



Calls to Action on Health Data Ecosystems

RECOMMENDATIONS FROM
MULTI-STAKEHOLDER ROUND TABLES

Digital Health Society & the European Institute for Innovation Through Health Data Calls to Action on Health Data Ecosystems

This contribution summarises the outcomes of two recent multi-stakeholder consultations to examine the acceptance criteria for societal trust in the use of health data and a recipe for trustworthy digital health: standards, architecture and value.

The Round Tables were developed and convened by DHS and i~HD neutrally and independently from the event sponsors, Johnson & Johnson and Microsoft. Each meeting was attended by around 27 online participants from EU institutions, national governments, industry, academia, hospital management,, healthcare professionals, regulators and patient representatives. DG Sante and Connect officials contributed to both events.

The recommendations and calls to action arising from these events were presented to large online audience at the Digital Health Society Summit in November 2020 and discussed by a multi-stakeholder panel. The recommendations, the calls to action and round table summaries are documented in the accompanying uploaded report, also available at www.

The recommendations cover the following key themes.

1 Raise the digital, literacy & skills of all stakeholders

2 Generate and value trustworthy Real World Evidence

3 Accelerate interoperability across Europe and globally

4 Demonstrate benefits to society from data access, use and reuse

5 Adopt a risk stratification approach

6 Build a trustworthy framework for data access and use

7 Adopt a transformational approach to health data



**Raise the
digital,
literacy &
skills of all
stakeholders**



Member States should set target standards for population and professional digital, health and data literacy and openly share these targets at a European level.



Literacy should cover, for the public:

- becoming fluent data users for their own health
- appreciating the importance of the data they create
- understanding their rights and protections over data held by and used by others
- understanding the benefits their data can offer to society.



Literacy should cover, for existing and future health professionals and managers:

- how to use digital health tools/ data science for patients and citizens
- how to educate and support patient/citizen users of health data and digital health tools
- how to respond to and escalate issues, readings of concern
- the importance of RWE and its quality
- how to understand data science and its contribution to healthcare practice.



Researchers, regulators, public health and political decision makers also need to be health data science literate.



Healthcare funders (ministries, regions, insurers) should publicly declare an annual budget they will invest in patient/citizen literacy resources and initiatives, and how they will cover age ranges, ethnicities and other population subgroups and leave no one behind.



Education providers targeting public and health professional education should be required to share digital health curricula and learning objectives (not course delivery materials). Equally these points should be applied to curricula for health and data literacy for the education of children.



Industry should contribute to this mission by sharing educational resources and the selective sponsorship of training places on literacy programmes.



**Generate
and value
trustworthy
Real World
Evidence**



National and Regional Health data infrastructure providers and coordinators, the research community, public health agencies and European data infrastructure programmes should increase and co-ordinate investments in:

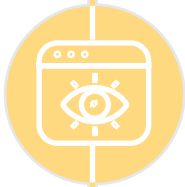
- education to raise the skills of those who need to generate real-world evidence, so they ask the right questions and generate comparable answers
- the kinds of research questions can be answered by distributed analytics, and which ones need to work on a dedicated patient level data extract
- improving data quality, starting with facilitating a more motivating culture within healthcare professionals and better EHR system user interfaces
- research into errors and statistical corrections for low quality data, and the generation of synthetic data e.g. for the training and validation of AI
- audit processes and traceability of the sources of data must be embedded into policies and architectures to ensure transparency.



**Accelerate
inter-
operability
across Europe
and globally**



Member States should embrace an alignment of standards adoption with other countries, such as on the EEHRxF, and reflect those as strong interoperability demands within national and regional procurement policy and specifications.



Future standards development strategies should involve representative data creators and users, especially health professionals and patients.



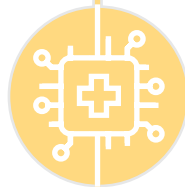
Healthcare providers should demand, from their EHR suppliers, explicit and independently verified interoperability against prescribed standards through procurement specifications and renewal contracts.



The extent of the interoperability a healthcare organisation and its supplier can deliver should be measured and made public.



Member States and the EC must support patients and citizens to become strong advocates of joined up (interoperable) health data balancing illness and wellbeing (prevention) needs.



The **EC** should more strongly encourage health data generated through its funded projects to be more widely reusable via the EHDS.

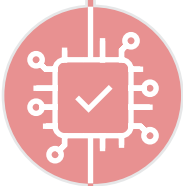


Interoperability between consumer devices which generate health data and EHRs will become increasingly important as this type of data grows in volume and relevance and must therefore be ensured through regulation or soft law.



4

**Demonstrate
benefits to
society from
data access,
use and reuse**



Data Permit Authorities and data sharing intermediaries should:

- publish lists of data uses they will normally support, and those they would not
- require the intended benefit of data use to be stated with each data request
- define the terms and conditions they will require from data users
- publish annually the benefits they have enabled, and lessons learned from reusing health data
- consult with the public to define societal benefits and value
- involve patients and citizens at decision making (board) levels
- promote and oversee good models of data altruism.



Industry should support and then adopt consensus practices on how best to communicate the benefits to society from their use of health data.



5

Adopt a risk stratification approach



The GDPR places too strong an emphasis on the identifiability of individuals from data through explicit attributes and does not give adequate recognition to unique data patterns that may enable data subject identification.



At an EU level a specific health scientific and research basis for reuse is needed.



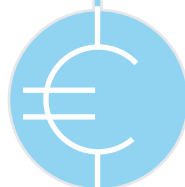
Pseudonymisation should not always be considered as personal data without taking into account the safeguards including the protection of linkage keys.



EU and national research funders should invest in further research on risk stratification methods for health data sets so that proportionate protections such as appropriate codes of conduct and suitable information security measures and can be applied consistently according to purpose and risk and not, as at present, in a piecemeal way.



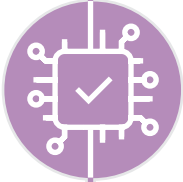
Data Protection Authorities and the European Data Protection Board should indicate willingness to develop and adopt risk stratification guidance on the use of data protection safeguards.



Member States and the EC should balance risks with the opportunity costs of not sharing health data.



**Build a
trustworthy
framework for
data access
and use**



Data Permit Authorities should:

- promote the development and adoption of multi-stakeholder Compacts regarding responsible data use, transparency, accountability, communication, by including the public (patient and civil society organisations) health funders, providers and health data organisations (public bodies and industry)
- hold open public consultation when developing governance frameworks and decision-making rules for health data uses and reuses
- include members of the public in the constitution of the European, national or regional decision making bodies themselves
- publish inventories of data use requests received, accepted, declined and of any investigations into misconduct
- conduct public awareness campaigns to explain to the public the research uses and benefits of using health data.



All public and private stakeholder should support the adoption of standards and Compacts for how data access requests are formulated and transparently reported on.



**Adopt a trans-
formational
approach to
health data**



All stakeholders should support and promote treating repositories of pooled anonymised health data as a societal good.



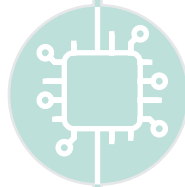
Investments should promote the uptake of federated data models to facilitate interoperability, connectivity and FAIR data access while upholding GDPR compliance.



Europe should now consolidate efforts on one or a small number of common data models so that data harmonisation methods, tools and skills can be scaled up to become a readily available and affordable resource.



Stakeholders should focus eHealth governance models, trust mechanisms and research infrastructures to contribute data to large-scale independent health data repositories that provide real-time continuity of data access for individuals, healthcare delivery and for population level analysis, with appropriate governance.



Synthetic data sandboxes should be developed to enable research into novel security approaches and the training of AI algorithms.



A transformation towards cross-organisational and independently run health data repositories will require radical change in ICT products and procurement, for which policy enablers must now be enacted.



Regional and national early adopters should be encouraged to collaborate across borders to develop best practices, lessons learned and accelerate the reuse of data and the development of benefits from it, sharing with other Member States and stimulating European competitiveness.

DHS Summit Feedback

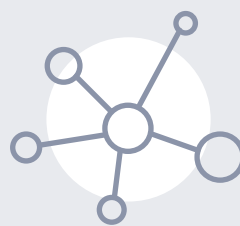
The above calls to action were presented and discussed by an expert panel which comprised Nicola Bedlington, Chair of Data Saves Lives, Jesper Kjaer, Danish Medicines Agency and Nigel Hughes Project Lead EHDEN & Janssen and Ioana-Marie Gligor, Head of Unit DG Sante. They agreed the two most important calls **were upskilling digital, data and health literacy, and generating and valuing trustworthy Real World Evidence.**

We conducted a poll of Summit attendees and the results were:

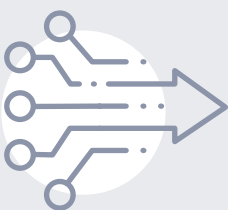
50% thought the most important call was adopting a transformational approach to health data.



82% thought if health data is to be a societal good it should be defined by a group formed of multiple stakeholders.



72% thought a list of data uses that would normally be supported and those that would not be supported should be developed by a group formed of multiple stakeholders.



73% thought that, to develop trust in data access and use, they would prefer to see a combination of written laws/regulations and multi-stakeholder codes of conduct.



REPORT OF ROUND TABLE
VIRTUAL MEETING ON

Acceptance criteria for societal trust in the use of health data

Round Table Summary

This report summarises the topics, discussions and conclusions of a multi-stakeholder Round Table held on Thursday 3rd September 2020 on acceptance criteria for societal trust in the use of health data. Its aim was to propose **criteria for building and retaining societal trust in the uses and reuses of health data**, across a spectrum across direct care, public health, health system improvement and research. 27 participants, comprising patient organisations, healthcare providers, payers, ministries, data protection authorities, industry and industry associations and representatives from the European Commission participated in a highly interactive half-day meeting designed and run by the Digital Health Society and the European Institute for Innovation through Health Data, sponsored by Microsoft and Johnson and Johnson. The Round Table sought to consolidate what society and decision makers would regard as **acceptable conditions and terms for access to large scale data resources**. It has been timed and offered as input to the scope, design and governance framework being developed for the European Health Data Space.

Right across the learning and innovation ecosystem, there is a growing need for large scale access to health data. A momentum for European cohesion on data access, harmonised criteria and governance, has been accelerated by the European Health Data Space. How-

ever, **there is less public understanding and therefore trust for uses of data that are not directly applicable to the individual and performed by organisations who seem less familiar within the health ecosystem**. Participants considered the challenge of societal acceptance criteria from three perspectives: the who, what and why of data use and reuse; technical and organisational safeguards; transparency and trust about use and value. Through breakout group and plenary discussions, the following themes emerged.

There is a big difference conceptually between data use to benefit the individual and larger scale data reuse that has the potential to benefit many (but might not include individual data subjects). **The public, and individual data subjects generally support data reuse if it clearly explained to them what the beneficial objective is**. It was recognised that there is no universally accepted definition of beneficial use, but that **illustrative lists of purposes that would normally be supported by decision making bodies, and purposes that would not be supported, are helpful for guiding the public and guiding decision makers**. It would additionally be reassuring and strengthen public support if list of approved uses, and denied uses, would periodically be published. **This is in effect a combination of transparency of intention and transparency of action**.

It has been found by a number of patient and public perception studies that ensuring that the use purpose will deliver a benefit to health systems and that this benefit will be affordable, without excessive profit taken, is the most important criterion for support. Whether the bodies involved are commercial or public, whether they are classically associated with healthcare or not, are less important factors. COVID -19 has demonstrated the level of public engagement and support that is possible if the purpose for data collection and use is clear and in society's interest.


The importance of transparency to the public as well as to individuals whose data might be reused was a dominant theme throughout the event, across all of the three breakout groups. This was perhaps considered to be the most important success factor. It was emphasised that **this transparency must be inclusive**, including vulnerable groups of people whose data are equally important and who should be inclusive beneficiaries of the outcomes from using data (a potential adaptation of the concept of reasonable accommodation was discussed). Inclusivity may have economic challenges in a single market, but on the European scale and through the use of European standards, inclusivity can be made economically viable. **Greater public and health workforce education, including data literacy, digital literacy and health literacy, are therefore essential success factors as well.**

Decision-making bodies, and the governance frameworks that they operate under, are more likely to operate at Member State and/or regional level rather than EU level, for legal, political and practical reasons, but the governance frameworks they utilise should be as consistent as possible, across Europe. **The public must be involved in developing their governance frameworks and decision-making rules** and should be included in the constitution of the bodies themselves.


Public fears about misuse, including fears that information will be used in some way to disadvantage or discriminate against individuals or minority sub-populations, must also be addressed by such bodies.

These fears are very powerful and if they are not addressed, they risk dominating over the perceived benefits of using health data. An important mitigation for this fear is, again, transparency. **The public have to know how their data is being used, and how it is not used.** Even when it is not feasible to give individual level control over all possible uses of data, **the public then need to have confidence in the organisations that are making decisions on their behalf.** In cases where data have been anonymised it is not easy to provide personalised feedback, but collective published feedback about how data have been used to populations may prove sufficient.

Many of the reuses of data, especially for research, public health and health service improvement do not need identifiable data, but they do often need fine-grained, close to real-time, data including longitudinal histories and increasingly including specialised data types such as genomics. The biggest concern for citizens is whether they could be identified from a dataset that is being shared or accessed. **The GDPR strongly distinguishes pseudonymised from anonymised data, but it was argued that fine-grained data can never be truly anonymous.**



This historic distinction is perhaps no longer viable, and a risk stratification approach which takes into account the way in which the data are being processed and protected and the benefits of use through information security measures is more appropriate.



Synthetic data, in which noise (perturbation) is added to the data in order to prevent individuals being recognised even from very rich data patterns, is a method that is gaining recognition as a method for some kind of population level research.

It is usually fruitful to think about the interplay between what we construct by means of technology and our social constructions (laws, codes of conduct, organisations, behaviour). Law should not be made without thinking about how technical constructs can help enable compliance and enforcement. Technology initiatives should not be developed without considering how and by whom these initiatives will be governed. A third, psycho-social (people oriented), dimension is also important.

There was considerable discussion about a code of conduct. Although a formal and possibly legally enforceable code might be developed at a European level, **there was support for organisations, especially health and health-related companies, to come together and to develop voluntary codes of practice that they agree to adopt: known as a compact.** The public would be most assured if this is **a single code developed through multi stakeholder engagement including patients and** the public, and was adopted by all health data user organisations (commercial and public, and including patient and civil society organisations themselves when they collect and use data).

Introduction

This report summarises the topics, discussions and conclusions of a multi-stakeholder Round Table held on Thursday 3rd September 2020 on **Acceptance criteria for societal trust in the use of health data**.

Its aim was to propose criteria for building and retaining societal trust in the uses and reuses of health data, across a spectrum from direct care, public health and health system improvement to research. The Round Table sought to identify what society and decision makers would regard as acceptable conditions and terms for access to large scale data resources. This report is therefore intended to help frame future European initiatives to develop better formalised models for data provision, use and governance, to better position new actors (e.g. industry) in roles such as healthcare delivery partner, care pathway redesign partner, analytics partner and knowledge partner. In particular, this Round Table and report have been timed and offered as inputs to the scope, design and governance framework being developed for the European Health Data Space. It may guide the development of any necessary enabling legislation and policy instruments, industry promoted standards or codes and innovations in information security safeguards. No individual stakeholder is able to solve the challenges and now more than ever we need a deep collaboration which strikes fair balances for all to enable the common good.

