abpi

Bringing medicines to *life*

Big data road map

Foreword

Rt Hon Earl Howe,
Parliamentary Under Secretary for Quality, Department of Health

The potential of health informatics to drive growth and provide business opportunities while helping achieve better health outcomes in the UK is well recognised. The Government has responded by investing in the UK's capabilities and infrastructure to develop a sustainable health informatics research base and help the UK to compete globally, including creating the Farr Institute comprising the four e-Health Informatics Research Centres, the Big Data Institute on biomedical research at Oxford, and support for the European Bioinformatics Institute at Cambridge. We recognise this is a continuing endeavour and welcome this timely report by the ABPI setting out a road map of potential actions over the next four years. In October, the Government published; "Seizing the data opportunity", our strategy for developing the UK's data capability. The strategy sets a clear vision for how the UK can lead the way in extracting knowledge and value from data. It is clear that partners must come together to deliver this vision for "big data" in the UK, and Government will continue to play a key role in this.

Rt Hon David Willetts MP,
Minister for Universities and Science, Department for Business, Innovation and Skills

Stephen Whitehead,
Chief Executive, ABPI

The ABPI is a proud partner to the research community in the UK. With this report, we hope to make an important contribution to building on the UK's potential for world-class data-led research. Building the UK's capabilities in this area will not only help us compete in the global race for investment and jobs in the knowledge-based industries of tomorrow, it will also help to advance healthcare through the development of new treatments and patient-centred services. This vision can only be delivered by working together. I hope that this report can be the starting point for new collaborations and partnerships to bring us closer to that vision.

The ABPI represents innovative research-based biopharmaceutical companies, large, medium and small, leading an exciting new era of biosciences in the UK.

Our industry, a major contributor to the economy of the UK, brings life-saving and life-enhancing medicines to patients. Our members supply 90 per cent of all medicines used by the NHS, and are researching and developing over two-thirds of the current medicines pipeline, ensuring that the UK remains at the forefront of helping patients prevent and overcome diseases.

The ABPI is recognised by government as the industry body negotiating on behalf of the branded pharmaceutical industry, for statutory consultation requirements including the pricing scheme for medicines in the UK.

The ABPI is grateful to KPMG for its contribution to this work and for developing the report.

This study engaged over 70 individuals from the pharmaceutical industry, the health sector and academia over several months in 2013. Their contribution has been invaluable and the study team would like to thank them for their commitment.

Contents

1 Introduction

90%

OF THE WORLD'S TOTAL DATA HAS BEEN CREATED JUST WITHIN THE PAST TWO YEARS.

IBM, 2012

A health and social care system under pressure

The UK is facing unprecedented levels of strain in our health and social care systems. On the demand side, like most healthcare systems around the world, we are confronting an ageing population with high and increasing levels of co-morbidities and rising patient expectations. On the supply side, rapid healthcare cost inflation, limited productivity gains and evermore constrained public resources are forcing policy-makers to look for transformational, rather than incremental, changes to health and social care services.

A pharmaceutical industry that is transforming

While the UK remains at the forefront of scientific research, the pharmaceutical industry in the UK faces a tough commercial environment with some of the lowest prices in Europe and slow uptake of innovative medicines in the NHS. Challenges also remain in conducting clinical research in the NHS with the UK's global share of patients enrolled in clinical trials falling from 6% to 1.4% between 2000 and 2010[1].

Shifting the balance from volume to value

Despite these challenges, there is a strong sense of shared opportunity for both the pharmaceutical industry and the NHS. A number of responses to these complex issues are emerging and some significant multi-stakeholder initiatives are underway.

Outcome-focused commissioning: The NHS in England has undergone a large scale reorganisation via the introduction of clinically-led commissioning and a focus on delivering high quality, safe care in the face of a financial challenge estimated at £30 billion, over and above the 'Nicholson challenge' to find £20 billion by 2015.

Medicines optimisation: This sets out four principles for healthcare professionals to improve patient outcomes with greater consideration of the patient's experience, the evidence, adherence and safe use of medicines.

Joint working: Partnerships are becoming increasingly common within and between the life sciences industries, healthcare providers and academia. Cross-stakeholder initiatives such as the National Institute for Health Research (NIHR) Clinical Research Networks, Translational Research Partnerships and the Academic Health Science Networks aim to promote innovation and collaboration and may act as real catalysts for change.

