare in Europe. The Think Tank brings together EIT Health Partners with other leading healthcare stakeholders to agree on means for ensuring that innovation reaches the citizens and patients who need it most. Through central and local exchanges, Think Tank members seek to identify healthcare needs and potential solutions to those needs.

By identifying key pressure points and catalysing discussion, the Think Tank drives need-focused innovation within the EIT Health community and beyond, to better meet the needs of European citizens.

The following are summaries of the proceedings of the EIT Health Think Tank UK-Ireland Roundtable Discussion, held 20 July 2018 at **Keble College, University of Oxford, UK.** The meeting involved three sessions. The first two sessions consisted of presentations followed by discussions. Key points raised during the presentations and discussions are highlighted. The third session involved identifying recommendations based on these discussions. These recommendations are summarised in two pages at the end of this document.

Contents

Participants

Chairman	
Paul Timmers	Independent Consultant and Visiting Research Fellow in Cybersecurity Policy and Digital Transformation, Oxford University, UK
Speakers	
Professor Martin Landray	Research Director, HDR UK Oxford; Deputy Director, Big Data Institute, Oxford, UK
Professor Dipak Kalra	President of the European Institute for Innovation through Health Data; Professor of Health Informatics, University College London, UK
Advisors	
Dr David Atkins	Chief Executive Officer, Congenica, Cambridge
Dr Ian Barrett	Associate Director, AstraZeneca, Cambridge
Professor Anne-Marie Brady	Head of School of Nursing & Midwifery, Trinity College, Dublin
Dr Craig Buckley	Head, Product Innovation, Siemens Healthineers GB & Ireland
Dr Dexter Canoy	Epidemiologist, The George Institute for Global Health (UK), University of Oxford
Dr Nick de Pennington	Digital Innovation and Population Health Lead, Oxford University Hospitals NHS Foundation Trust
Mr Tom Denwood	Member of the NHS Digital Executive Team
Mr Sybo Dijkstra	Senior Director, Philips Research UK
Dr John Dinsmore	Health Innovation Lead/Deputy Director of the Trinity Centre for Practice and Healthcare Innovation, Dublin
Dr Lydia Drumright	University of Cambridge
Dr Rebecca Ghosh	Senior Researcher at the Clinical Practice Research Datalink
Dr Darren Lunn	Real World Data Manager at the Clinical Practice Research Datalink
Professor Gil McVean	Professor of Statistical Genetics at the University of Oxford and Director of Oxford's Big Data Institute
Professor Melinda Mills	Nuffield Professor of Sociology, University of Oxford
Mr Chris Molloy	Chief Executive Officer of the Medicines Discovery Catapult
Professor Alison Noble	Technikos Professor of Biomedical Engineering, University of Oxford
Mr Jan Nygaard Jensen	Deputy Head Novo Nordisk Research Centre Oxford; Scientific Director, Head of Bioinformatics & Discovery Technologies
Professor Kazem Rahimi	Deputy Director, The George Institute for Global Health
Mr Bleddyn Rees	International Healthcare Lawyer, Osborne Clarke
Professor Jens Rittscher	Professor of Engineering Science, University of Oxford
Dr Mark Sheehan	Leader, Oxford Biomedical Research Centre (BRC) Ethics Group; Oxford BRC Ethics Fellow at the Ethox Centre, Nuffield Department of Population Health, University of Oxford
Dr Jan Wolber	Head of Imaging Technology, GE Healthcare
Dr Tingting Zhu	Independent Stipendiary Research Fellow, University of Oxford
Observers	
Jan-Philipp Beck	Chief Executive Officer, EIT Health
Leslie Harris	Managing Director, EIT Health UK–Ireland
Roberta Giammaria	Public Affairs and Communications Lead, EIT Health
Mayra Marin	Executive Officer, EIT Health
Murray Gardner	Industry and Research Partnership Manager, University of Oxford
Martin Kerr	Industry and Research Partnership Manager, University of Oxford
Karen Wolstencroft	Medical Writer and Communications Consultant to EIT Health
Henrik Cyrén	Managing Director, EIT Health Scandinavia
Merike Leego	Innovation Manager, EIT Health Scandinavia
Marco Pugliese	Managing Director, EIT Health Spain

Agenda

Roundtable Meeting Chairman: Paul Timmers

09:00–09:30	**Welcome and introductions**	
	Introduction of participants; agenda for the day. Background to the EIT Think Tank on Big Data; objective for the day.	Paul Timmers Jan-Philipp Beck
9:30–11:00	**Session 1: Research: data availability and integration**	
	Big Data in health research **Focus on:** Challenges regarding research data availability, access and integration, e.g. data from multiple sources; data quality including bias; integration of databases and interoperability; linking clinical & industrial data R&I	Speaker: Professor Martin Landray
11:00–11:30	*Coffee break & networking*	All
11:30–12:30	**Session 2: R&I collaboration and framework conditions**	
	R&I collaboration and framework conditions **Focus on:** R&I partnership models for the UK; collaboration between academics – industry – authorities; codes of conducts for R&I partners as well as for partnerships; current and emerging legal and policy frameworks and their impact on R&I collaboration and R&I opportunities. This will be complemented by addressing capability building and training as a key enabler and area of action. Collaboration and framework conditions address issues amongst others of data access & use, and data flows and liability; outcomes measurement; HTA, device, and clinical trials regulations; patient recruitment; scaling-up and interconnection of databases, e.g. European Electronic Health Records.	Speaker: Professor Dipak Kalra
12:30–13:30	*Lunch*	All
13:30–15:00	**Session 2: Discussion continued**	
15:00–15:30	*Coffee break & networking*	All
15:30–17:15	**Session 3: Actions and recommendations**	
	▪ Communicating about big data and health ▪ Concrete research & innovation projects to pursue in EIT Health UK/Ireland ▪ Research & Innovation in near-future EIT Health work-programme / Innovation Platform ▪ Recommendations on collaboration and framework conditions to participants, UK authorities, European/international level	Moderator: Paul Timmers