The Round Table was an invitation-only, multi-stakeholder and highly interactive half-day online event with 27 participants, dividing for some of the time into three virtual breakout rooms for deep dive topics. The agenda is given in Appendix 1. The participants included patient organisations, healthcare providers, payers, ministries, data protection authorities, industry and industry associations and representatives from the European Commission who are architecting the European Health Data Space. The list of meeting participants is given in Appendix 2.

The event was jointly run by the Digital Health Society (represented by Bleddyn Rees) and the European Institute for Innovation through Health Data (represented by Dipak Kalra). It built on the Digital Health Society's Summit in Helsinki with the Finnish Presidency last December when both organisations collaborated on the data and digital content.

It was sponsored by Microsoft and Johnson and Johnson, who contributed financially for preparing and running the event but did not control the structure, hosting, content or reporting of the event.

OPENING PLENARY SESSION:

Scene setting

Bleddyn Rees and Dipak Kalra welcomed participants.

Bleddyn set the scene for this Round Table, which has built on prior related events over the past several months, starting with a Digital Health Society Summit in December 2019, which highlighted many of the issues and challenges that impact on how the public and patients understand and indicate preferences for, or control over, the uses made of health data. In the spring of this year DG Santé has run a series of consultation workshops with strong DPA participation, moderated by Petra Wilson. In May, the DigitalHealthEurope project ran a virtual focus group for industry about company aspirations and potential contributions to the European Health Data Space (EHDS). The industry participants highlighted the special opportunity for architects of the EHDS to develop a coherent governance framework that could be adopted by other European data initiatives, thereby helping to harmonise approaches adopted across Europe. This Round Table was designed to contribute to that aspiration, by

taking a deeper dive on societal acceptance factors for data reuse that might be taken on board when developing the EHDS governance framework. A future event is planned by Digital Health Europe on the perspectives of patient organisation representatives on this topic. These events have some organisers and participants in common, and are sharing outputs so that their progression is complementary and additive.

Dipak reminded the audience that there is an explosion of the opportunity space to learn more from health data, as more and more kinds of data are being captured about and by patients and citizens, and are potentially combinable if this is permitted to build up rich pictures of healthcare, health outcomes, wellness and wider influences on health etc. We must take a future looking vision on the availability of data. We should especially note the most rapid growth area will be citizen generated data, and our approaches to governing data use must actively win citizens' trust in sharing their data.



"Personal sensor data is expected to grow to 90% of all stored information within the next decade"

"90% of the data in the world today has been created in the last 2 years"

"By 2021 there will be almost as many personal assistant bots on the planet as people"

"> 1billion have access to mobile broadband internet"

Right across the learning and innovation ecosystem, there is a growing need for large scale access to health data. Many of the innovations we are developing or foreseeing need to benefit from vast volumes of health data. This may be from conventional healthcare sources (e.g. detailed EHRs), patient and citizen generated, medical devices and non-health sources such as pollution. The analysis needs are for this to be fine grained, individual level data (normally anonymised) so that precise and novel analyses can be undertaken. Pre-compiled aggregated data sets, or data warehouses refreshing their data every few months, are no longer adequate. There is also an increasing demand for the data to be close to real time (so that real time feedback systems, for example driven by AI, can be developed), and for this to be longitudinal, reflecting health, wellness, disease trajectories and outcomes.

This growing data need has stimulated national and EU level (and international) investments in large scale data resources and networks that offer these opportunities. The different existing and emerging infrastructures comprise a mixture of eHealth (digital health) services and research infrastructures. They are implemented via a mixture of centralised and federated architectures. These different infrastructures are often set up quite differently. They may process different kinds of data, from different sources, serve different purposes and user types and have different governance frameworks. This makes it difficult to win public trust at a European level.

The COVID-19 pandemic has stimulated public awareness and support for generating rapid insights from data, and the attitude to data collaboration amongst data using organisations.

This has included important for the collection of personal data by a range of organisations (such as restaurants) that would not normally have this and allowing location data to be used by agencies that do not normally have this either. However, it would be wrong to assume that this societal goodwill is going to be permanent or can now be relied upon for many other desired uses of health data.

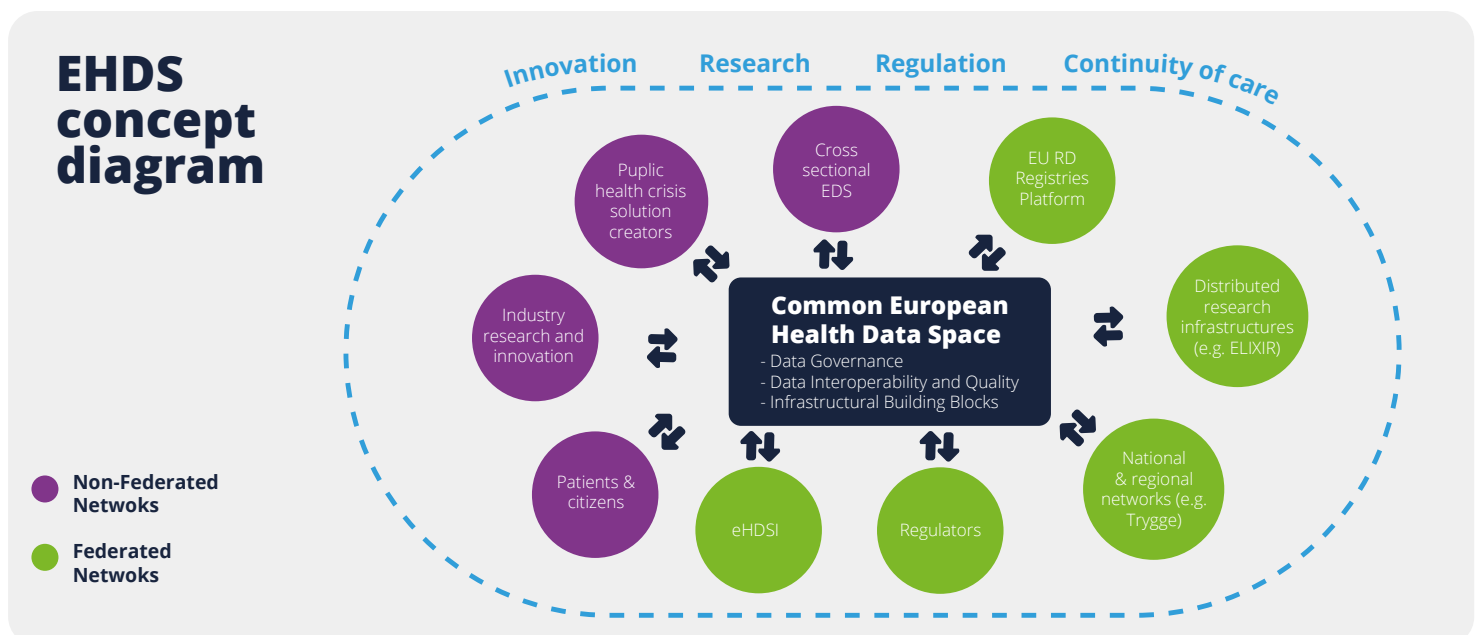
A momentum for European cohesion on data access has been accelerated by the European Health Data Space (EHDS). The details of how this space will be designed, what data sources it will contain and which ones it will network to, and how it will be governed including its uses as well as the terms of use are still being defined. However, the following concept diagram, developed by DigitalHealthEurope, seems to be a plausible concept model. Europe already has several existing data network infrastructures that might be interconnected through the space, which could then offer a single portal for access to permitted data extract from these networks. This includes the eHealth Digital Service Infrastructure (which presently exchanges patient summaries and electronic prescriptions between some member states, but has a roadmap to extend the range of electronic health record data sets to be communicated), the DARWIN network being developed by the EMA and the national medicines agencies for the exchange of medicines information including for pharmacovigilance, the European Reference Networks that especially connect centres across Europe caring for patients or conducting research in rare diseases, and the life sciences research infrastructures. There are additionally, stakeholder groups which have data accumulations that could be contributed into the space, as physical data or as network

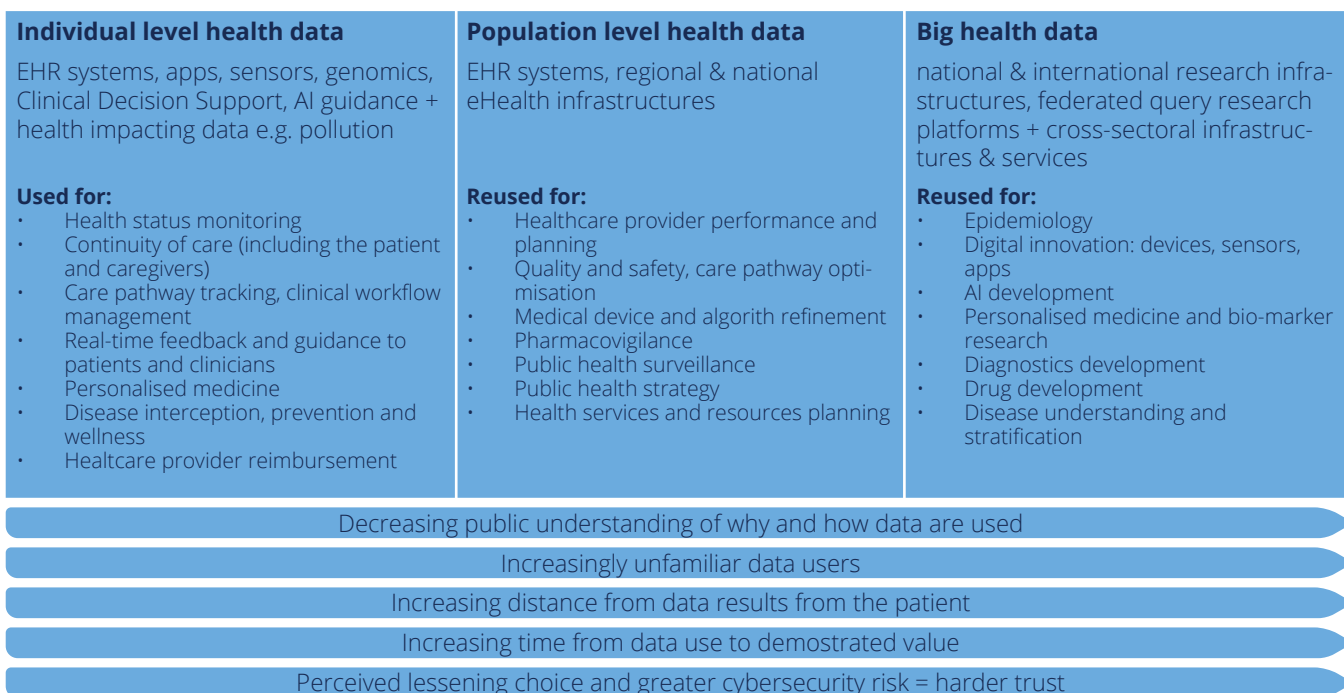
connections, subject to suitable agreement and terms. This includes industry, such as Pharma, MedTech, Telecoms and large ICT companies, and public health agencies that are starting to accumulate data in response to emergencies such as COVID-19. Patients and citizens are an important potential data contributor, as well as data user, as mentioned earlier. These different data sources span healthcare, research, tech and regulation, and the uses of the space may cover any or all of these areas. Data quality, interoperability, collaboration and governance have to be right. The public have to be on board for this to succeed.

The diagram below illustrates, in the upper portion of the diagram, many examples of the potential agreed uses of health data that occur at an individual level (close to the patient or citizen), at the level of regional and national health systems (for public health or health service improvement purposes) and at an even larger scale for the conduct of research. All of these example uses occur today, but some of them are relatively local and are only just beginning

to scale up. Personalised medicine and AI, for example, will become increasingly important.

However, when shifting from left to right on the diagram there are several challenges that are faced by the public when it comes to accepting and supporting these data uses. This includes the more limited widescale public understanding of the right-hand side uses, how they are undertaken, with what kinds of data. Additionally, the kinds of organisation that become involved in undertaking those uses are less familiar as health stakeholders to the public. Knowledge derived from populations of patients may take a long time to feed back to visible public benefit, and the benefit may be perceived by different people from the ones whose data was needed to generate a knowledge. This all creates a progressive disconnect between patients and the public and the uses and uses of data, with reducing direct engagement and choice, making it harder to win public trust and provide public assurance. We need to develop a new consensus on acceptance criteria for the uses of data.





Europe has seen many public attitude and patient attitude surveys in recent years, which have been conducted by many different organisations, using different methods and especially using differently framed questions (some of which have not been well worded to yield precise answers). The result is possibly a more confused picture of public opinion than a helpful one, and we need to recognise that public confidence and trust in the uses of data they understand less well, such as genomics information, is lower than for classical clinical data. However, there is a general consensus that the public does support the use of health data for quality improvement at research that is targeting new or better solutions for diagnosis, treatment and prevention, but not for non-health purposes.

This was the starting point for the Round Table. Its objective, through three breakout groups and subsequent plenary discussion, was to

identify some ground rules for building and retaining societal trust in the uses and reuses of data. This should be across the spectrum of purposes and users, consider different access and governance models, how transparency should be demonstrated and what acceptable societal benefit should look like. It is hoped that an eventual set of endorsed acceptance criteria would give greater confidence to data providers and data users about data availability and what access arrangements are permissible, acceptable and serve to catalyse greater data availability and data use.

Participants were divided into three breakout groups to discuss specific facets of this challenge:

1. The who, what and why of data use and reuse
2. Technical and organisational safeguards
3. Transparency and trust about use and value

SUMMARY OF DISCUSSIONS IN BREAKOUT 1:

The who, what and why of data use and reuse

Moderators: Dipak Kalra and Zoi Kolitsi

1. Individual level health data v. Big health data

General observations

- There is a big difference conceptually between data use and big data reuse (“is my data being used for my care or in a research setting?”). Dividing the data use landscape in more ways, as in the figure shown earlier, might not be helpful when it comes to acceptable data use.
- It is very important for people to be able to trust the people who use the data. Lines get blurred when citizens don’t know what will happen to their data → transparency was agreed to be key in order to build the necessary trust.

A common viewpoint participants had encountered is that if patients are told clearly how a particular data use will lead to a direct benefit to them (e.g. for a new type of treatment for them) or a benefit to others (e.g. treatments for other patients, but might not include the patient him or herself), they willingly give their consent.