Incentives to invest: The Strategy for UK Life Sciences outlines the Government's approach to providing incentives for industry. These include tax incentives, investment in skills and capabilities, improvements to public datasets and provision of a streamlined clinical research infrastructure. It also includes a drive to increase the rate at which the NHS takes up new innovations.

The 'involved' patient: More fundamentally, the shift in the role of patients from passive recipients to active managers of their own care offers new opportunities for improved patient outcomes at lower cost.

[1] NESTA: 'All together now: Improving cross-sector collaboration in the UK biomedical industry', March 2011

The role of big data

Big data is a common theme underpinning many of the proposed solutions to the challenges facing the NHS, the life sciences research community and the pharmaceutical industry.

Big data technologies make it easier to work with large datasets, link different datasets, detect patterns in real time, predict outcomes, undertake dynamic risk scoring and test hypotheses. Other sectors have already proven the value of big data. For instance, technology and web-based companies such as Google and Amazon rely on big data techniques to develop, target and enhance services. We have begun to observe big data in action in our industry. Human genome sequencing has undergone a paradigm shift from its initial 10-year project to a less than one week task. Already, the information from this endeavour is improving our understanding of the genetic basis of many human diseases. There is potential for big data approaches to go beyond current levels of capability in order to address the bigger picture issues such as identification of new therapeutic targets for complex conditions, improved efficiency in clinical trial recruitment and generation of innovative technologies.

Electronic medical information can now be linked securely on an unprecedented scale, potentially revealing how diseases manifest and how treatments are used in the real world. Social media and patient websites also contain increasing amounts of information on actual patient experiences. When combined with health monitors, genomic and clinical trial data, it will be possible to gain new insights for the health of individuals and whole populations.

However, if big data is to be a powerful catalyst for improving patient outcomes, its uses and limitations need to be understood by everyone from company Board members and academics to hospital managers, clinicians and patients. The fundamental challenge is not the technology. Unlocking the potential of big data for the health and wealth of the UK requires a strong collaborative culture, an overarching robust governance framework and a shared data ecosystem – one that involves stakeholders aligning around a vision of delivering better patient outcomes at lower costs, with common standards, methodologies, processes and services to support practical action. We believe failure to fully grasp this and build on the developing ecosystem will result in a significant missed opportunity for both public health and UK plc.

WHAT ARE THE TYPES OF QUESTIONS BIG DATA COULD HELP ANSWER?

- How can we identify new therapeutic targets more quickly and effectively?
- How can we identify patients with a high risk profile for certain diseases before they present with symptoms?
- How can we better prevent chronic conditions?
- What are the best care pathways to manage patients with different combinations of co-morbidities?
- How can we better identify patients appropriately for clinical trials?
- How can data support better adherence to national and local guidance, such as NICE recommended pathways?
- What is the best way to model the relative impacts of different healthcare interventions?
- Which factors influence patients' adherence to healthcare interventions or advice and reduce wastage?

Overview of this road map

This road map sets out a four-point plan to direct progress in this area over the next four years. We outline the opportunities and challenges that big data presents, we give an overview of the UK's big data assets and issues, and we recommend specific actions required to create the conditions for success. Success in this area must be defined in terms of a long-term and multi-sectoral ambition: to improve patient health, increase NHS cost-effectiveness and maximise the potential for global pharmaceutical industry investment in the UK.

This report is a first step in a fast-changing technical and competitive environment. We intentionally set out actions only for the next four years – to look beyond then seems unrealistic. Nevertheless, the uncertain, distant horizon should not justify inaction. These short- to medium-term actions are practical enablers of long term success and may determine a future turning point for the UK pharmaceutical industry.

EVERY

TWO DAYS NOW WE CREATE AS MUCH INFORMATION AS FROM THE DAWN OF CIVILISATION UP UNTIL 2003.

Eric Schmidt, Chairman, Google, 2010

2 The big data challenge

A call to action

Data, data, everywhere...

Unmistakably, the pharmaceutical and healthcare industries and academic researchers are now at the centre of a rapidly proliferating data ecosystem which could prove to be one of the most significant drivers of change for both the public and private sectors over the next decade.

WHAT IS BIG DATA?

Big data is what we get when we can access, store and analyse vast quantities of complex and variable data. The term big data is used to describe two things: The characteristics of the data ...

- **Volume:** the amount of data generated by organisations or individuals
- **Velocity:** the frequency and speed at which it is generated, captured and shared
- **Variety:** the new types of data generated.

... plus the analytics – the approaches and methods used to generate and visualise the information and insight held within the data.