Introductory session

Background to the EIT Think Tank on Big Data in Healthcare

PRESENTER:

Jan-Philipp Beck, CEO of EIT Health

Jan-Philipp Beck, CEO of EIT Health, gave the following overview of EIT Health and its goals, as well as the objectives underlying the Think Tank on Big Data in Healthcare and the desired outcomes of the day's Round Table (RT) Meeting:

About EIT Health

EIT Health's slogan is "Together for healthy lives in Europe". The organisation's goal is to identify unmet clinical and economic needs within the complex healthcare systems around Europe. Based on these identified needs, EIT Health then initiates and facilitates innovation by bringing together leaders in education, business, and research.

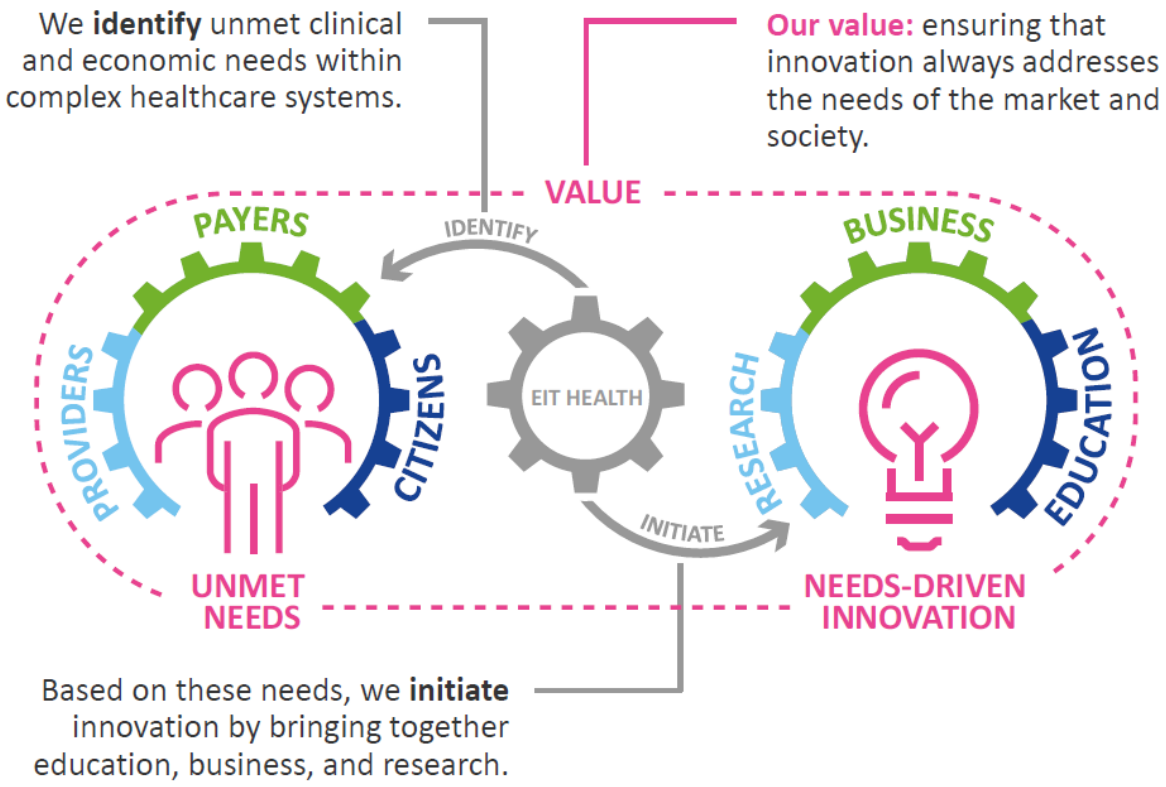
EIT Health is Headquartered in Munich, Germany, and the network comprises six regional offices (Co-Location Centres) in UK–Ireland, Scandinavia, Spain, France, Germany–Switzerland and Belgium–Netherlands – as well as seven further developing innovative regions, the EIT Health InnoStars. The strength of EIT Health lies in the expertise and resources of its unique, diverse and growing network of partners, which currently comprises 140+ leading organisations across all key areas of healthcare (the pharmaceutical industry, medical technology, payers, public and private research institutions, and universities).

EIT Health believes that to bring innovation into the healthcare domain, all these different actors need to be connected, so they bring together world-class thought leaders from partner organisations who can contribute their visions and share their organisations' assets to accelerate innovation.