- » It was stressed that, as experienced through COVID-19, people understand that putting data into a pool contributes to the greater good. They do not need to get personal benefit in order to participate (a crowd-sourcing spirit). They are nervous and much less trusting if the benefits cannot be clearly explained.
- The paradigm of rare diseases (RD), which was agreed to be a situation where patients do not see a hard divide between care and research. They want their data used to improve diagnosis and treatment. This gave rise to the following questions:
 - » How far are we from extending the approaches we adopt for RD patients, in which research is an integral part of care, to other patients and conditions? Might RD patients help to champion the research agenda for other patient groups?
 - » Can these RD concepts be easily understood as we present the benefits to society?
 - » Is it productive still to separate primary and secondary uses as we do today (care vs. research) or can we make things easier by looking at the bigger picture? Would it be realistic to progressively blur the boundaries between primary and secondary uses for commoner conditions?
 - » It was agreed that the area of rare diseases shines a positive light on the use of data for research, given that the medical need is so high. Patients are much more willing to make their data available, but also are contributors to finding the solutions.

» RD should not be completely separated from commoner disease research. These should be linked: common disease research will progressively be based on smaller populations of patients, which might eventually be applicable to a single patient.

» We all agree there has to be a link between care and research but there are separate demarcation lines – a single patient is of no interest to the research data user, but care needs to have nominative (identifiable) data.

- Can we use these different groupings of purpose to better understand the worries that patients have about how their data is used? Would it be productive to look at the groupings their acceptance in a different way, to understand the worries and the solutions to those worries?

Concerns expressed

Misuse

- Preventing the misuse of data, having transparency about proper uses and the reasoning behind these uses is important. There must be understanding and awareness about what it means to have access to a patient's data and for this to be used for the broader good.

“Fear is the strongest emotion”

- As seen and concluded from the enquires run at the European Commission over the last three years in data sharing, fear is the strongest of the emotions relating to data protection, rather than the potential for good.

The fear that data is being held to a patient's disadvantage

- Testimonies of patients have shown that in spite of various explanations offered to them, they fear that data they have shared will then be used to their personal disadvantage (whether this be penalising by insurance companies, for marketing, discrimination, preventing their career progress e.g. at work, cyber-criminality: blackmail after the hacking of private information).
- It was agreed that it is difficult to fight these fears about health data misuse due to a lack of faith in policy making and political decisions, as well as a deep-rooted fear about big private companies/industry taking advantage of the data but then pushing excessive pricing of products onto health-care systems. There may be value in listing as prohibited those misuses that the public have greatest concern about.

Monetising data

- The topic of monetising data isn't something that is spoken about much these days but was rumoured in the past to be lucrative. However, it was agreed that this issue has not gone away. It was agreed that this also links to data ownership. Some ICT companies (e.g. app developers) offer their customers free services in exchange for permission to use their health data. Another point was made that many hospitals don't understand – but sometimes overestimate – the value of their data. (A comparison was drawn with the data of an individual being of little value in that regard).

- We need to use a lot more data efficiently and effectively and safely, and to bring the public along. We can only overcome the fears mentioned above by showing the public the benefits from making good uses of data.

Assurance principles/core principles

- It was agreed that it is important to illustrate the positive examples of how using the data of large populations can lead to more effective treatments, with concrete case studies.
 - » There was a recent Belgium consultation on genomic data undertaken with engagement of students and teachers and a wider engagement of citizens which highlighted what the fears are. We need comparative qualitative studies to better understand what the issues are, and then to see what the impact of education is on those fears. Patients may then become actors, not just passive data donors.
 - » The example of Citizens' Juries, conducted by the UK Connected Health Cities programme, was discussed¹. In order to understand the scenarios for which participants would agree to the use their data for research, they were offered a three-day education course, which encouraged people to be engaged in understanding data use and to voice informed decisions. This helped to highlight what aspects of a proposed data use influenced decision making the most. However, this brought up the question as to how feasible it would be to carry out a similar exercise on a whole population.

» The public has to understand why, how and by whom data can be usefully used.

- The issue of trust enablers and transparency was discussed alongside a criticism of GDPR.


» Although GDPR is a means towards achieving trust, it was discussed that the Regulation is still in infancy when it comes to helping to ensure that citizens can find out "what the system knows about me".

With regard to matters of control, citizens should be able to decide (authorise) who may use data and for what purpose. Citizens cannot easily find this out. Although this is an area that GDPR tackles, it is pragmatically not happening in real life. If data is used under a public interest legal basis the citizen is not asked, does not decide, does not control and is not normally informed about that use.


A third important trust enabler is the need for greater public knowledge and access to education and literacy (health, data and digital) – from primary school upwards, about how to treat your own health data, about how to deal with common health conditions.

» These are three empowering enablers.

- However, how realistic is it for patients and the public to be able to read and digest about all of the different possible uses of health data, and to understand these well enough to be able to make informed decisions? We cannot ignore the role of governments here, to create integrated and co-ordinated governance frameworks, where citizens can see the principles, the values that governments respect, and what are the monitoring measures being taken.
- It was agreed that there is an interesting dynamic between citizen responsibility and government assistance. It was concluded that a citizen needs to be active and to locate/find if there is a governance framework that they can be part of designing or influencing, part of which is transparency.
- On the other hand, we must get away from the notion that everything must be put onto the shoulders of the citizen. There must be an agency that they can trust to make decisions about data use on a collective behalf, and that they can monitor.
- Some of these ideas were seen to be mostly applicable to situations in which personally identifiable information is being used. We also need to make more transparent use of solutions that don't include identifiable data or patient level data. Distributed networks and analytics hold promise for removing that need.



We need education showing the public how their data can be useful without their identity ever being divulged to researchers. This could be through lists of example uses and more detailed case study examples. With the evolution of analytical methodologies, we can do virtually everything researchers need with data without making personally-identifiable data accessible.



BREAK//////////

2. Purposes and people

- It was agreed that the ways forward for encouraging people to be fear-free and to share their data is through transparency about having done the right things with data. This is paramount. The conversation then focused in that regard on the purposes and players.

Purposes

- An increasing number of organisations are becoming part of the health data ecosystem.
- Similarly to the growing number of food ingredients, a list of acceptable purposes may

help to address the need for transparency and trust building. It may be illustrative.

- It was agreed that there is no need for a comprehensive list of purposes. It may be too generic (or would need to be infinite!).
- The need remains for ethic committees and local government bodies to approve or disapprove a research data use, as well as educating citizens about the role of such committees.
- A blacklist of the purposes for which data will never be used would also be useful.
 - » Attention was brought on the Guidance note for researchers and evaluators for social sciences and humanities research 2010 – which sets boundaries through a list of what data may not be used for.
 - » It was also mentioned that a list such as this exists in France². This includes prohibited examples such as marketing about a pharma product and insurance purposes. In France public interest is also a ground for allowing data use (though this is not enacted by written law) – but this should be formally defined and ideally then enacted/ adopted into written law. There is patient organisational involvement in that decision making.
 - » Within this context, it was also mentioned that perhaps the Declaration of Helsinki could also contribute to determining a blacklist.
 - » A list of purposes that would be prohibited is also useful as it is illustrative of the boundaries a decision-making body would not cross.

Organisations/Players

- An illustrative list was shown through the slides and the discussion showed that public surveys have come down against having a fixed list of approved organisations.
- The conversation led to questions such as whether any kind of organisation should not be granted access to health data and how the public perceives their data is being used
- To address this, the findings of a recent Citizens' Jury in Scotland were touched on. The most important decision-making factor was the purpose for which the data would be used, but not with whom the data would be shared. There was support for improvements to their own health, the health of others or the health system, but not use for a purely financial purpose.
 - » Specifically, it was mentioned that participants asked, "Why is my data not already used (like in banking, to improve experience)".
 - » On the topic of data donations, if the purpose is to save a life or to serve public good, the common answer in the Jury was definitely yes, we would be very keen to share our data.

SUMMARY OF DISCUSSIONS IN BREAKOUT 2:

Organizational approaches: Safety and Acceptance

Moderators: Paul Timmers and Nathan Lea

1. What standards should be set around anonymisation, given the challenges with genomics, fine grained location data, rich clinical profiles, rare diseases...?

- Are we getting to the stage where tech is moving so fast that it is difficult to make anything pseudonymous/anonymous?
 - » Rare disease
 - » Is this true of synthetic data?
 - » Avoiding the wicked question: how do we define a consent-based model?
 - Need an open and honest dialogue with citizens
 - They raise different concerns to medical professionals
- Transparency of how citizens' data is used
 - » If you can give them a basis to consent and be aware of data use it builds trust
 - » Need to have a risk framework around different kinds of data – e.g. time limited trust, certain data like family history of rare disease. If shared more widely can cause issues and have implications for others.
 - » Truly synthetic data³.
- Main risk is of citizens handing over control of their data if they can be identified again.
 - » Maybe many citizens would like to donate fully personal data – perhaps altruistic plus personal benefits (different motives).
 - » Citizens would likely be more ready to hand over detailed data if they knew how it was protected, where it was going and who would see it.
- If citizens knew it could be used for advancing medical innovations, they may be happier to share, but they may be worried about insurance company interests and how data may be used against their interests.
 - » But each of us is different – what level of risk will you tolerate? May have a generic risk band we opt into, for health data sharing. Anything out of someone's band needs special permission.
 - » Near real time use – real time management decisions. Recognition of social determinants of health may be increasingly important for guiding care.
 - » Decisions are made on trusted data – currently about 16% of that which is available.
 - » Blending of citizen data with health service data in a single environment is key.

³ The term synthetic data is used variably, at present. In this context the term was used to describe real health data extracts that have been modified to a sufficient extent that individuals could not be identified from the data. Techniques include blurring, rounding, date shifting, making coded terms less granular... An alternative meaning is when data is generated entirely by a computer algorithm that has parameters which ensure the fictitious data value ranges resemble real population norms but resemble no specific individuals.

2. When and how (even, if) pseudonymised data can be safeguarded enough to be considered effectively to be anonymised to a data recipient/user?

- Risk stratification, where there is no binary yes / no. Unlikely to hard code anything in particular because all data is so different.
 - » Could we develop tool sets to test that: e.g. could the data be reassembled?
- Aggregated data implies the need for less protection. GDPR still applies to pseudonymized data, so that determines how you proceed.
- This is hard – an issue is that anonymous data has a use but in many cases it is next to useless. This is about a world that is evolving – the big games are now in personalised prevention where anonymous data is not useful
- Who makes the determination?
 - » Data Trust – a third arbiter that represents that patients' say. The right construct. Political cycles make them less appropriate.
 - » Can Data Trusts/arbiters keep big platforms under control?
 - » Trust brokers – independent of all players but ensuring data subjects' concerns are addressed in their determinations. Charities and patient organisations may play a role here.
 - » Data Permit Authorities...?
- GDPR created to develop common rules within the digital single market (and beyond)
 - » Rule harmonisations
 - » Definitions of trust – are these culturally bound to different jurisdictions?
 - » Can we have common rules?
- The wild west – trading privacy for convenience
 - » Some discomfort with what convenient use implies for wider data use.
 - » How can you - the citizen - use your data as well as for the common good? Can you yourself use your data to negotiate insurance? Trade your data for health? You can defend your own interests, since you define them.
 - » COVID-19 Apps – to what extent is that choice enforced?
 - » Look for game changers in the privacy / convenience trade-off

3. Rules regarding authorisation and access controls, restricting indiscriminate “internal reuse” of the data by large (public & private) organisations?

- Endgame is that the citizen controls their data, or it is controlled on their behalf with a full audit trail
 - » There would be exceptions under consent
 - Study consent – use only as long as consent and participation permits
 - Where are these scoped by contract?
 - » Dynamically withdraw data for wider healthcare access
- The nature of a study is different. Right to be forgotten may not be realistic given the seemingly contractual nature of studies and participant responsibilities as an active participant.
 - » Find mechanisms for changing minds as well
- Ethical determinants around non-punitive opting out and cannot force people to give their data
 - » Issues are often covered by GDPR
 - » Is it more useful to have something at EU level harmonizing the selection of legal bases?
 - » Harmonized interpretations? Probably... There are variations on use of consent, public task or others for collection.

» Is individual more secure if they know how data is secured and across a more harmonized framework?

» Avoid shopping around where there are variations in protection.

4. Concluding points

- Architecture is the enabler if supports everything we have covered, and it is feasible.
 - » Tech is malleable
- Codified real time data for life – clean. Obligation to citizens to make this data available to support their care and care of others
- Keen to see a more harmonious and harmonized health regime in Europe.

Great potential especially with AI to revolutionize medicine – find a way to strike the balance with concerns around privacy and sensitive data.

- This is about better tech and organisational safeguards. What is “better” in this case? Is it faster?
 - » Make it possible
 - » Tying ourselves in knots in the status quo
 - » Opportunity to make things smoother
 - » Fairness relates to how data is acquired and its accuracy, but it is hard to find what is fair.

SUMMARY OF DISCUSSIONS IN BREAKOUT 3:

Transparency & Trust about data use and value

Moderators: Bleddyn Rees and Carina Dantas

Context for the session

Many studies have suggested that the overriding discriminating characteristics in the eyes of the public about the use of health data mainly refer that the use should deliver clear public benefit, especially benefits to the health system and to citizens, and should not primarily be used for personal or organisational gain of the organisation using the data. Although deeper dive research has danced around this topic, it has so far proved difficult to formally specify ways of defining public benefit and health benefit.

1. Demonstrating health or societal value

- Do we need a definition of health or societal benefit? A definition is challenging.

Engaging patients/citizens for co creation and public-private partnerships (extending IMI to other industries e.g. tech, MedTech, automotive and energy etc.) on projects to meet unmet need has real value.

- Health and societal value should be seen in the perspective of “common good” – But does common good apply only to the entire population or specific groups/segments of the population (e.g. rare diseases). How large does a segment have to be or serious/valuable the challenge being tackled?
- Information should be provided to citizens / patients not only at the beginning and end of a study but throughout its duration.
- When a study fails, the results should also be shared – lessons learned has real value and failure can avoid future failings. Failure should not be stigmatised and can be as important as success in terms of learning.
- The value of published research, product and services was discussed; they are all valuable and understood as diverse but all forming different parts of the chain of innovation. Value will vary but as long as passes a minimum value of benefit to patients/people that is acceptable.

Value is also seen differently depending on the type of organisations (e.g. research, industry and patient). Criteria to judge value might be helpful e.g. reducing, inequalities, improving outcomes, improving financial stability of health systems.

2. Earning trust

- Can we, should we, incentivise citizens or patients to share data and build trust? Patients are usually more willing to share data than a “common” citizen – this distinction needs to be addressed. Also, incentive is a term that may be misunderstood. Need to consider what incentives or benefits might be possible and desirable.
- It is important to know/understand mainstream opinions but also the different segments/minorities. Can the concept of “reasonable accommodation” (taken from disability legislation) be adapted for data sharing? Use of facial recognition might not be 100% compliant with GDPR but its use for persons with disabilities/vulnerable people may be justified as “reasonable accommodation” and an alternative to consent. A form of exemption allowing societal benefit to override compliance (think public interest justification).

Literacy (digital & health) is essential to empower people (they need to understand) and it is also key to provide strong safeguards that assure people their data will not be misused. More work needed on these possible safeguards.

Messaging needs to be clear and balanced – clear information, transparency and consents in order to develop data trust and altruism.