The technology industry also often adds two more 'V's to describe some of the practical challenges of big data:

- **Veracity:** as reliance on external sources of data grows, to what extent can we trust the data?
- **Value:** how can we avoid drowning in big data and ensure a return on investment?

VOLUME

THE AMOUNT OF DATA GENERATED BY ORGANISATIONS OR INDIVIDUALS

VELOCITY

THE FREQUENCY AND SPEED AT WHICH IT IS GENERATED, CAPTURED AND SHARED

VARIETY

THE NEW TYPES OF DATA GENERATED

VERACITY

TO WHAT EXTENT CAN WE TRUST THE DATA AND THE OUTPUT?

VALUE

HOW CAN WE ENSURE A RETURN ON INVESTMENT?

Where is the data coming from?

Exciting recent developments in technology and research have led to huge advances in our understanding of the molecular and genetic basis of disease – vast amounts of information are being generated from human genome analysis. At an unprecedented scale, electronic record data from all parts of the healthcare system are being linked and have the potential to reveal cross-sectional and longitudinal trends relating to how diseases manifest and how treatments are used and work in the real world. Social media and websites also contain vast amounts of information on patients' experience of both diseases and treatments. The rate of increase in data volume is also growing. Consider this: it is estimated that when the Square Kilometre Array radio telescope goes live in 2020, it will produce more data within a few days of operation than the entire current size of the internet.

In short, in theory we now have the technology to collect, store, transform, access and analyse vast amounts of data at modest cost. This includes performance and clinical data from GPs and hospitals, data from clinical research by industry and academia and data from patients and the public generated through social media and other sources.

What are the applications?

The profusion of data – in and of itself – does not necessarily guarantee value for patients, academic researchers, healthcare managers and professionals or the pharmaceutical industry. We need to be able to process it into easily usable formats, analyse and link it together if we are to generate insight and knowledge to identify better treatments, help improve patient outcomes and healthcare cost-effectiveness. For the pharmaceutical and healthcare industries and academic researchers, the analytical applications can be described in four categories of increasing sophistication:

- **Processing large volumes of data** – such as genome sequencing, complex imaging, data mining
- **Tracking indicators over time to detect patterns** – such as real time clinical monitoring, benefit/risk profiling, compliance monitoring and digital engagement monitoring
- **Linking data sources** – detecting correlations across multiple datasets for, example, for population stratification, trial cohort identification, co-morbidity analysis, safety signal verification, comparative outcome analysis, data verification
- **Modelling complex systems or decisions** – decision support for inclusion/exclusion criteria and/or outcome prediction in clinical trials, development risk profiling, pricing models.

These data-driven applications span the value chain from drug discovery to healthcare delivery. There are strong, shared benefits for healthcare and life sciences sectors with the long-term potential to underpin real breakthroughs in clinical development and safety monitoring of medicines, including the shift to stratified medicine and the associated challenges for value assessment.

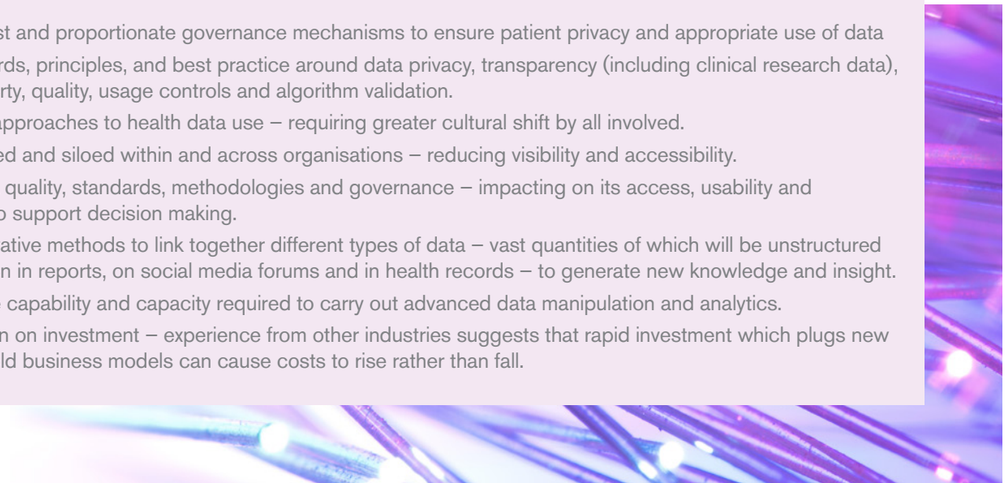
What are the challenges?