The value EIT Health brings is in ensuring that this innovation always addresses the needs of the market and of society, to ultimately benefit citizens and patients while supporting economic development in European regions. This focus on "needs-driven innovation" was critical for the day's RT discussions for the UK-Ireland region.

About EIT Health Think Tank

The EIT Health Think Tank is a forum of experts and thought leaders cooperating to shape the future of healthcare in Europe around key topics of relevance and importance in this field. Big Data is one key topic. The Think Tank brings together EIT Health Partners with other leading healthcare stakeholders to agree on means for ensuring that innovation reaches the citizens and patients who need it most. Think Tank members seek to identify healthcare needs and potential solutions to those needs. Think Tank dialogue feeds back into EIT Health activities to ensure resources and innovation are optimised to address healthcare's most pressing issues. It enables a two-way exchange between those who demand health innovation and innovators in the EIT Health partnership who supply it.



The aim is to bridge the gap between the reality facing healthcare providers with ambitious goals at a European level, recognising that while many innovations are currently being developed and tested around Europe, a large proportion are not implemented and never reach citizens. For each key topic, the objective is to reach a consensus on the challenges faced, to understand the barriers at opportunities, and to agree on needs.

The Think Tank spans both European and regional levels: focusing on topics of European relevance, barriers and opportunities are defined at a regional level (via the RTs), and these findings are then consolidated to derive recommendations for initiatives or policy changes.

Expected outcomes of the Round Table meeting

By the end of the day's RT meeting it was hoped that the following would be achieved:

- Actionable ideas for projects (quick wins) that could be implemented regionally and have measurable impact
- Actionable ideas for projects that can feed into the EIT Health Innovation Platform and inform programming
- Input to the policy-making process at regional and/or European level

Three key messages to help ensure these objectives are achieved, are:

1. Let's not only report but also build
2. Do we understand where we have disagreements?
3. Identify actionable outcomes

Summary of discussions and recommendations

KEY POINTS AND RECOMMENDED ACTIONS:
Key points and recommended actions that may be of potential interest to EIT Health for future initiatives are highlighted in bold blue letters. These points inform our Summary of Recommendations.

SESSION 1: **Research: Data Availability and Integration**

PRESENTATION: **Big Data in Health Research**

Professor Martin Landray, Research Director, HDR UK Oxford; Deputy Director, Big Data Institute, Oxford, UK

Professor Martin Landray gave an overview of how Big Data, in particular information from the NHS and other healthcare systems, has the potential to transform medical research.

Big Data in health requires a new strategy and this will involve new collaborations and partnerships across a range of academic disciplines and traditional organisational and sector boundaries. Alongside these requirements, new methodologies and IT infrastructure are needed to facilitate data acquisition, management, analysis and access, as well as advanced computing and controlled management of large, sensitive datasets of healthcare information.

Since this is a whole new discipline, it will require a new cadre of scientists in terms of the type of research fellowships, and the content of doctoral training and continuing professional development. Currently, not many data scientists are skilled in understanding the language of medical science and healthcare. Similarly, few medical researchers can speak the language of data scientists.

New regulatory and governance approaches are also needed: many of the traditional approaches are outdated and a significant amount of work will be required to develop new models. Critically, it will be important to maintain the trust of patients, the public, and the clinical community while implementing these changes, so this needs to be built into any initiatives going forward.

Health Data Research (HDR) UK (https://www.hdruk.ac.uk/) comprises six sites, including one in Oxford, and 22 organisations around the UK. By harnessing health and biomedical data in the UK, HDR UK intends to develop and apply data science approaches to address health research issues. The challenge is to take the data science advances developed by the big academic institutions, apply them at a local level, and ensure that they are accessible. HDR UK's priority areas are: actionable health data analytics, precision medicine, modernising public health, and implementing 21st century clinical trials.

Professor Landray illustrated the value of Big Data with an example of the correlation between ischemic heart disease and systolic blood pressure (BP) when stratified by age. In this case, large numbers of data values give greater clarity and distinction of risk between the different age groups; and this clarity is not possible with a smaller number of values.

Cohort studies are useful to understand the causes and consequences of disease, including genetic components. The challenge is to convert these complex data into meaningful phenotypes and to develop methods to enable phenotyping at scale, in depth and over time. A wide range of data sources are available and need to be considered when developing these approaches. These sources include: routine healthcare records (clinical events), sensors and web-tools to assess behaviour and function (e.g. activity, cognition), imaging and investigations (e.g. MRI, ECG monitoring), and high throughput laboratory data (e.g. proteomics, metabolomics)

Professor Landray used the example of an ongoing UK Biobank (http://www.ukbiobank.ac.uk/) project, which is being undertaken in partnership with the Big Data Institute (BDI) in Oxford (https://www.bdi.ox.ac.uk/) and is looking at a comprehensive assessment of phenotype using a vast array of different data sources, including routine healthcare data. The example shows that, ultimately, linking back to the clinical picture is key. UK Biobank has about 36,000 participants in Scotland, 21,000 in Wales and 446,000 in England. Although the healthcare systems differ between these three countries, collection of routine NHS data provides valuable observational information on deaths, cancer cases (date, stage and grade), admissions to hospitals (dates, diagnoses, and procedures), and primary care data (dates, diagnoses, symptoms, signs, prescriptions, and laboratory tests). This data aggregation enables the identification of trends and patterns in health.