- Trust can be guaranteed by which measures? Legislative, information, (some countries have evolved systems e.g. data altruism, data protection), others – is there the need for concerted actions in multiple perspectives? Insurers use of health data appears to raise concerns especially in countries with insurance funded systems which could be controlled via legislation.

3. Missing further information citizen perspective on data sharing

- Information to citizens, that are not only misinformed but sometimes also suspicious – this is a key element to be addressed
- Why health data is secure and its use safe.
- What data is being shared and the benefits.
- Dynamic consent was discussed as a potential tool. It was highlighted it is necessary to understand what it implies, so that it is not a burden to citizens/patients [very short discussion as ran out of time].

Is a digital contract a way forward?

A code of conduct or a “compact” would be a good idea. Compact denotes more common agreement /commitment to each other and purpose with organisations collaborating to achieve those purposes. It needs to be practical and pragmatic and trades associations should also be involved (linked to GDPR, responsible AI governance and real-world evidence). A compact must be transparent and create accountability. A clear advantage is a compact/code is able to adapt and be amended much quicker than legislation.

- Compact/code could be promoted and provide assurance with kitemark etc which signatories could use.
- Should there be separate Compacts/ Codes for data used for AI and Genomics or one for all data?
- Maybe take advantage of ongoing initiatives (e.g. European Framework for Ethics in the ICT Profession; COVID-19 as a use case) to advance on a compact that could be commonly agreed by multiple stakeholders and the basis of a social commitment.

Principles / Criteria

Based on these discussions some key Principles / Criteria were presented and discussed, inspired from the B2G Data Sharing for the public interest:

1. Proportionality in the use of the data
2. Purpose limitation
3. Do no harm
4. Conditions for data use (contract / terms)
5. Transparency
6. Accountability
7. Fair and ethical data use
8. Citizen Involvement and Centric uses
9. Clear definition of societal/public benefit
10. Duty to share
11. Clarity of legal issues
 - a. liability for data quality
 - b. IP rights
 - c. competition law

The above need to be tailored to healthcare and detail develop for each.

Plenary feedback and discussion from the Breakout groups

Break-out 1: rapporteur Zoi Kolitsi

Are we comfortable that we can have a discussion around different situations in the use of health data?

» **Feedback:**

- It can sometimes prove challenging to have these conversations, especially if
- the purpose is research rather than care.
- Reflected on the RD paradigm – very specific case.
- Are we moving towards a situation where, with better education of citizens, we can reach a better perspective on the uses of data?

» **Question that we discussed:** what are the concerns and what are we developing criteria for?

- Data misuse came up most as the concern that people are most worried about. They fear that data can be used against them as individuals, such as insurance or marketing.

» **The assurance principles:**

- Informing people better, leading to more knowledgeable citizens participating in governance and in decisions

- Putting data visibly to good use for society
- Trust enablers to address misuse of data
- The role of government and appointed bodies is important to monitor proper governance

» **Question:** if data plays such an important role, but has risks, would it be helpful if any governance body take on responsibility to control or check on the uses of data? Elaborate criteria on appropriate uses of data and what should be prohibited?

- Agreement – The group discussed a good list and a blacklist.
→ FR has a positive and a blacklist – we will look into this.

» **Question:** the role of industry and if we should define the types of organisations that should have access to health data.

- Real life intervention from Scotland from findings from Citizens' Jury
→ Patients are willing to share but need to understand the purpose and the financial benefits to the user of using that data.
→ Didn't need to know if it was a private or public organisation.
→ There was a definite yes to donation for public benefit.

Open discussion

» There was discussion about the whitelist and blacklist Dipak had presented. It would be helpful for the public to have an indication of the kind of areas that wouldn't be acceptable.

» It was suggested the lists could be developed further in consultation, and might become stable, but should not attempt to be complete (because this would not be possible, and there would be new additions periodically). These could be illustrative and used to guide a decision-making body. They could be complemented by a growing list of actual permitted and refused decisions.

» The lists can be fine-tuned and can be a living document – any authority could use this and point to it.

- There was agreement with the principle of having such lists.
- It was recognised that further work was needed to complete the lists.

» There was discussion about how such lists would work. Are these examples or are they criteria? It was agreed these are examples to help to convey the approach to deciding on a request to use data.

» Should the governance of the lists have a legal basis (who has the right to maintain the lists and accept/deny requests etc)?

» The lists should be governed by a body that has a legal basis, even though the lists themselves might not be part of legislation.

» We need to balance the need for the patient to know vs. the need to allocate the responsibility to a governing body to do this work for the patient.

» A decision-making body should have authority given to it either in law or via other instrument that makes its decisions binding and for which it is accountable.

» We should not seek to establish a new European authority. It is unlikely all Member States would delegate to such a body. We should be able to depend on countries to implement this. Most of the data reuse requests would have to pass via national data governance, but for these data committees lists would help with defining what bona fide research is and what it is not. Lists could be maintained by a European (not for profit) institution. We should start with guidelines rather than hard law.

» oResearch Data Scotland has shown that a lot of health data has been collected. If you want to use this data effectively, we should focus on what you should do (emphasise the positive side, more than what you shouldn't do). An outcome of today should be to promote what a company could do to make a difference to citizen lives.

- There was consensus that having a list that was used to illustrate the (health beneficial) purposes that would generally be supported by decision making bodies, and a list of those purposes that would not, would be reassuring to the public. This would be transparency of intention, and over time transparency about decisions made could be added.
- Further discussion is needed about whether decision making bodies operate at national or a European level.

Break-out 2: *rapporteur Nathan Lea*

- The group had discussed the challenges for anonymisation/pseudonymisation.
 - » Is pseudonymisation and anonymisation really the main issue here? Is it realistic to consider any data to be truly anonymised?
 - » It may be more important to focus on transparency of how data is used and how concerns about this can be addressed.
 - » The main risk of concern is whether citizens can be identified, either when their data is handed over to others or when they're empowered to share their data themselves.
 - » There is a continuous risk spectrum – much depends on use and context. It is difficult to hard code what is acceptable or not acceptable.
 - » Anonymous data has uses but it is limited for the work we are now undertaking, so we need more detailed data that could be considered identifiable.
 - » We need to have a transparent discussion with citizens and other bodies, about risk stratification, to see what would be acceptable going forward.
 - » Who should decide if particular data is anonymous or pseudonymous?
 - Rather than the regular players in this space, it was felt that this should be an organisation that is independent from the processing of the data, and that would prioritise the concerns of the data subjects and citizens.

» GDPR was drafted to develop common rules for this, in the Digital Single Market – is that still possible?

» When a citizen uses digital tools, it is a bit like the wild west – trading privacy for convenience. Can we empower citizens to negotiate about the use of their own data, protecting their own interests as they see them? If they have more control, the data portability requirements are also better met.

Rules on authorisations and access

» The end game should be that the citizens control their data, or that a controller does this on their behalf, with a full audit trail. There would have to be exceptions to this control, for example not being able to exercise the “right to be forgotten” if a person has started to participate in a clinical trial.

A lot of these complex areas are already covered by GDPR, but we have differing interpretations across Member States. Can we harmonise rule making and interpretation e.g. on lawful basis: do we all need to follow consent for processing, or are there other bases? This is happening nationally, causing irregularities, but can could be determined on a macro EU scale.

AI is prime example where we need to be more transparent about how data is being used, not only to citizens but also to regulators. There is a need to explain what is happening during AI development in order to be able to assure it. This is an area we need to do more on.

(The breakout group have decided to have an additional discussion on this topic.)

Wrap up and answers

- » Architectures for these data infrastructures is the enabler, but we have to be sure the technology supports the needs we have discussed. The technology needs to be malleable.
 - Ethical obligatory duty on those dealing with sensitive personal data, to ensure they provide clean, codified, real time data to support the care of individuals and others.
 - Need a more harmonised data protection regime in Europe (how interpret and regulate)
- » Potential is there to for AI to revolutionise medicine, but we need to find balance with protecting privacy regarding sensitive data.
- » Make things possible – we must not tie ourselves down within the current status quo and must bring citizens into a more

transparent discussion. We also need to be more accurate on how we define “fair” uses of data, including the quality of the data being used.

Open discussion

The way pseudonymisation is handled by the GDPR is a missed opportunity. There could have been proper rules for pseudonymisation, and maybe even certification rules for those who undertake this, so that it could be treated as anonymised. It was questioned whether this distinction is so very important: it would be better to evaluate reidentification through a risk management framework, which takes into account the context in which the data is being used. Longitudinal data is often important, which makes the data quite distinctive even if there is no pseudonymisation key.

- » It was suggested that consent or using some type of privacy preserving analytical method are the only two options that can be adopted. Consent is fine for deep dive research studies e.g. adding genomics to already existing clinical data. For other for

broader epidemiological analysis the only solution would be a privacy-preserving analytical method. This can include distributed machine learning, using advanced data approaches (where the data remains with the custodian). This has been used, for example, in IMI Melody.

» The group had also discussed analytical and synthetic data – allowing analysis of a population at an intermediate stage of the research, to gain insights before going on to using the real data.

» GDPR requires a Data Protection Impact Assessment, which could be linked a risk stratification framework. Its disadvantage is that it is not cast in stone, so you need to have regular update of that risk framework and monitoring it. Responsible organisations should be able to do this amongst themselves, and get a suitable approach recognised by data protection authorities.

» GDPR does not easily accommodate technology advances like privacy preserving analytics. Distributed analytics raises new competition concerns – a new set of problems that don't belong to GDPR but perhaps to competition law.

» Synthetic data (meaning original data that has had noise added to it) and distributed analytics will become more often used. We need to keep track of this emerging field and encourage knowledge to be shared.

» Synthetic data is not useful for making decisions about individuals but can be very useful for studying cohorts. The challenge with synthetic data is how much noise has to be added and how that impacts on the accuracy of the analyses that are performed.

» Synthetic data raises fitness for purpose and risk questions, which might usefully be the subject of future European Commission call topics.

Break-out 3: rapporteur Bleddyn Rees

The group had focused discussion on the B2G data sharing principles for the public interest, and added to them, recognising that all of the principles need to have more health specifics added.

Demonstrating health or societal value

» Definition of societal benefit: about a common good and won't always cover the whole of society but rather segments of society.

» Information needs to be provided at the beginning and throughout a study, not just at the end. If a study fails, the lessons learned can be as important as successes.

» Research vs. innovation: all are valuable - the value depends on the kind of organisation conducting the data use.

Earning trust and how incentives work

» Topics touched on were if incentivizing can easily be misunderstood. Considering the example of AI enabling blind people to be able to "see". This may require some rule changing. This is called 'reasonable accommodation', which could be adapted for health data use (and also including safeguards against the misuse of data).

» This is a legal concept that is embedded in the European Framework Directive for equal treatment in the workplace. Disabled people should expect reasonable accommodation for their needs, such as equipment to enable them to participate in workplace activities e.g. meetings. This concept may be transferred to the data environment, by ensuring that vulnerable people are included in decision making about data. Everybody should be included, and that should enable all of society to feel confident in responsible health data sharing.

» Rules should not be used to disadvantage people. Inclusivity should cover both the benefits from the use of data (e.g. being able to use real time feedback systems because they have accessibility features) and being sure to include all people within the data to ensure scientific validity and applicability of the results.

» How then can dynamic consent also be delivered in an inclusive way, including ensuring vulnerable people being appropriately informed?

» In the physical world, legislation based on rights, such as non-discrimination, can be difficult to transfer to digital space and the internal market. Some legislation such as the web accessibility Directive is internal market based: a rights-based route is hard to deliver. But is there an economic argument that can be made in favour of reasonable accommodation in the digital world. The scientific rationale mentioned above is directly applicable. An economic argument might take time to formulate.

» A market argument usually favours delivering to the majority of customers: for marginal features the costs often outweigh the customer revenue. European values do recognise the importance of inclusivity, but it is a difficult legal basis since accessibility features are usually expensive. However, Europe is a large market and even smaller subgroup needs can scale up. Adopting standards can reduce the costs, which unlocks the economic value.

» The GDPR focuses mostly on permissions (e.g. having a legal basis) but the discussions today have really emphasised the importance of transparency. The communication of purposes and benefits to the public has to be inclusive. This is not as much a GDPR concern.

Legislation and guidelines

- » Digital contracts and codes of conduct
 - Much prefer the name compact (meaning commitment of various stakeholders to each other, to whatever rules and criteria that have been established).
 - This term may help by changing the language and would equate to adhering to an agreement on a voluntary basis rather than through legislation.


» The group had recognised the value of real time consent, though were concerned not to overburden on citizens. However, does it have value by making consent easier?

» COVID-19 provides an important opportunity to develop a framework at a time when this use case is an ongoing priority


» – presents interesting opportunities going forward.

Discussion on a code of conduct to complement the European Health Data Space

- » Note: EC participants had technical issues at this point in the meeting and will contribute afterwards to the written report.
- » EFPIA is working on a code of conduct relating to the GDPR. This is starting with Randomised Clinical Trials: the legal basis, data controllers and processors. Should this be reframed as a compact? It would make more sense to have a multi-stakeholder voluntary code rather than a pharma only developed code – for pharma they would get patients involved which is more likely to allow access to data and build trust.
- » One needs to distinguish authorship from applicability: multiple stakeholders should be involved in agreeing a code (or compact) and the resulting code should be applicable to, and adopted, by, all of the data user stakeholders (not just by pharma, e.g. also by MedTech).
- » Patient organisations are increasingly managing data so they could also abide by a code that is commonly agreed.
- » We should not foster silo codes. Data use is a value chain, and silo codes will inevitably leave gaps.

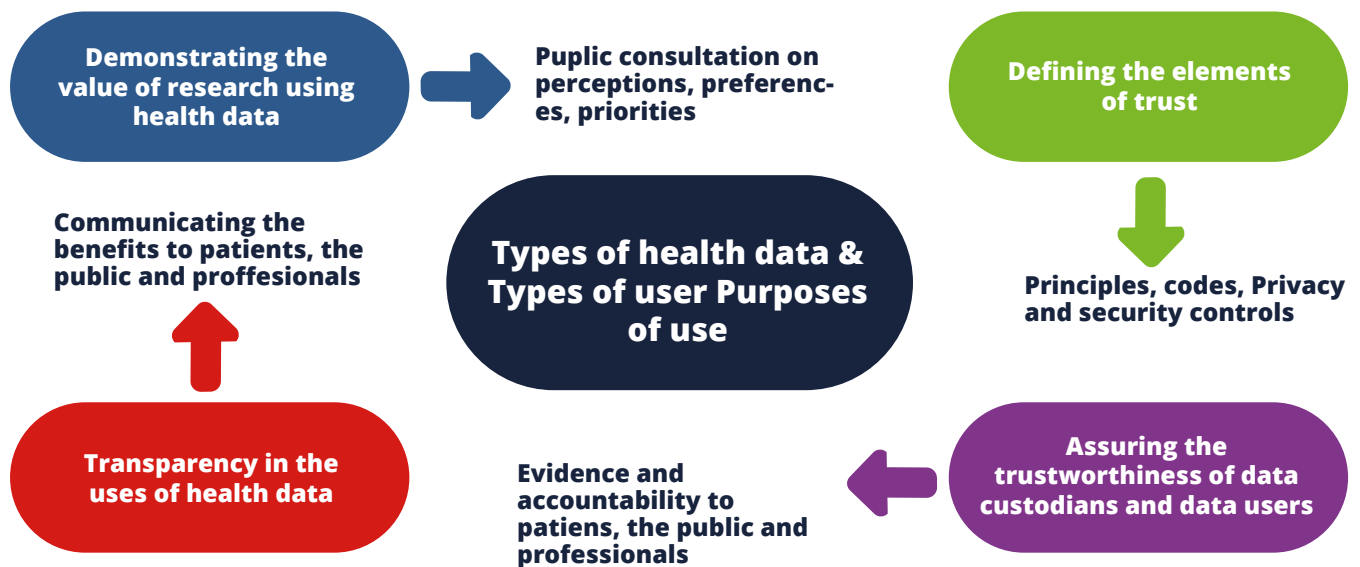


Data literacy is very important – to access, use and understand the data, and transparency explanations. We need to raise the importance of this in the code. COVID-19 has given the opportunity for all to see how citizen involvement can be an enormous opportunity.