Turning to clinical trials, Professor Landray said there is a crisis in this field, and an urgent need for more efficient approaches. Clinical trials are critical for informed healthcare – treatment not based on randomised, controlled trials (RCTs) could potentially harm patients. In addition, clinical trials provide access to a rich resource of information for possible combination with other datasets and further analysis for different research projects.

The cost of late-phase trials is becoming prohibitively expensive, resulting in fewer, smaller, shorter trials that are often inadequate to address important clinical questions.

Alongside this, there has been a disengagement of the drug development industry, which changed focus from developing treatments for chronic disease and multi-morbidities to more expensive drugs for rare conditions. This has resulted in disillusionment amongst the clinical investigator community as well as reduced entry to clinical trials for patients most in need of new therapies. In addition, over-interpretation of outdated trials regulation ('Good Clinical Practice') has stifled innovative approaches to improve trial quality and efficiency.

To develop efficient RCTs, the following are needed:

- Efficient recruitment using routine clinical data
- Effective assessment of safety and efficacy using digital technology and routine data
- Excellent study quality through software engineering and statistical monitoring

The BDI in Oxford and NHS Digital have been collaborating on a feasibility assessment for a Phase 3 cardiovascular outcome trial that is targeting recruitment and randomisation of 12,000 patients from 90 UK hospitals. The aim is to identify centres where there are eligible patients so that they can be targeted precisely, and efforts focused accordingly, making recruitment efficient and more rapid.

In terms of post-trial follow-up, there is often concern about the quality of data when routine clinical data is collected, but in the context of randomised trials, accuracy of data is not an issue. Accurate data does not necessarily equal reliable results, so it is possible to use real world data in the context of randomisation. As an example, Professor Landray cited the extended follow-up of the Heart Protection Study Collaborative Group trial of the LDL-cholesterol lowering effects of simvastatin, the results of which were published in the Lancet in 2011 showing that benefits persisted for 5 years[1].

Another example of effective use of routine data for clinical research is the 3C trial, which was undertaken to assess the safety and efficacy of different immunosuppressive strategies among 800 kidney transplant patients [2,3]. Various data sources were used including a transplant registry, a disease-specific registry, and NHS hospitalisation data. The study revealed a greater risk of transplant rejection and serious infection with sirolimus versus tacrolimus as well as no significant improvement in kidney function at 18 months.

These studies confirm that real world data can be used in the context of randomisation to answer important clinical questions very cost effectively.

When developing any new technology, it is also important to assess its impact on the healthcare system. An RCT was undertaken by Wilson and colleagues to assess whether a new device – an automated electronic alert for acute kidney injury – could reduce the severity of such injury and improve clinical outcomes in patients in hospital. The device was in fact found to have no impact on patient outcomes despite incurring healthcare personnel time.

Professor Landray concluded by summarising the potential future opportunities of Big Data in healthcare research:

Cohort studies

- Enhancement of existing, mature studies
- Disease-focussed cohorts (e.g. common and moderately rare diseases)
- New large-scale cohort studies enriched with certain patient groups

Efficient trials

- Trials of sufficient size and duration to assess treatment effects reliably
- Commercial, government and charity funded
- Assessments of new and existing treatments, and digital technologies

Extension to other datasets

- Current and future NHS secondary care datasets (e.g. NHS Digital)
- Other data sources (e.g. public health, disease registries, primary care)

References:

1. Heart Protection Study Collaborative Group. Effects on 11-year mortality and morbidity of lowering LDL cholesterol with simvastatin for about 5 years in 20,536 high-risk individuals: a randomised controlled trial. Lancet. 2011 Dec 10;378(9808):2013-2020.
2. 3C Collaborative Group. Alemtuzumab-based induction treatment versus basiliximab-based induction treatment in kidney transplantation (the 3C Study): a randomised trial. Lancet. 2014 Nov 8;384(9955):1684-90.
3. 3C Collaborative Group. Campath, calcineurin inhibitor reduction, and chronic allograft nephropathy (the 3C Study) - results of a randomized controlled clinical trial. Am J Transplant. 2018 Jun;18(6):1424-1434.
4. Wilson FP, et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. Lancet. 2015 May 16;385(9981):1966-74.

DISCUSSION:

Key points raised

FACILITATOR: **Paul Timmers**, Roundtable Chairman, Independent Consultant and Visiting Research Fellow in Cybersecurity Policy and Digital Transformation, Oxford University, UK

Infrastructure and human resources

- We now have the capability to use modelling with Big Data to solve some clinical research problems, like the example presented by Professor Landray, but a significant challenge with Big Data will be the human aspect, in terms of resourcing, training and infrastructure.
- We need to ensure that all relevant skills are represented in Big Data initiatives in clinical research, including data scientists, engineers etc, as well as clinicians and researchers – a multidisciplinary team.
- Ongoing investigation and analysis of data requires a continuous funding stream, which is often difficult to secure.
- Often there is inadequate capacity to interrogate data. We need to build sufficient tools and staffing into new projects.

Analytics and reporting

- May need to think about new analytic methods and different ways of reporting data from clinical trials.

Data access and integration

- Access to data is an ongoing challenge, especially in primary care. In the UK for example, England, Scotland and Wales record data in different ways so it would be helpful to have integration/ interoperability across borders.
- Need more linked central repositories of available data, in both the UK and across Europe.

Data capture and curation

- Standard forms of data capture do not always represent functional measurements well, e.g. the 6-minute walking distance test. Devices such as accelerometers might capture functional data better and more accurately.
- Can we develop technology to assist with effective data curation?
- We need a standardised approach to data input as well as good metadata.

Data sharing

- We need to develop business and incentive models to encourage data sharing between organisations.
- These will help build public trust in collaborations between healthcare and industry.

Predictive value of data

- **Preclinical data: biological assays can provide predictive information – a potential area for funding support as it is cost effective compared to clinical trials.**
- How do we provide information on which data is helpful and predictive and can help inform decision support?

Diversity of data

- We need to embrace differences in sample types and systems, and develop the ability to work with raw data.
- Genetic data: we need to ensure population diversity in what is captured. For example, 70% of UK Biobank data is generated by the UK, US and Iceland, and so is limited to specific populations.

- Imaging data: advances have been made in recent years in standardising imaging file formats, however work still needs to be done in terms of the type of imaging technology used, and how data is collected and interpreted.
- Regarding ultrasound: proprietary technology and formats are a barrier. As with any proprietary technology, there is limited ability for deep analysis of the data.
- Wearable sensors need to be acceptable to patient or they won't be used. Also, we need to generate good data, otherwise processing time is wasted if the results are not usable/valuable.

Privacy and security

- We need to ensure that adequate privacy controls and agreements on data access and sharing are in place while maintaining an efficient, automated data system.
- Examples cited of existing initiatives that might provide useful information:
  - National Data Guardian in UK (https://www.gov.uk/government/organisations/national-data-guardian)
  - The Wellcome Trust: https://wellcome.ac.uk/what-we-do/topics/data-sharing
  - NHS Digital: https://digital.nhs.uk/data-and-information/request-data-access
- Is there a future role for the privacy enhancement technology currently being developed in other sectors? An example is research work being undertaken by the Royal Society on privacy enhancing technologies to address questions of trust around digital systems (https://royalsociety.org/topics-policy/data-and-ai/topic/).

Consent and ethics

- **New dynamic consent models are needed to put the citizen/patient at the centre of process. We need work/research to understand what is acceptable – some may give consent for data to be used in a certain way but not in others.**
- **Is pan-European citizen consent possible/feasible? We would need to ensure this is embraced all of society not just middle-class, well-off and generally healthy people.**
- We need to design a consent process so that data can be used and repurposed in a meaningful way in the long term, not just for one-time use, otherwise valuable data is wasted.
- Is "consent" obsolete and should other norms/social contracts be considered for healthcare data?
- Consent/ethics should be embedded in all Big Data initiatives.

Engagement

- **Research into models and methods for listening and responding to citizens and patients would be beneficial – in particular to understand from them the negative attitudes toward data sharing, and how this can be changed. We need a dialogue to develop a common understanding and to counter the negative perceptions of participation.**
- Citizens need to know exactly how their data is being used and the results that are generated from that use – seeing the value and benefits will help build trust.
- We need to engage and incentivise clinicians to address the benefits of data quality and standardisation of data input.
- Need clarity and simultaneous communication of the benefits of Big Data initiatives to all stakeholders; these should include benefits in four dimensions: for citizens/patients, healthcare systems, academics/researchers, and industry.

Decision support

- **There is a potential for projects/research into areas where clinicians would benefit from access to more basic data, that can influence clinical decisions and drug choices.**
    - **A considerable amount of patient health data exists that could be of value to clinicians when deciding treatment pathways for their own patients ("patients like mine").**
    - **However, the challenge is to access, combine, analyse and present data in such a way that it provides simple but meaningful information on which clinicians can base treatment decisions.**
    - **A good use case/pilot project would be multi-morbidities; most care pathways and clinical trials currently focus on a single comorbidity, so information on multiple morbidities would be of value.**
- We need to research into how, using simple analytics, big datasets could be used to provide easily accessible decision-making tools for clinicians. While EHR are valuable, a processing step is necessary to make the data useful.
- It is important to determine what data is relevant for decision making, and what is not relevant. There is a constant evolution of knowledge in every disease area and it is necessary to publish and communicate both positive and negative results to give a full picture.
- Similarly, we need to expose potential bias and uncertainly to decision makers at the point where it matters.
- We need to ensure robustness and reproducibility of data (from different systems) when supporting decision making. What is the impact of missing data in the context of the question being asked and the clinical decision being made?
- Validation of data from clinical trials: EMEA/FDA requirements are that new drugs need to be validated in large populations. Computer software and modelling techniques are evolving and may change during the trial, presenting a challenge for validation.

Innovations for health improvement

- **How can we use artificial intelligence (AI) and decision support to help patients with multiple conditions set smart health goals?**
- With machine learning (ML) and AI, models can be limited, providing information on correlation but much less on causation. We need good statistical models and more complex analytical models to explore causation in more depth. This would be of particular value in multimorbidity.
- How can we verify that AI and deep learning models are safe before they are deployed?
- What are the robustness and validation standards that the ML/AI community would consider acceptable?