Concluding discussion

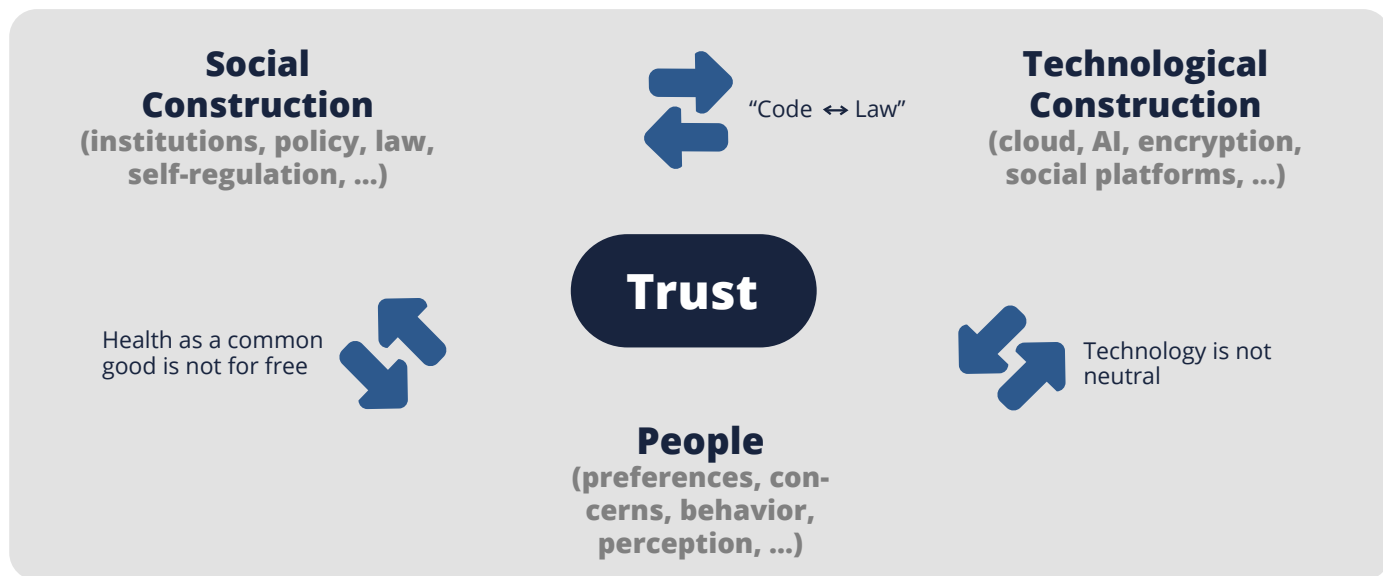
There was discussion about the diagram below which presents four interconnected dimensions of trust.



Paul Timmers reflected that it is usually fruitful to think about the interplay between what we construct by means of technology and our social constructions (laws, codes of conduct, organisations, behaviour): a technological and a sociological construction of reality – each side may treat the other side as malleable to accommodate their own perspective. For example, the previously discussed lists (whitelist and blacklist) are a tool, a technological construct: it may even one day be managed automatically,

as we do today by detecting improper content on social media. On the other hand, codes of conduct do not have a good reputation unless they are accompanied by transparency and accountability. These two sides can each enable the other to be successful. Law should not be made without thinking about how technical constructs can help enable compliance and enforcement. Technology initiatives should not be developed without considering how and by whom it will be governed.

It was suggested that a third, psycho-social (people oriented), dimension is important as well, including a cultural shift. Paul kindly provided afterwards the slide below that depicts these three perspectives.



APPENDIX 1:

Round Table Agenda

13.00	Connection hassles (They always happen! - but please connect at 13.00)
13.05	Welcome, virtual meeting logistics
13.10	Scene setting <ul style="list-style-type: none">• the context for this meeting, opportunities and problem statement• meeting objectives• brief clarification and scoping questions
13.30	Breakout group discussions on acceptance criteria Participants transfer to their assigned virtual breakout rooms 1: The who, what and why of primary & secondary data use (moderator Dipak Kalra) 2. Technical and organisational safeguards (moderator Paul Timmers) 3. Transparency about data use and value (moderator Bleddyn Rees)
14.15 - 14 25	Comfort break Followed by 15 minutes, still in groups, to consolidate the acceptance criteria and propose how to implement them
14.40	Feedback from the breakouts: three categories of acceptance criteria Per group: 5 minutes reporting, 5 minutes open responses and suggestions Consolidation of the main criteria
15.10	Plenary discussion moderated by Dipak Kalra on how to operationalise the criteria: Level 1: European Union (e.g. code of conduct as one of the instruments) Level 2: Member States Level 3: Public & Private Sector
16.10	Next steps, closing remarks
16.30	Close

APPENDIX 1:

Participant list

Name	Breakout	Organisation
Prof Mahmoud Adil	1	Public Health Scotland
Nicola Bedlington	3	DataSavesLives
Leonardo Calini	3	Microsoft
Karla Childers	3	Johnson & Johnson
Maria Christofidou	1	i~HD
Prof George Crooks	2	CEO Digital Health & Care Institute Scotland
Carina Dantas	3	ECHAlliance & Digital Health Europe
Ander Elustondo Jauregui	2	DG Sante Unit B3
Cornelia Kutterer	3	Microsoft
Rachel Dunscombe	2	Tektology Ltd
Ioana-Maria Gligor	3	DG Sante
Tomas Gornik	1	Better Care Health IT expert, co-chair of openEHR
Dipak Kalra	1	i~HD
Zoi Kolitsi	1	Digital Health Europe & I~HD
Nathan Lea	2	i~HD
Licinio Kustra- Mano	1	DG Sante
Angel Martin	3	J&J
Prof Catherine Paugam—Burtz	3	Director General Assistance Publique-Hopitaux de Paris
Bleddyn Rees	3	Digital Health Society & ECHAlliance
Hannu Hamalainen	3	SITRA
Michael Strübin	2	MedTechEurope
Louisa Stüwe	1	The French Health Data Hub
Paul Timmers	2	University of Oxford
Yiannos Tolia	2	DG Sante
Bart Vannieuwenhuysse	1	European Federation of Pharmaceutical Industries & Associations
Veronica Zilli	1	Johnson & Johnson
Invitees who could not attend but have contributed to the report		
Elena Bonfiglioli		Microsoft
Michele Calabro		European Patients' Forum
Brendan Barnes		European Federation of Pharmaceutical Industries & Associations

REPORT OF ROUND TABLE
VIRTUAL MEETING ON

**A recipe for trustworthy
digital health:
standards, architecture
and value**

Round Table Summary

This report summarises the topics, discussions and conclusions of a multi-stakeholder Round Table held on Friday 30th October 2020 on what is needed from interoperability, quality, standards and architectures to deliver the best value from trustworthy digital health and how to accelerate their adoption and use.

27 participants, comprising patient organisations, healthcare providers, payers, ministries, data protection authorities, regulators, industry and industry associations and representatives from the European Commission participated in a highly interactive half-day meeting designed and run by the Digital Health Society and the European Institute for Innovation through Health Data, sponsored by Microsoft and Johnson & Johnson. The Round Table sought to identify the most important positive disruptions and the necessary actions that will enable us to make better use of health data, and especially how these may achieve value to all from the European Health Data Space (EHDS). This event was timed to provide input to the scope, design and governance framework being developed for the EHDS.

Opening plenary presentations

The meeting started with plenary scene setting presentations.

Dipak Kalra suggested that, after decades invested in standards development, many patients are yet to see evidence of joined up health information to support their care. Substantial effort is also still needed to combine multiple data sources to answer care pathway optimisation, outcomes improvement, public health strategy and research questions.

Ceri Thompson presented the European Commission's European EHR Exchange Format which may now help to accelerate standards adoption through a portfolio of recommendations on interoperability principles, priority categories of health record data to exchange and proposed standards and profiles to communicate this across borders. This initiative, now being actioned by the x-eHealth project, is involving a wide range of stakeholders with a strong emphasis on patients and clinicians, on transparent prioritisation and ensuring robust data protection.

Nigel Hughes explained that the IMI EHDEN project is establishing the largest European network of data sources to support federated research, via distributed analytics. After only two years it has certified an impressive number of small-to-medium sized enterprises (SMEs,

currently 26 in 16 countries) to enable the mapping of data sources (62, in 14 countries) to the OMOP common data model and is in parallel growing a network of data partners to join the federation.

Jesper Kjær showed how European regulators are also proactively growing their capability to leverage real-world data, through the DARWIN initiative, to complement RCT data for pre-approval and post-approval decision-making, better pharmacovigilance and preparing for personalised medicine. He predicted that data generated by patients and citizens may become the largest proportion of the real-world data used for decision-making.

Summary of the group and plenary discussion points

The adoption of federated and hybrid architectures

Most of the “big data” networks in Europe, such as EHDS and the Multiple Sclerosis Data Network, have adopted a federated architecture, using distributed queries communicated to a network of data repositories that only share query results but no subject-level data.

These and other real-world data and cross-border eHealth initiatives may in the future become interconnected through the European Health Data Space, which could provide a unified portal to discover and interact with these data resources meaningfully and near to seamlessly.

Distributed analytics requires new skills to better formulate research questions. We still need to learn more about what kinds of questions

can be answered by distributed analytics, and which ones need to work on a dedicated patient level data extract.

Training is especially important for composing distributed and big data queries, so these are well formulated (e.g. generate reproducible results) and properly address the insight-generation needs of public health and research.

In practice we are likely to need a hybrid approach to analysing data sources, most federated but some needing to be individual databases or fine grained data set extracts.

The EHDS might be a mechanism for implementing and exemplifying this hybrid approach.

This could only work at a European level if our trust and trustworthiness also operate at a European level, for example by having a unified governance framework so that citizen's rights apply and are enforceable evenly and transparently across borders.

Data quality improvement is vital.

Data quality assessment and improvement should be encouraged through benchmarking and disseminating good practice. This requires investment, and a business model for that.

Pressure should be put on the ICT sector (EHR vendors, new app and wearable vendors) to embed data quality support in their products.

Health data must be labelled in a standardised way with provenance metadata so everybody can trust its origins and safely interpret data from its original context. This metadata must be efficiently incorporated automatically by EHR and PHR systems, not by adding to the data entry burden.

A transformational approach to health data, as a pooled resource

The ambitions we now have from data indicate the need for a new kind of disruption: for us to consolidate health data in shared health data pools or banks, separated not only from single ICT companies but also from single healthcare providers.

Large scale (regional, national) health data pools should provide real-time continuity of data access for self-care by individuals, healthcare delivery and population level analysis.

The custodianship and governance of future health data pools must uphold the perspectives of citizens and patients, reflect common European values and principles, be independent of the EHR system vendors, of the companies that collect data via apps and sensors and of the companies that make use of data.

Interoperability standards have conventionally focused on connecting heterogeneous vendor specific systems to enable data flows to follow the patient between providers, or within a health data network. This pooled data approach requires common data models for data at rest.

A transformation towards cross-organisational, and perhaps independently run, health data pools will require radical change in ICT products and procurement rules.

Data must deliver value, as a common societal good

We should consider pooled and reusable health data as a societal good. We must ensure that the uses of health data, for example through big data networks and pooled data, deliver societal value.

Our focus for that value should not only be on developing innovations but on improving health outcomes, in the immediate future (i.e. today, not tomorrow).

The value propositions must ensure there are equitable benefits for all.

Uses of health data should be considered more like the example of higher education: society strongly supports investment in education, but does not expect to see or judge the short term financial return on that investment, but instead regards it as an investment for society as a whole.

Regulators and public health agencies must lead by example to demonstrate how to make best use of real world data for reliable evidence generation and decision making.

We need to improve the digital literacy and the data literacy of all stakeholders so that they become competent and fluent data creators and users.

All data creators must better appreciate the importance of the data they create, including its completeness and the quality of the data.

Data literacy for health professionals should be included within training and continuing professional development programmes.

Patient and citizen generated data will be an important complement to professionally generated data in the future. As personal computing power grows, patients may in the future be the primary holder of their complete EHR. We must make substantial investments in a pan-European scaling up in education for citizens, patients, families and carers. This education must convey transparently both the benefits that data

can bring to society as well as the risks that have to be mitigated through data protection. We must acknowledge that cultural differences in attitude exist across Europe including varying trust in governmental agencies that might be defining and applying data access and use rules.

Digital health literacy and tools are also needed to support people in making more use of their own data themselves, in self-care and prevention e.g. enabling comparisons of their status and progress with an anonymised pool of similar people. We should facilitate communities such as neighbourhoods collaborating on health promotion through their data, possibly supported by health insurers and regional health authorities.

The EHDS is a potential enabler for new positive disruptions, persuading multiple stakeholders to move together

Changing the perspectives and actions of one stakeholder group, when the stakeholders that connect with them do not change, does not lead to much improvement. This tends only to happen through disruption. COVID-19 has been one such disruption, which has catalysed change but also shown us how unprepared we are to collaborate on generating and sharing intelligence at scale.

The EHDS could now act as a catalyst and forum for agreeing common European ethical and governance principles, quality and interoperability requirements.

In conclusion

The Round Table presentations and discussions as a whole proposed a transformational role in the way we perceive and handle health data, and where data originates, where it resides and who controls who can use the data in order to deliver value. It was generally felt that the future lies in enabling wider multi-stakeholder uses of harmonised and pooled, meta-data-labelled and quality benchmarked, health data as a societal good.

Introduction

This report summarises the topics, discussions and conclusions of a multi-stakeholder Round Table held on Friday 30th October 2020 on what is needed from interoperability, standards and architectures to deliver the best value from trustworthy digital health, and how to accelerate their adoption. Its aim was to identify some of the most important actions and disruptions that are now needed to enable us to make better use of the growing number of health data sources across Europe, and especially how these may achieve best value to all from the European Health Data Space (EHDS). The COVID-19 pandemic has helped to highlight the importance of having high-quality and interoperable data, that can be combined at a European scale quickly in order to inform public-health strategy and decision-making. However, the difficulties we faced with achieving that large-scale intelligence gathering now challenges us all to rapidly transform the data landscape for health.

This Round Table and report have been timed and offered as inputs to the scope, design and governance framework being developed for the EHDS. This important new initiative, which builds on the Commission's Recommendations last year on the European EHR exchange format (EEHRxF), has the potential to provide a European scale portal to many data resources - provided that their format and governance are compatible. This initiative is also recognised as a catalyst for bringing together big data stakeholders across Europe, to foster alignment be-

tween them across their own initiatives as well.

The Round Table was an invitation-only, multi-stakeholder and highly interactive half-day online event with 27 participants, dividing for some of the time into three virtual breakout rooms for deep dive topics. The agenda is given in Appendix 1. The participants included patient organisations, healthcare providers, payers, ministries, data protection authorities, industry and industry associations and representatives from the European Commission who are promoting the EEHRxF and architecting the European Health Data Space. The list of meeting participants is given in Appendix 2.

The event was jointly run by the Digital Health Society (represented by Bleddyn Rees) and the European Institute for Innovation through Health Data (represented by Dipak Kalra). It built on the Digital Health Society's Summit in Helsinki with the Finnish Presidency last December when both organisations collaborated on the data and digital content.

It was sponsored by Microsoft and Johnson and Johnson, who contributed financially for preparing and running the event. The companies did not control the structure, hosting, content or reporting of the event. Individual experts from those companies contributed as meeting participants and are listed in Appendix 2.

OPENING PLENARY SESSION:

Scene setting

Bleddyn Rees and Dipak Kalra welcomed participants and introduced the plenary speakers and moderators for the breakout groups.

Bleddyn set the scene for this Round Table, the context for which is the European Health Data Space. It offers the opportunity to bring together experts from different areas to consider what solutions now need to be strongly promoted, and to formulate calls to action.

This meeting builds on prior related events over the past year, starting with a Digital Health Society Summit in December 2019 that highlighted many of the issues and challenges that impact on how the public and patients understand and indicate preferences for, or control over, the uses made of health data. In the spring of this year DG Santé ran a series of consultation workshops with participation from national Data Protection Authorities, moderated by Petra Wilson. In May, the DigitalHealthEurope project ran a virtual focus group for industry about company aspirations and potential contributions to the European Health Data Space (EHDS). The industry participants highlighted the special opportunity for architects of the EHDS to develop a coherent governance framework that could be adopted by other European data initiatives, thereby helping to harmonise approaches adopted across Europe. The first sponsored Round Table was held in September 2020 and was designed to contribute to that aspiration by taking a deeper dive on societal acceptance factors for data reuse

that might be taken on board when developing the EHDS governance framework. A subsequent event was recently held by DigitalHealthEurope on the perspectives of patient organisation representatives on this topic. These events had some organisers and participants in common and have shared their outputs so that their progression could be complementary and additive. This second sponsored Round Table examined more socio-technical topics: how to disrupt and thereby accelerate the adoption of interoperability standards to enable data sharing and combined data analysis, the adoption and interconnection of centralised and distributed data architectures and infrastructures, and how to ensure that we have data of sufficient quality to make trustworthy inferences so that decision makers can better trust real world data insights.

Dipak explained that objectives of this Round Table were to determine the most important data-related success factors, and the calls to action which now need to be championed in order to accelerate that success. However, he cautioned the audience that decades have gone by with substantial efforts and investments in interoperability standards development, but still patients and the public see limited evidence of joined up health information to support their care. Substantial effort is still needed to combine multiple data sources to answer care pathway optimisation, outcomes


improvement, public health strategy and research questions. From his experience working in standardisation bodies, it seems that we have been successful at engaging a lot of “technology geeks” to develop technological standards such as modelling, terminology, ontology & workflow representations, API specifications. However, we have poor at engaging the people who actually create and use health information. Standards bodies only manage to engage token technically aware clinicians, but very few grassroots clinicians, few patients, few public health experts and not many from the clinical research sector in helping to ensure that our standards prioritise and deliver what they most urgently need to collect, share and analyse. Maybe it is time for us to redefine which stakeholders we position to most strongly determine the priorities for interoperability, connectivity and collaboration.

Our health landscape today is filled with perverse incentives that do not favour interoperability and the sharing of data. The biggest concern of the end users of health ICT systems, the clinicians, is to cope with their workload, and they desperately need their EHR system to help them with this workload, and not be a data entry burden. This leads to a compromise between the complete picture of the patient and what is actually recorded. It would be nice if they could instead have more cognitive capacity and time - from better-designed systems - to consider how best to leverage their EHR to support patient care and outcomes. In contrast, providers and the purchasers of health ICT solutions are jointly focused on how to optimise business efficiency. Their concern is how the EHR can support them with workflows and resource utilisation, providing management reports and dashboards to help run


the healthcare organisation most efficiently including optimising reimbursement. These are the functions of the system that the purchasers focus on, and the vendors therefore provide. Then, if we look at the health and care funders, they want to determine how they can optimise investments and get the best value, and they want healthcare providers, and therefore the systems of the healthcare providers, to deliver that information. If we want to change these perverse incentives, we need to turn the financial and reward levers to favour care collaboration, and require systems that collect data of societal value and help track and improve health outcomes, which usually depend upon care collaboration. This is, in Dipak’s view, what we need disruptions to achieve.

Ceri Thompson, Deputy Head of the eHealth, Wellbeing and Ageing Unit of DG Connect, presented the Recommendation on a European EHR Exchange Format, which was formally adopted in February 2020. This is a major milestone, a building block to enable the sharing of health records securely between Member States, towards realising the 2018 EC Communication on the Digital Transformation of Health and Care. This Communication laid out three objectives:

- Secure access and exchange of data across the EU
- Health data pooled for research and personalized medicine
- Digital tools and data for citizen empowerment and person-centred healthcare



The ability for citizens to access their own health and care data still varies greatly across the EU. Member States still have internal interoperability challenges with sharing data between their providers as well as between countries.



The Recommendation seeks to facilitate cross border interoperability of health and care. It sets out an overarching framework for the further development of a European EHR exchange format, with three main components:

1. A set of principles governing the access to and exchange of EHRs across borders
2. Common technical specifications for the cross-border exchange of data
3. A joint coordination process for the development of the European EHR exchange format to ensure that there is involvement of the wider stakeholders

The strategy was a deliberate response to concerns that technical specifications could otherwise arise in a non-transparent way. The value of access to health records has especially been promoted to citizens, health professionals and public administrators, across European countries, for example if patients need referral to specialists across borders or live near a border. There are several important guiding principles for the exchange of health records at a European level. These include that the design should be centred on patients, that the EHR data should be machine readable, com-

municated with adequate protection, audit and strong systems for regulating access. Trust is vital, given that this is health data and we need the highest standards for security. The EC have recommended the creation of a National Digital Health Network in each country, to oversee the introduction of these often-novel areas of organizational interoperability.

The Recommendation proposes a set of common technical specifications (as a baseline). They specify an initial set of health information domains: patient summaries, ePrescriptions, lab reports, medical images/reports. It lists the recommended interoperability specifications (based on standards and profiles that are already in use). It has advocated an incremental and selective approach for adopting, refining and maintaining the specifications as the European EHR exchange format.

A Joint Coordination Process includes the Member States (MS), European Commission, wider stakeholders including HealthTech to support this process. This includes supporting common adoption approaches whilst recognizing that different countries will be able to proceed at different speeds. The Commission can support this at a European level through its research, innovation and deployment actions. A Support Action project X-eHealth is bringing together Member States, clinicians and ICT specialists to further develop the draft technical specifications. A survey is being conducted to determine the baseline of where countries are now in the use of the specifications and will monitor how these are taken up in the next few years.

Nigel Hughes the Project Lead from Janssen, presented the federated architecture scale up programme of the IMI European Health Data & Evidence Network (EHDEN), which runs from 2018-2024 as a project, but as a long-term European infrastructure.


EHDEN aspires to be the trusted observational research ecosystem to enable better health decisions, outcomes and care. Its mission is to provide a new paradigm for the discovery and analysis of health data in Europe by building a large-scale, federated network of data sources standardized to a common data model, OMOP.

In effect it is building a network of “railway tracks” on which different healthcare and research users can run their analyses on a large collection of data sources. A key challenge it is tackling is harmonising data from technical and process standpoints, using OMOP. EHDEN has so far had two calls for SMEs to learn how to map data to OMOP and become certified. There is also an open call for data partners, who can qualify for a grant of up to €100,000 to harmonise their data and make it available on the network, according to a code of conduct. Data partners may themselves be users of the network to run studies. The EHDEN Academy is free to join, and offers e-learning resources

about EHDEN, the OMOP common data model and how to use the available tools to conduct federated query studies. Nigel outlined the EHDEN consortium and work plan structure which comprises three main pillars. The technical community now includes 26 SMEs in 14 countries who have been certified, and a growing list of data partners, currently 62, whose data can be discovered via a catalogue, which will be public in 2021. A recent third call for data partners has attracted 34 applications. There is a workstream pillar dealing with tools development, the portal, dashboards and security measures. Nigel illustrated the dashboard and catalogue, which aim to greatly reduce the effort to discover and determine the suitability of data sources to a study. Several research use cases have been pursued including on COVID-19. The EHDEN Academy will help to sustain this initiative post project, and there are open conversations with the EC about how EHDEN can connect with the EHDS, and with the EMA on an interface with EU DARWIN.

The COVID-19 study-a-thon demonstrated how processes within a research workflow that might normally take a few months each, and take two years altogether, could be compressed to a matter of a few days. It was possible to work efficiently with data sources already mapped to OMOP from all over the globe, to perform phenotyping

and characterisation studies, to examine variations in COVID-19 presentation, the use of drugs and prediction models. Much of this work has already been published and is starting to be used by public health authorities.




Speed can be greatly increased but quality also preserved. Future research in collaboration between EHDEN Partners and EMA will examine the effectiveness and safety of the new COVID-19 treatments and vaccines.

Whilst having started with healthcare provider data as the “low hanging fruit”, EHDEN hopes eventually to extend to include citizen generated data.

EHDEN's strategy seems to be in line with that of the EC for the EHDS, and EMA for EU DARWIN, and it is proving to be a thought leader in this dynamic European environment.

Jesper Kjær, from the Danish Medicines Agency (DMA), spoke to the question “What do decision makers need in order to trust real world evidence?”.



Jesper started by introducing the ambition of the DMA's new Data Analytics Centre (DAC) to increase the accessibility of safe and effective medicines and medical devices through the use of clinical trial and RWD, and advanced analytic methods. It will generate quantitative data insights and support a better regulatory framework for scientific advice, pre-approval and post-approval decisions, evidence for the EMA PRIME priority medicines scheme and to prepare the DMA for precision medicine.



The DAC will connect real world data silos, from the Danish Health Data Authority, industry trial CDISC patient-level data, pharmacovigilance data and other data sources, using the National Genome Centre supercomputing resource. Jesper outlined the DAC governance structure, comprising an External Steering Committee and Clusters of Excellence across the Medicines Agencies. Jesper drew the audience attention to a recently published article in *Clinical Pharmacology & Therapeutics*, on “Randomized

controlled trials versus real-world evidence: neither magic nor myth". Its punchline message is that RCT and RWE need to be combined.

He explained that the EMA HMA Big Data Steering Group has now produced a work plan that includes topics such as data quality, discoverability and networked analysis. These tasks will define what regulators need in terms of data quality and what is needed for decision making. He noted the complementarity with EHDEN. The DAC and EFPIA have established a data science forum, to discuss data quality, lineage, representativeness and metadata including FAIR data objects. Recent experiences of COVID-19 have highlighted the impact of different testing strategies on the data and how it can be used. The HL7 Project Vulcan is now exploring how to connect HL7 FHIR and CDISC to support regulatory use of RWD. New technologies such as blockchain will also help to evidence lineage and provide traceability. Jesper drew attention to the retracted Lancet hydroxychloroquine paper that showed how important data transparency is.

Jesper offered, as an audience provocation, that industry has traditionally seen RCT data in CDISC form as the source that overarches everything else (this is where the funding has gone). Healthcare sees the EHR as its predominant data source and sees data in CDISC/RCT form is just a fragment. The reality may be neither, but with the data from outside of the healthcare system (from citizens such as activity trackers, sensors, grocery shopping etc.) being by far the largest source of RWD, all of which deal with and involve health data.

The meeting then divided into three breakout groups:

- A. Improving the (re)usability of data through standards**
- B. Architectures enabling large scale, secure and timely data access**
- C. Sustaining data sharing and access by demonstrating value and trustworthy decision-making**

The briefing text, discussions and plenary feedback of each are summarised next.

BREAKOUT A:

Improving the (re)usability of data through standards

Moderators Tomaž Gornik and Dipak Kalra

Briefing text circulated in advance

Interoperable health data, ideally captured at source into systems that incorporate standards-based representations, is a critical success factor for information use and reuse. Interoperability standards have been available since the late 1980s, and their number has grown massively. However, their adoption remains piecemeal across vendors, systems, providers, countries and data types. It is often said by new entrants to the field, such as startups, that the health information standards landscape is confusing, that there seem to be multiple standards covering the same purpose and that the standards that we have don't seem to be capable of easy concurrent use. (They do not fit together well.) Do the Refined EIF and the more recent EEHRxF offer us the answer? This breakout will start by exploring whether we have invested over the past few decades in the right kind of structural and semantic standards. Where are we strong and what are the gaps?

A new information architecture is key to unlocking the power of digital technologies and creating the connected health ecosystem of tomorrow. Today's solutions tightly couple data to applications. As health and care data is for life it needs to outlive applications so there is a clear need to separate the two. Architectures

for the future will have a vendor-neutral data layer at the centre, so data can be used seamlessly by all apps, applications and algorithms. To bridge the gap between current and future state, standard APIs will allow legacy EHRs and new open platform-based systems to coexist, enabling innovation during the transition. What needs to be done to accelerate the shift to open data platforms?

Despite having generic information models and very large terminology systems, there is a layer in the middle that has not seen sufficient promotion and investment: building communities of practice to define core clinical data sets, clinical models, value lists or other practical representations that could shape consistent clinical documentation practice, focus data sharing on relevance and enable the more precise targeting of clinical decision support and analysis queries. Why have we not yet invested in growing such communities of practice: clinical data standards makers and champions?

Patient generated data (apps, sensors, citizen-controlled data clouds) is the next exciting growth area. This new ecosystem brings in non-healthcare players, some of who may bring a different approach to standards and adoption from other sectors. Do we have a satisfactory portfolio of standards to represent what patients might capture, for healthcare, prevention and wellness management? This community of developers seems at times to be even less motivated than EHR vendors to adopt standards, possibly less aware of them.

There is no strong incentive for them to build standards into their developments, to invest in those skills and no high-level governance of the personal health ecosystem. How can we best enable the professional generated and patient generated data ecosystems to co-operate?