SESSION 2: **R&I collaboration and framework conditions**

PRESENTATION: **R&I collaboration and framework conditions**

Professor Dipak Kalra, President of the European Institute for Innovation through Health Data; Professor of Health Informatics, University College London, UK

Professor Kalra emphasised that there are huge challenges across Europe in delivering next-generation healthcare. He said these challenges fall into three broad categories:

1. Economics:

- Legacy of the financial crisis: high debts and deficits
- Continued increases in public health spending anticipated
- Concerns about how this will be paid for (sustainability of public finances)

2. Population health:

- Ageing and rising levels of chronic disease and comorbidity
- Public health problems and inequalities

3. Health systems:

- Challenge of responding to changing population needs
- Need for structural reforms – e.g. integrated care, eHealth
- Evidence of marked variation in clinical practices and significant levels of "waste"

In the current climate, healthcare systems are struggling to determine the changes they must make to meet these new challenges. Professor Kalra advised that the best way to understand how to transform is to have a better understanding of what you are doing by looking at the data – evidence-based transformation is key.

One important development is that there is now a commonality of purpose and convergence of needs between clinical research and healthcare. Clinical research concerns include enhancing access to real world data, optimising clinical research processes and generating new evidence for precision medicine and value-based care, while major healthcare concerns involve improving the quality and safety of care, supporting patients in self-care and health maintenance, and improving efficiency of care. However, both clinical research and healthcare now have a shared need to improve access to combined health data from multiple sources. It is apparent that it is no longer meaningful to keep these data sources separate and that the two sides should work together to maximise the use of data for the benefit of patients and citizens.

The European Institute for Innovation Through Health Data (i~HD; https://www.i-hd.eu/) is a neutral body that aims to bring key stakeholders in healthcare together to co-create solutions for capturing and sharing better quality health data – and for facilitating its trustworthy use for smarter healthcare and efficient research. The problems of Big Data in healthcare will not be solved by one single stakeholder, so it is important to bring them together and agree an alignment of stakeholder values and interests.

Can we find ways to achieve stakeholder alignment? What are the common value propositions that will bring the resources they each have together for a common purpose – smarter healthcare and more efficient research?

Professor Kalra outlined the common challenges to achieving the effective use of health data for person-centred care, and the re-use of health data for clinical research, which related mainly to privacy/security, quality and interoperability, and value.

Concern for privacy protection, ethics and security create a broad range of challenges, and stakeholders will need to work together to ensure trustworthy access and use of data.

It will also be important to find effective ways to demonstrate value in data transformation, so that data use and handling are not seen as a cost burden. Instead, data must be seen as a valued asset that delivers value. The pharmaceutical industry is now increasingly interested in providing value-based healthcare. Rather than just being the purveyors of medicinal products they want to be a healthcare delivery partner.

What are the current models of this collaborative way of working between healthcare and pharma, and where are they successful?

To assure public trust when reusing Electronic Health Records (EHRs) for research, organisations need to demonstrate that they are compliant with data protection legislation, both at a national level and also across all EU Member States. However, it is inefficient for every project to develop its own policy locally. What is needed is consistent information governance across Europe to give greater confidence to citizens, as well as greater confidence and reduced risk for data providers and users.

i~HD is part of a wider governance landscape that is emerging across Europe with several organisations developing mutually-recognised standards and codes of practice for the secondary use and sharing of medical data for scientific research projects (e.g. BBMRI-ERIC, RD-CONNECT, CORBEL, EMIF). i~HD is now involved in developing processes to implement these codes of practice. There is a need for processes to guide educating and training research and ICT staff, certifying service providers and EHR systems, and promoting consistent practices across Europe – but they can be quite hard to operationalise.

Can we do more to promote European cohesion regarding governance and codes of practice?

Professor Kalra moved on to discuss interoperability standards – not in the context of technical specifications, but also standards that give clinical meaning. There are currently a range of assets: concept presentation (hierarchical terminology systems, ontology etc.), EHR information models (EHR reference models, clinical document registries) and clinical models, (archetypes, templates, datasets, data dictionaries), but nothing is currently bringing these together.

It will be important to work with end-user stakeholders to determine what information will be of value to them, so we can capture the data well and share it easily. It was noted that mobile health is a growing field creating new-generation information silos.