One of the arguments for why standards have been so poorly adopted across Europe is the lack of business models for EHR system vendors, and possible also procurement rules that do not favour innovative (perhaps even experimental) procurements. Many health systems still reimburse today on the basis of the activities performed by single healthcare providers, which gives the providers no incentive for to procure standards except those needed internally (e.g. for PACS, labs). Vendors will do what their customers will pay for. What incentives could now accelerate EHR/PHR standards adoption?

Summary of the discussion

This breakout group focused on the concept of interoperability, what this means, and whether current interoperability standards will help us to achieve the usability and reusability of data. Has the time now come to put the data at the centre – concentrating on the data “at rest”?

Tomaž Gornik gave a scene setting presentation. He pointed out that we have focused our efforts in developing and promoting interoperability standards to enable the connectivity and transfer of data between different systems, as a contrast and a means to break up the data monopoly of single vendor products. However, in this model each application stores data differently, which is fine for that application, but we now better recognise the need to keep data for the lifetime of the patient if not longer, to access and use the data across multiple systems. This means we need to migrate the data between systems, via interoperability standards, which has costs and almost inevitably involves some data loss. The direction for the future, as presented by Gartner and EY for example, emphasises the importance of a shared common health data resource, which holds patient-level health data in a single form and with which multiple applications interface with to read and write data. When using these consolidated health data repositories, the institutions at provider and regional levels could release most of their data stores or, where this is not yet possible, they could agree on the same representation and governance, thereby contributing to a virtual personal health record for each patient. An example of this is the HIGHmed project in Germany, part of the national Medical Informatics Initiative, which has developed a single

generic and scalable platform architecture to which a variety of hospitals contribute data.

There was discussion about whether the data centric landscape portrayed by Gartner and EY is the design we should all now promote and follow. Several participants agreed that this is indeed the right approach. However, it was recognised it is challenging to introduce this kind of change without considering what this means for, and the impact on, health systems. There are substantial variations in the ways in which data are stored and consolidated today, and we therefore need to examine how we could transition to this new model and what the costs would be. Health systems tend to make incremental changes and any approach has to factor in the impact a change would have on the incentive chains of the key stakeholders. It may be helpful to consider foremost the clinical/patient benefits and to look for some immediate quick wins that this approach could deliver, as exemplified recently through COVID-19. It may then be possible to make evidence-based propositions to policymakers and to the European Commission.


Apart from transitional incremental steps, endorsing a data centric approach can be taken on board when making new procurement and major ICT upgrades. The EHDS might be a mechanism for implementing and exemplifying this approach.

The strongest driver here is not in the adoption of standards but in enabling access to information. The data and what can be done with data inevitably needs to the paradigm of “follow the money”. There was discussion about whether the concept of standardising data “at rest” (in a common data repository) is an irresistible value proposition, and how that would align with the possible federated approach of the EHDS? It was noted that standardising data at rest does


not mean that it is not federated: multiple standardised data repositories can be connected. The EHDS rightly puts data at the centre of our thinking, rather than the communication of data. It is also important to put the data to good use, not just through existing applications, so that the data does deliver value, for example to improve health system efficiency (i.e. following the money).

The risk is that as we bring data into a common environment but from very different contexts of capture. We may end up with a large mess without context, rather than today's distributed messes in context. We therefore have to make sure we look after metadata, standardise it, and that part of our investment in data literacy should help people to capture this metadata so that relevant data can be discovered and meaningfully interpreted. Users of data from a pool have to know the origin of the data and how they can trust it, including its quality. However, we cannot impose even greater documentation burdens on busy clinical staff, and therefore we need smart solutions to make their data capture and the capture of metadata to be even more efficient than it is today.

Data quality therefore also becomes an important factor, and it was felt that the appeal of data quality would increase if it became part of the international standards family.



Digital literacy is also important. We are constantly asking clinicians to collect better data, and we need to think about how we could automate that process more, especially the capture of metadata.



The MedTech sector is less directly affected, unless companies are themselves asked for data. The industry is waiting for standards like HL7 FHIR and IHE profiles to be more widely adopted commercially, which would enable them to avoid having to make their own data conversions.

It was suggested that the EC could incentivise the adoption of standards and the sharing of data though requirements placed on its funded projects. However, it was noted that the timelines of most EU projects are too short. Interoperability is a long and hard road that doesn't fit within those timelines, which is a weakness of the current EU funded programmes. However, the two levers of the EEHRxF and EHDS should be used.

The group saw huge opportunities for the EC to leverage the momentum behind the Exchange Format and the prospect of a European Health Data Space – which already emphasises the data. There is also a potential opportunity for EC projects to be incentivised to make better use of standards and to contribute their data in a standardised format. We recognised though, that their short time frame might make this challenging to deliver.

BREAKOUT B:

Enabling large scale and timely access through adapted architecture

Moderators *Nigel Hughes and Licinio Kustra-Mano*

Briefing text circulated in advance

- RWD is messy, We're drowning in data, but thirsting for knowledge, and the quality of RWD is extremely variable in terms of being able to answer research questions
- Underlying systems have to date in-optimally supported primary clinical care, and inadequately serviced secondary research use
- The need for insights and level of inquiry is rising exponentially, but most often the right data is not the right place at the right time to answer the right question (COVID-19)
- Socio-technical constructs, built around federated networks, that utilise the Internet, common data models, standardised analytics and advanced distributed machine learning are nascent, but rapidly expanding internationally
- Meanwhile, citizen rights and involvement, assurance of privacy and security, and ethical research conduct, need to be balanced with our societal research interests

- Data sufficiently open for research, but closed enough to protect citizens
- Could new socio-technical architectural approaches, utilising 21st century tools, e.g. federated networks, meet our needs in addressing our challenges, and/or will we see new ones we need to mitigate against?
- How can a Federated Data node facilitate a quid pro quo amongst Data Partners and Researchers?
- For Europe, can we create a European digital railway network enabling digital study trains, running on a standard gauge networks across borders to address multiple stakeholders needs?

The 'data supply chain' for clinical, primary use, and for research, secondary use suffer from archaic data capture systems, lack of conformity, consistency and continuity, and a reductionist general approach that captures the minimum, often inadequately for research purposes. Furthermore, structured data is often the low hanging fruit we work within health research due to challenges of curating semi and/or unstructured data, which often inhibits deeper enquiry. Technology innovation is impacted positively on this now in the 21st century, but we retain at best 20th century primary capture systems.

Whilst facing a challenge of making sense of human and machine language issues, syntactic and semantic harmonisation and the sheer effort of data curation, we are often needing to identify what questions we are able to ask with the data available. This is all within a socio-technical construct with parallel aspects of governance, privacy and security, ethical research conduct and ultimately balancing the rights of citizens and patients versus a societal need to understand health, from our biology to real world outcomes. If nothing else, COVID-19 has cast a lens focusing on all of these issues acutely, within the European region of multiple Member States, across borders, localities and institutions. The right data has not been in the right place at the right time to answer the right question.

National health infrastructures have tended to centralise the storage of data most needed for strategic decision making and surveillance, on the other hand, we are seeing more federated networks for clinical research, where data remains local, within countries (e.g. Germany, France) at a European level (EHDEN, EHDS, EU DARWIN) and for diseases (ERNS). What are the strengths and weaknesses of each architectural model, including from a governance and citizen control perspective? Is the future fully federated or hybrid?

Within federated networks, as per a Venn diagram, there are differing types of networks, with differing architectural designs and underlying assumptions about the source data, common data models, and standardised analytics, supporting rapid research, whilst ensuring quality is not impacted. In essence they embrace the local differences, while responding to them through harmonisation and interoperability steps.

Our ambition is to get closer to real-time data access, which we need for public health surveillance, the calibration and optimisation of algorithms and medical devices, for pharmacovigilance and for real time feedback loops to clinicians (e.g. current antibiotic resistance) and patients (e.g. AI alerts). Are the data architectures that we are growing capable of delivering real time data access, and of connecting to patients (for feedback) and not just to their data?

What is state of the art in ensuring data protection and information security, especially in federated architectures where data sources are being asked to open up a channel into their data for a wide range of purposes of parties that are not usually pre-specified and pre-agreed? Can these security solutions protect privacy even when the patient numbers are very small, such as in rare disease networks?

Federating the federations: if there are a growing number of regional, national, multinational, public funded and privately funded and public private partnership funded networks, are we at risk of creating a new wave of “distributed silos”, or can we obtain benefits from this regional interoperability and harmonisation effort? Meanwhile how do improve the underlying messiness of real world data?

Summary of the discussion

There is a lot of expensively generated data about each of us, largely created by other actors, which helps to form a picture of how each individual progresses in their illness and in response to treatments and interventions. Europe now has many great examples of secondary use, and secondary use infrastructures. It is challenging to consider how these could become integrated via the EHDS.

When considering what may be developed over the next 2 to 3 years in support of using health data for research, efforts should be focussed on the uses of data for clinical innovations that improve health outcomes, for today and not for tomorrow.

Research currently can take up to 17 years to deliver tangible benefits. Primary and secondary uses of data should not be so de-linked, and we should consider every patient as being an N=1 study. Everyone's experience is important. Technology is facilitating this capability, for example to enable the rapid validation of an AI model so it can be used in clinical practice. COVID-19 has positioned us at a technological inflection point, and we are now accelerating in our reuses of data.

We have seen many great examples of the secondary use of data, but how can we integrate these into the EHDS and into healthcare systems. Our joint focus should not only be on

innovations but on improving health outcomes today – not tomorrow, which matters to both healthcare and research (i.e. to primary and secondary use). These should not be decoupled, as they largely are now. We are at a technological inflection point as a result of COVID 19, which is accelerating the trends and the opportunities.

Public health seems to be overarching the primary and secondary uses of data: it spans research and healthcare. However, there is a need for more clear questions addressing more clear public health knowledge needs.

It is important for us to learn how to better formulate the questions. Too many projects are still looking at the answer side of the Question-Answer equation, based on the data that is available rather than the questions that need to be addressed. We need to upscale those working in public health, and increase the interoperability of the questions rather than only considering the interoperability of the data.

A large proportion of studies cannot easily be reproduced, which is a concern and may be due to a lack of clarity about the way the question was framed rather than how it was answered from the data.

There is also a need for us to better understand how federation works in practice: which are the relevant questions to be answered via federated networks and distributed analytics. We probably need to adopt a hybrid approach for different kinds of research question. It would be naïve to think that there should be a homogenous approach at a European level, and therefore for the EHDS. We need a heterogeneous architecture with interoperability support for diverse questions. For this we need to take a design view, not being wedded to one architecture, and with several philosophies: ask the right questions with more appropriate prioritisation; see this is a broad church and not a cult, embrace diversity; take advantage of rapidly advancing technology to facilitate temporal confidentiality and access, support hybrid architectural models.

To facilitate trust, reciprocal trustworthiness and to better manage citizen rights there needs to be shared ownership of the problem as well as of the solutions. There must be governance transparency in responding to issues such as data breaches, which need to be capable of being handled at a European level, which means a macro (EU) level agreement on the governance rules.

The focus of the EHDS should be on the secondary use of data targeting improving primary use needs in clinical care: better healthcare, better evidenced public health and societal benefit, and support of research and innovation. There are open topics to be addressed about its approaches: legal and governance (how do we regulate this?); data quality (what does it mean in practice, is there a maturity model, and how can we trust in data?); infrastructure (even if we have a hybrid approach, what can we trust as the basis for moving data?) and capacity building (by what means will we invest in people?).

Even within a hybrid model we will need an infrastructure that can support question and answer flow, and the upskilling of the user workforce. There are blocks that need to be overcome, such as intellectual property regarding data and data insights. We must learn from failings and from things that have not worked in the recent past.

It is vital for sustainability that we develop the value propositions. There is a need for a narrative that defines equitable benefits for all actors and avoids a fight over who pays more for data.

For example will the patient need to pay more, or pay less, or get paid, if their data is used to develop innovations? Who should get paid, or should be charged more? Since value is in the eye of the beholder, we need a definition

or multiple definitions. The absolute quid pro quo must be improved outcomes for all, and we must avoid a conflation of funding mechanisms to achieve this. We should consider if the value from data is an intrinsic property and right: everyone gains from it in the same way as we consider providing higher education. This is about a concept of societal good with incremental good being delivered per step. There are examples already, such as cross-border e-Prescription, the public health requirements that were met during COVID-19. These uses of data augment human rights. The EHDS should be considered necessary tool for human health in Europe.

There is a delicate balance between using data for transparent good, and commercialisation. The issue is not about ownership of data or results, but of ensuring always that there is a societal good. There is an increasing number of laws and applications to support this, such as in Denmark. It is important to balance the risks and benefits, not only to focus on the risks. Societal discourse is vital here, and this requires digital literacy, and upskilling of the population, so they understand the issues and can also understand what their own data tells them.

The group concluded that the EHDS is a necessary tool for human health & societal good, and should be substantiated as such, beyond traditional ROI, including just financial.

BREAKOUT C:

Sustaining data sharing and access by demonstrating value and trustworthy decision-making

Moderators Zoi Kolitsi and Bleddyn Rees

Briefing text circulated in advance

Many decision-makers still strongly favour and primarily use evidence from randomised controlled trials (RCTs). Real-world evidence is often still regarded as being of poor quality, not trustworthy, because of the inherent noise within real populations, the highly variable clinical practice and diverse care pathways used in different settings and countries, and the lack of a formalised discipline in data collection. On the other hand, there are strong counter arguments that promote the importance of real-world data reflecting the true spectrum of actual patients, from a wider age range, including the extremes of disease severity and a realistic pattern of comorbidities, than RCT trial subjects. Both RWE and RCTs evidence are required and it is not an either or choice.

This breakout group will start by examining the arguments against trusting real-world evidence and how those arguments might now, or in the future, be addressed. The following topics may be explored.

- A. **How busy junior healthcare staff can be motivated to capture good quality data,**
- B. **How EHR vendors and healthcare providers can be encouraged to adopt consistent data sets and value sets so that their data are combinable and comparable,**
- C. **Whether statistical methods can, or might in the future, be able to compensate for missing data and shifts in clinical documentation practice over time, or developing new methods and technologies to get out of the straight jacket of old fashioned medical record keeping**
- D. **Is trust primarily a matter of educating decision-makers and the people whose data it is about the benefits of/ real-world value of RWE?**
- E. **How could we build greater credibility in real world evidence, possibly through evidence quality benchmarks?**
- F. **How can we persuade stakeholders to re-connect to a newer, more fit for purpose, data value chain?**

Summary of the discussion

This breakout group was set in motion by addressing several questions that paved the way for the discussion and were addressed, one by one from two different angles:

- Why are we not trusting RWE today?
- How can these issues be addressed?

How can busy healthcare staff be motivated to capture good quality data?