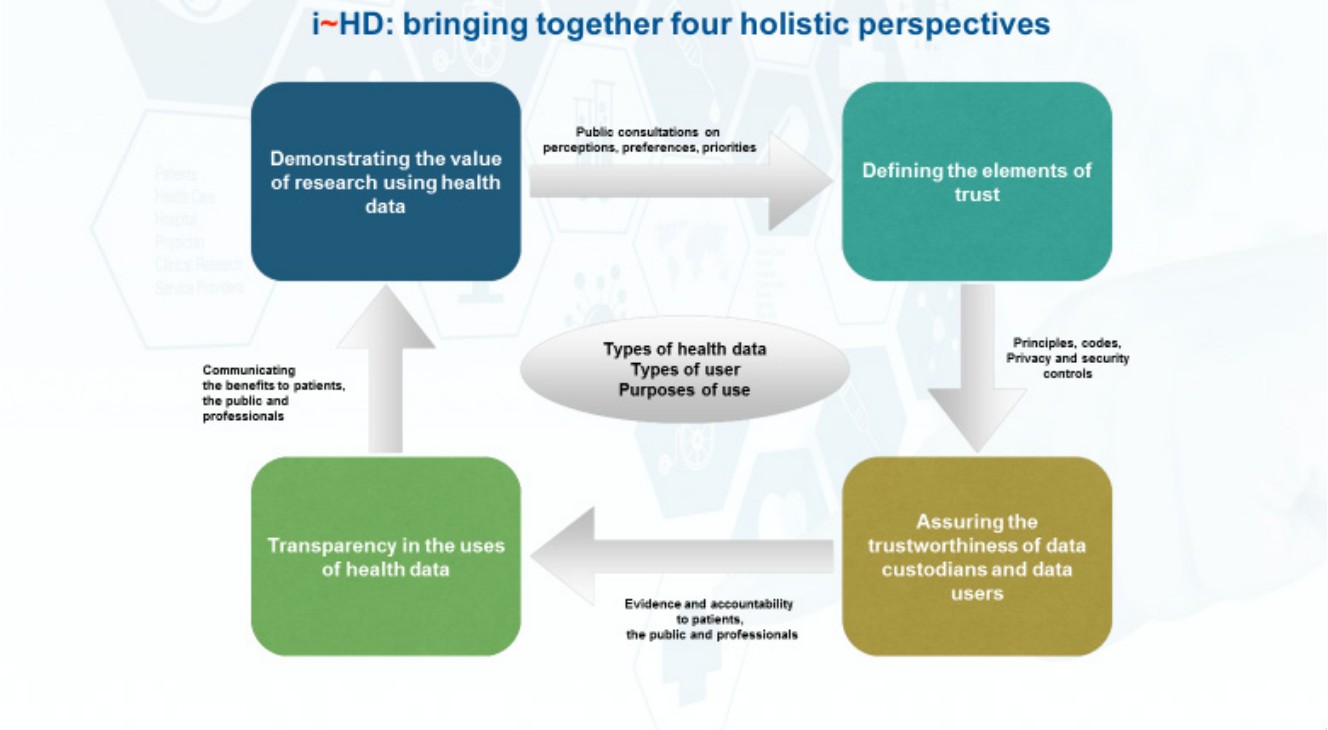
Find out from the research community what data points they need so these can be built into routinely collected data.

There is a digital health continuum – the population is neither 100% patient nor 100% citizen – so both must be considered in any new healthcare innovations. To ensure their engagement, society needs to understand the value of research using health data. Public consultations on perceptions, preferences, priorities will enable us to define the elements of trust. The development of principles and codes for privacy and security will help assure the trustworthiness of data custodians and data users.

Is there a way to facilitate public consultations on perceptions, preferences and priorities for the security and privacy of personal health data?

Evidence and accountability to patients, the public and professionals is needed to ensure transparency of the use of health data.

Communicating the benefits of research findings to patients, the public and professionals will help demonstrate its value.



DISCUSSION:

Key points raised

FACILITATOR: **Paul Timmers**, Roundtable Chairman

Governance and ethics

- There was disagreement on the need for law/regulation and ethics to be separated. Some considered this was necessary while others thought they were intrinsically linked.
- **Ethics is driven by social norms, which differ from country to country. Potential research project: societal views of ethics and transparency around Europe.**
- Within the NHS, different models of consent exist for use of patient data:
  - Research without consent – the legal route (Section 251; http://www.legislation.gov.uk/ukpga/2006/41/part/13/crossheading/patient-information)
  - Follow-up of study cohorts – consent given as part of study
  - De-identified data – consent is less of an issue
- However, industry partners may have different regulations and standards and may not recognise NHS standards. A simplified, consolidated set of standards would be of considerable value.
- Problems may also arise when trying to combine data from individual studies/datasets that are subject to their own separate agreements.
- We could start by considering specific challenges and make solutions generalisable and scalable.
- **There is potential for an EIT Resource Centre – a common set of best practice tools across Europe to establish and governance framework for collaborative data sharing. Common standards might help address the "paranoia" and "overcompliance" that is currently inhibiting innovation.**

Personal data

- What defines "personal data"? Imaging data and that from rare diseases is difficult to de-identify completely. Also, we need to consider the rise of facial recognition and where it fits.
- We must understand the actual nature and the public perception of privacy and consent regarding personal healthcare data, so we can reach agreement on how to apply any regulations.
- Could we have a fixed period of use? A "protected time" for use?

Patient and public engagement

- We need to share evidence of the impact of data collection and use for patients, e.g. the benefit of breast screening.
- Results of data use need to be publicised and fed back to citizens and patients in a way then can understand. Experience from the UK Biobank shows that 95% of donors would do it again once they understand the benefits. This communication and feedback loop needs to be embedded in projects.
- There is no direct feedback between patients and corporate sponsors of clinical trials. Corporations are often seen as profiting from data. Patients may think: "my ill health should not be sold".
- Could participation and engagement be widened by announcing/appealing for citizen feedback?

Transparency

- We need to balance the need for transparency on data use with the need for company privacy regarding project development.
- For citizens and patients, the purpose of data use is key.

Multi-stakeholder agreements

- Need to improve collaboration in multi-stakeholder agreements – alignment of stakeholder values and interests to ensure sustainability. Currently, there is significant competition between academics and institutions, so the wheel is constantly being re-invented in terms of agreements.
- Need to create multi-stakeholder agreements including information about transparency, the benefits for the healthcare system, and society as a whole. Principles need to be defined for how data is handled. We need to avoid the 2016 DeepMind data sharing controversy.
    - ♦ **Potential project for EIT Health: form an international multi-stakeholder group (academia, industry, healthcare etc) to discuss and accelerate the process of developing agreements and principles, with clear milestones and a time limit.**
- It would be helpful to get an overview of the different R&I data collaboration agreements that exist – in some cases data is shared for free while in other cases there are cost-based models.
- How do we quantify the value of real world evidence versus clinical trial data?

Interoperability

- The challenges are how to combine data to make it meaningful, and then how to translate it so it makes sense to a clinician.
- May be a role for an independent facilitator to bring different parties together to drive interoperability for mutual benefits. The more that stakeholders can see the benefit of interoperability standards, the more likely they are to adopt them.
- Ideally, we need a common data model across different countries in Europe.
- We need to link back to a value proposition to ensure adequate funding.
- While there is currently a degree of variation in data systems and a need for tools to ensure consistency, this requires a considerable resource and cost investment. We do need to be realistic and work with what we currently have, even if it's imperfect.
- It may be valuable to involve people and expertise from non-health sectors in this discussion. An example given was Silicon Valley which developed rapid, agile technical solutions to similar challenges.

Capacity building and training

- The need for multidisciplinary teams was recognised, not just including data scientists. One individual does not hold the entire skill set needed.
- Training is needed so people can better understand the perspective of other disciplines and translate between them.
- It is a challenge to get good data scientists to work in the field of healthcare; we need ideas for how to attract them.
- Look broadly to identify barriers to innovation in the workforce, beyond data scientists.
- How are we preparing people for careers in a Big Data environment within healthcare? What are the options and are we creating barriers to attracting top talent?
- We need to ensure that digital content (Big Data, AI etc.) is embedded in educational curricula and ongoing development of healthcare managers and healthcare professionals. We should also use behavioural science – what behaviour do we want to change in society?
- Other non-healthcare sectors have progressed further on this – what can we learn from them?
- Could access to data and interoperability be factors that attract students and top talent to the healthcare arena – giving them a sense of value to society?

SESSION 3: **Summary of Actions and Recommendations**

SUMMARISED BY: **Paul Timmers**, Roundtable Chairman

1. Data access	There is a wide variety of healthcare data, but the challenge is to access and use raw data in a meaningful way. **There is potential for projects/research in areas where clinicians can benefit from access to more basic data that is analysed and presented in a meaningful way, so they can make informed clinical decisions and drug choices. Pilot project idea: data from patients with multimorbidities.** **Access to preclinical data: biological assays can provide predictive information. This is a potential area for funding support as it is cost effective compared to clinical trials.**
2. Data quality	Ensuring quality involves good capture of data and good curation of data. **Find out from the research community what data points they need, so these can be built into routinely collected data.**
3. Right data and the right time (and "right now", i.e. quick wins)	We should find simple solutions that can deliver high value. We can undertake validation of dynamic solutions.
4. Complex analysis and transparency of this analysis	Potentially a big area for research. **Investigate how we can use artificial intelligence and decision support to help patients with complex multiple conditions set smart health goals.**

5. Development of an "EIT Health Data Area"	**An EIT Health Data Area could involve a service function for both stakeholders and contributors to the Big Data infrastructure, ie:**
	• **A collection of best practices, for example for data protection.**
	• **Establishing a governance framework for collaborative data sharing.**
	• **Interoperability (technical, organisational, semantic).**
	• **Actual data sharing, interconnection of data sources.**
	• **Developing trends:**
	  ♦ **Today – static**
	  ♦ **Future – dynamic**
	• **Possible stepwise programme, both static and dynamic, i.e. making a start from today and, based on trends analysis, adapting for the future:**
	  ♦ **Catalogue standards**
	  ♦ **Combine standards**
	  ♦ **Add guidance (in the form of advice rather than a full service)**
	• **Undertake a "gap analysis" of what currently exists in this arena to avoid duplication.**

6. Citizen engagement and demonstrating value	**Develop new dynamic consent models that put the citizen/patient at the centre of process. Doing this requires work/research to understand what is acceptable to patients regarding how their data is used.** **Research possible ways of developing pan-European citizen consent.** **Need research into models and methods for listening and responding to citizens and patients, to understand from them the negative attitudes to data sharing and what would change this.** **Is there a way to facilitate public consultations on perceptions, preferences and priorities for the security and privacy of personal health data?** **Potential research project: societal views of ethics and transparency around Europe.**
7. Transparency	For all stakeholders in the collaboration. Focus on benefits. One important purpose for transparency is to enhance trust. **Develop new ways to achieve stakeholder alignment and agreements. What are the common value propositions that will bring the resources stakeholders each have together for a common purpose – smarter healthcare and more efficient research?** **Potential project for EIT Health to form an international multi-stakeholder group (academia, industry, healthcare etc) to discuss and accelerate this with clear milestones and timelines.** **Research the current models of collaborative ways of working between healthcare and pharma, and identify where they are successful.** **What can more be done to promote European cohesion regarding governance and codes of practice for data sharing and use?**
8. Training	Training is critical to ensuring that we have the capabilities to meet the demands of Big Data initiatives in healthcare. There should be a link with EIT Health Campus initiatives.

EIT Health is supported by the EIT,
a body of the European Union