Regarding this issue, three main challenges were discussed. The digital skills of health and care professionals are often not sufficient for this aim. Moreover, even when they are, it is still not clear how this may benefit their work, being frequently seen as an additional burden to their daily tasks, when the time to see patients is many times already reduced. The outcomes that may result from high quality data for the patients is also not always clear, thus being an additional factor to discourage this collection. While data capturing might be highly repetitive and manual, understanding the importance of the activity and the potential positive and tangible impact on people's lives represent motivational drivers. The limited perceived value of EHR systems, and the delegation of data entry from doctors to other healthcare team members, risk poor quality data and low reliability. Interactive technology applications can make data capturing process less cumbersome. EHR systems should check the content of the data as much as possible and prevent inaccurate data entry by non-acceptance of it. The data collected by them can support their own use of decision support in patient care, which is connected to the quality

of data which they capture. It is also important to continuously communicate to healthcare staff the importance of individual contribution for collective success. Safety of the patient care needs quality data because this is key for the other health professionals involved in the care process.

To address these challenges, two main actions were discussed focusing on education and supporting infrastructures. On one side, the inclusion of data management and the transformation of healthcare supported by digital tools in the health professionals' curricula; additionally, for those currently at work, proper training and retraining is needed. We have EU level co-operation but not enough regional and national regulations and efforts towards interoperability. Financial incentives were considered possible, however these are not always a long-lasting solution and the main aim should be the integration of data management procedures in the workflow.

This is mainly an issue of developing that culture. ICT platforms should support interoperability and integrated workflows by design and include feed-back tools, responding to the “need-to-know” of professionals.

How can EHR vendors and healthcare providers be encouraged to adopt consistent data sets and value sets so that their data are combinable and comparable?

Healthcare providers require accurate and consistent documentation of patient data into EHRs. These systems have largely yet to deliver on the promise to increase productivity, data interoperability and ability to generate insights for better decision making. The starting point would be to make EHRs more user friendly and ease the burden on practitioners during data collection. Healthcare purchasers and EHR system vendors should respond to transparent demands for data to be combinable and comparable. Some ideas were discussed about how to foster interoperable platforms and solutions, namely by promoting national frameworks and regulations on the use of standards and encouraging European co-operative guidelines and recommendations. Data interoperability between EHR systems (across vendors and for same vendor across software versions) should be built in by design.

Making funding conditional on these requirements, or providing rewards for implementing them, is also a viable option, even to promote the release of specific funding lines for that purpose, aiming at a broad implementation in Europe. A more disruptive option would be to introduce a “once-only” patient right: checks and test should only be undertaken once, and shared between care providers, not duplicated or repeated unless there is a new clinical need.

Can statistical methods, or might in the future, be able to compensate for missing data and shifts in clinical documentation practice over time, or there is the need to develop new methods and technologies to get out of the straight jacket of old-fashioned medical record keeping?

In this regard, the main challenges identified are the lack of interoperability, completeness and data quality. Considering that there is often missing data in EHRs and that many EHR platforms are not well designed to collect reliable and quality data, one main question discussed was if we have sufficient data to create statistical models that can in the future compensate for missing data elements. There is a research priority on this topic and at the same time the production of open data sets that can be used to properly train algorithms and support machine learning. There is still little knowledge about synthetic data, and for how complete and accurate data actually have to be. AI-based statistical methods can be used in analysis to compensate for low quality EHR data, or using synthetic data as a source for research analysis. The University of Oxford’s Big Data Institute has shown robustness against missing and incomplete data of findings on the efficacy of certain drugs. One main action that could address this is the promotion of links and joint activities between academia and hospitals, supporting programmes to create such data and synthetic models. It is important to better understand the relevance, breadth and depth of the data quality issues before determining strategies for statistical compensation for low quality. As an additional potential way forward, it was also debated if there should be incentives and an open culture for sharing properly-protected RWD for research within


the research community (since there is still a culture amongst the clinical research community of retaining “ownership” of collected RWD for each person’s own research).

How could we build greater credibility and trust in real world evidence?


The need identified in this question is the improvement of the quality of data and for that one of the possible actions discussed is benchmarking and labelling for the time being, although maybe in the future it will be desirable to develop metrics and work towards certification. Data collection and using benchmarks as a learning tool within the healthcare facility can be a good start, before real world evidence benchmarks starts to take shape.

Trust is a large topic. For example, how can we trust how organisations and service providers handle the data and what security and safety measures they adopt – this must be based on transparency. RWE should be based on scientific knowledge and methods which are transparent and can be verified.


One of the debated questions around this issue was the variability of trust cultures across the EU towards governments, as could recently be seen when comparing citizens’ reactions to the track and trace COVID apps. These were received by the public in different ways in different countries. Finland and Nordic countries have strong trust in governments and data collection was well received. This was not the case in southern countries like Portugal.




Education of citizens is essential for promoting trust. Education on what apps and systems are used for, presenting transparently how data will be used, respecting privacy, acknowledging the cultural differences and especially realising that one size doesn’t fit all. Moreover, it is essential to demonstrate how this is done in practice and what results are delivered. Diffusion of innovation, starting from thought leaders, will be important for this.



How can we persuade stakeholders to re-connect to a newer, more fit for purpose, data value chain?



It is likely that healthcare provider organisations will need to invest in improving the quality and interoperability of their routinely collected data. This may mean investing in training, upgrading their electronic health record systems and paying for new interoperability interfaces, appointing data quality managers etc.



However, individual healthcare provider organisations have rather weak business models and financial incentives to make their data more shareable and reusable. The lack of a joined-up value chain for health data has been highlighted over many years, but very little seems to have changed. Healthcare reimbursement models in Europe do not reward integrated care or person-centred care. However, a 'one size fits all' might not be suitable for such a complex topic. It is critical to work closely with public systems to be able to implement this at scale. Healthcare providers are largely not yet required to demonstrate their impact on health


outcomes to external agencies. Furthermore, the parties who most want to analyse data at scale are almost never closely connected to the parties that create the data and purchase the systems in which the data originally sits. It seems that multiple stakeholders have to shift their business priorities at the same time. Engaging stakeholders may be asking for new business models, but the challenge will be how to persuade stakeholders to re-connect to a newer, more fit for purpose, data value chain. Is it a possibility to pay for access but also pay for data quality? This was a question to be further developed in the discussions to come. EIT Health has included a chapter on business models within a white paper on optimising innovation pathways¹.

¹ <https://eithealth.eu/wp-content/uploads/2020/08/EIT-Health-Optimising-Innovation-Pathways-Think-Tank-White-Paper.pdf>


Final discussion on the enablers, disrupters and calls to action

The final discussion session, moderated by Dipak, focused on which stakeholders could really disrupt the data ecosystem, and what the enablers are for that disruption. Stakeholder change seems to happen very slowly. Changing the perspectives and actions of one stakeholder group, when the stakeholders that connect with them do not change, does not lead to much improvement. We need multiple stakeholder to move together. This tends to only happen through disruption. COVID-16 has been one such disruption, which has catalysed change but also shown us how unprepared we are to collaborate on generating and sharing intelligence at scale. The EHDS is a potential enabler for new disruptions.

Less has been said about the role of patients and citizens as potential disruptors. They are an agent who has historically been ignored as a catalyst for change but may now play that role given the growing importance of illness self-management, prevention, and citizen generated data.



In reality, though, there is a lot of rhetoric about citizens being in control of their health data. Much of the data about an individual is generated in a healthcare setting and hosted there. Personal health records and portals are mostly on a small scale. They are driven by champions and enthusiasts and activists, but on a societal level we are still using an approach that has a 'drop in the ocean' effect.



We also often treat patients as a homogenous group. In practice is everyone in a position to handle their own data: can they understand the concepts, does everyone want to manage and understand their own data and are there the correct mechanisms in place for those that do? We are far from a reality where the society is ready for this. How do we make sure that how can we give control back to citizens? This

does not only mean citizens determine who else can use their data but making sure the data flows to them in ways that they can get personal benefit from. We need to consider not only individuals but also their carers and family: how can a carer gain access to health data, raise questions about treatment decisions etc.? This issue needs to connect with the wider societal issue of increasing digitalisation and making data accessible to individuals to make use of.

We want patients to be involved in the collection of their data. We need that, especially as there is some data that only the patient can gather in their own environment. The only strong case we see for making use of data is in clinical research. We are still challenged to make better use of data in healthcare, and this is where the patient could play a strong role. Patients could ask their doctors for information, understanding, comparisons. Carers and families might ask as well.

We have a lack of definitions covering who is involved and how actors make decisions about data capture and data processing. Patients and citizens should have an input to prioritise how to use their data, areas they are concerned about, the translational impact the data needs to have and how to reflect this in decisions that can be actioned. It would be a great advance of the EHDS could develop and use digital tools to address some of these challenges and show how to action them in real time.

If we think of the EHDS is like a bank, we should be able to use digital tools to consult patients on the use of their data. Individuals might make different choices about the use of different categories of their data e.g. about different diseases they have. We need to make sure the way these choices are handled is determined by Europe, not the European Commission or any other single body. The concept of Independent Health Record Banks (IHRB) was first published in 2004, and perhaps we are now ready for this to become a mainstream model in Europe. The IHRB has to be independent of the data interests. There was a suggestion that Apple is doing this by holding health data and allowing the customer to control the way their data is used. However, this is a technology specific (company specific) approach. The control element as a feature of a product isn't what we need in terms of sharing these decisions and prioritization. They do not offer a basis to contribute opinions on whether the data is being used, nor is it transparent about the uses or the company trajectory. The digital tools and the citizen choices must be operated by independent, neutral and transparent parties, with no interest in the data. The governance rules

have to be determined by a collective of the people represented through the data.

There was a question about whether health-care payers should be using their financial reimbursement levers to drive greater care collaboration and therefore stimulate greater demand for interoperability. However, an alternative driver could be to champion health neighbourhoods: a collective benefit rather than the multiple individual benefits. Could we establish a neighbourhood data collective bank that could use the data and tools like AI to optimise the health of its neighbourhood?


Amazon's health team have shown use cases from the US where voice access to the EHR is given to patients, although under the control of their healthcare providers. The hospital decides what information a patient is permitted to access. If Europe isn't careful, there will be a consumer lead demand for these examples of disruption coming but which do not fit easily in European models. This is where the EC and the EHDS should be careful not to be left behind or have to follow and adopt approaches that do not fit well with European models and values. When looking at digital giants we look at the US – but how can we take advantage of Europe's lead in population level health services and preventive health to stimulate innovation in the European vendor space, to build a competitive advantage in the economy of the EU? Our biggest asset is our people!

It was also noted that we need to take into account that the computing power of mobiles is increasing. The argument used to be that individuals cannot have control over their complete EHR because of its size, but this may not be the case within the next few years. There


may be decentralization of the individual record towards the citizen. This may be a disruptive force without intending to be a solution per se. (In energy, solar panels and self-generating power has changed the energy market completely.) Where the data are held, especially if by citizens themselves or by independent banks on behalf of individuals, could be a powerful disruptor.

We need to see health outside of healthcare, and not be too focused on health as seen traditionally. Disruption may come first from well-being and prevention services, who are more willing to take innovation risks. However, a lot of people seem to be held back from making more use of or exerting more control over their data by not knowing what they're allowed to do. Risk reduction for people is an important paradox to tackle, as they do not always read or feel helped by guidelines. We lack interpretation layers, such as codes of good conduct that are written for easy understanding of what you can and cannot do, that could de-mystify the legal landscape (such as GDPR).

We also still have very poor levels of societal debate about the benefits and risks of different kinds of data use. We focus too much on the risks. We also invest a lot on new technology solutions but invest too little and only afterwards on communications, education and upskilling. Most of our challenges, even for interoperability, are socio-technical and related to humans, not to machines.



Have we over-invested in the provision of data, and under-invested in the uses of data? Regulators are an example stakeholder who could lead by example in making greater uses of health data from RWD including registries, stimulating momentum through Medicines Agencies within Member States.



Practical examples of beneficial uses of health data, in countries where EHR adoption is more advanced, will help inspire all countries and reduce the perceived risks. There is also already a momentum to use new regulated technologies. Innovations in software as a medical device can already transform health services by enabling patients to manage themselves. They may in future do this relying mostly on data already on the patient's phone, citizen generated, rather than needing to access the EHR.

Concluding remarks

The Round Table presentations and discussions as a whole have suggested a transformational role in the way we perceive and handle data, and where data originates, where it resides, who controls access and who gains access in order to deliver value. Interoperability standards will still play an important role. It seems as if the future lies less in supporting traditional point to point data communications and more in enabling shareable data as a societal good, potentially catalysed by the EHDS and reflected in its legislative proposal, digital services, liability framework and infrastructure. We might need a series of mini disruptors acting as critical enablers of this data ecosystem transformation.

APPENDIX 1:

Round Table Agenda

13.00	Connect to Zoom
13.05	Welcome, virtual meeting logistics: Bleddyn Rees
13.10	Meeting objectives, clarifying the problem space we are focusing on The importance of clinical and patient engagement in standards and quality: Dipak Kalra
13.20	Summary of the European EHR Exchange Format: Ceri Thompson, The EC, DG CONNECT
13.35	Federated architectures and the experience of IMI EHDEN: Nigel Hughes, Project Co-ordinator
13.50	What do decision makers need in order to trust real world evidence? Jesper Kjær, Danish Medicines Agency
14.05	Breakouts A: Improving the (re)usability of data through standards B: Architectures enabling large scale, secure and timely data access C: Sustaining data sharing and access by demonstrating value and trust worthy decision-making
14.50 - 15.00	Virtual coffee break, then consolidate the discussion points
15.15	Plenary feedback: 5 minutes per group + 5 minutes discussion
15.45	Plenary interactive discussion: Gently, gently standardisation isn't working. What are the disruptions, and who are the disruptors, we now need to mobilise? What role should key stakeholders each play to make sure the EHDS is a success?
16.15	Next steps, closing remarks
16.30	Close

APPENDIX 1:

Participant list

Name	Organisation
Ain Aaviksoo	GuardTime Health
Brendan Barnes	European Federation of Pharmaceutical Industries & Associations
Nicola Bedlington	DataSavesLives
Leonardo Calini	Microsoft
Maria Christofidou	i~HD
Catherine Chronaki	HL7 Europe
Carina Dantas	ECHAlliance & Digital Health Europe
Mathias Ekman	Microsoft
Ioana-Maria Gligor	DG Sante
Tomas Gornik	Better Care Health IT expert, co-chair of openEHR
Nigel Hughes	Janssen
Dipak Kalra	i~HD
Zoi Kolitsi	Digital Health Europe & I~HD
Jesper Kjær	Danish Medicines Agency
Jonathan Laszlo	Johnson & Johnson
Nathan Lea	i~HD
Licinio Kustra- Mano	DG Sante
Henrique Martins	EU eHealth expert, Professor of Health Management & Leadership
Birgit Morlion	DG Connect
Bleddyn Rees	Digital Health Society & ECHAlliance
Michael Short	UK Department for International Trade
Michael Strübin	MedTechEurope
Louisa Stüwe	The French Health Data Hub
Paul Timmers	University of Oxford
Ceri Thompson	DG Connect
Aneta Tyszkiewicz	European Federation of Pharmaceutical Industries & Associations
Veronica Zilli	Johnson & Johnson
Invitees who could not attend but have contributed to the report	



Calls to Action on Health Data Ecosystems

RECOMMENDATIONS FROM THE DHS & I~HD ROUND TABLES

THIS INITIATIVE IS SUPPORTED BY