The healthcare, academic and pharmaceutical sectors have specific challenges, beyond the core technical and analytical requirements. It is relatively easy to apply big data approaches within one organisation, but where opportunities for further analysis of data sit across organisations, there are potential commercial, legal and ethical barriers. The NHS and academia consist of many thousands of related organisations which have limited freedoms or incentives to share data at the level that will be needed in future. Despite a rapid global shift towards greater data transparency, big data makes it increasingly clear that we are not accustomed to dealing with data as an asset beyond the walls of a single organisation.

The pharmaceutical and healthcare sectors also face greater potential risks than many other sectors, since wrong treatment decisions can result in serious or irreversible harm to patients. As big data analytics and algorithms will increasingly lie behind commissioning and clinical decisions, they will have a significant impact on individuals and may shift public expectations and perceptions of accountability. Finding interesting data correlations, without understanding why, is often good enough in other sectors, but less so in health. Researchers are familiar with the risks – though not necessarily immune to them. As access to large data sets becomes easier, it may also be easier to make analytical mistakes unless rigorous standards and governance controls are applied. It is also possible that better analytics will also deliver some uncomfortable insights into the marginal value of some medicines.

COMMON BIG DATA CHALLENGES FOR ALL

- Developing robust and proportionate governance mechanisms to ensure patient privacy and appropriate use of data - defining standards, principles, and best practice around data privacy, transparency (including clinical research data), intellectual property, quality, usage controls and algorithm validation.
- Embracing new approaches to health data use – requiring greater cultural shift by all involved.
- Data is fragmented and siloed within and across organisations – reducing visibility and accessibility.
- Inconsistent data quality, standards, methodologies and governance – impacting on its access, usability and trustworthiness to support decision making.
- Developing innovative methods to link together different types of data – vast quantities of which will be unstructured information written in reports, on social media forums and in health records – to generate new knowledge and insight.
- A shortage of the capability and capacity required to carry out advanced data manipulation and analytics.
- Managing a return on investment – experience from other industries suggests that rapid investment which plugs new technology into old business models can cause costs to rise rather than fall.



What does big data look like?

What's new?

Today opportunities for innovation frequently lie in the further analysis of data – beyond the primary use for which the data was generated. New technologies and policies are beginning to improve access to, and analysis of this data while ensuring protection of individual privacy.

HEALTHCARE DATA

Structured and unstructured data from anonymised electronic health records and biometric data sources, such as images

Examples

Phenotyped patient cohorts for stratification
Prescriptions, compliance and response patterns
Pathology results and laboratory tests
Probabilistic models of outcomes

GENETIC AND GENOMIC DATA

Ever growing amounts of new gene sequencing data

Examples

100K Genome Project and Biobank data
Tumour mutation patterns
Whole genome sequencing

AUTOMATED SENSORS AND SMART DEVICES

Phone apps for home monitoring, handheld and sensor-based wireless or smart devices

Examples

Movement monitors
Dosage monitoring apps
'Smart' insulin pens
Digital pills

SELF-GENERATED DATA AND DIGITAL ENGAGEMENT

Every click or post says something about you... but what?

Examples

Healthcare professional and patient forums for sharing experience
Mobile phone data for tracking epidemics
Social media

PUBLIC DATA RELEASES

Government, academic and commercial data open to the public. Approximately 3,000 open datasets on data.gov.uk are health-related.

Examples

Aggregated population health data
Commissioning data
Summary clinical trial results
Open research data

Where are the opportunities?

A CALL TO ACTION

Big data analytics have the potential to link diagnosis and treatment more effectively; to accelerate and justify the shift toward stratified medicine; to get treatments into use faster; to enable the industry to shift from volume to value and from supplier to partner in the healthcare system.

No one organisation or sector can address these opportunities or challenges alone. If we are to leverage big data for the benefit of patients, the industry and the wider healthcare ecosystem, we must tackle the challenges together. At stake are the economic health of the UK pharmaceutical industry, the quality of healthcare services being delivered and ultimately the health of the public.

EXAMPLE BIG DATA APPLICATIONS

	Life Sciences				Healthcare		
	DISCOVERY	DEVELOPMENT	MANUFACTURING, SUPPLY AND DISTRIBUTION	HEALTH TECHNOLOGY ASSESSMENT (HTA)	COMMISSIONING	PATIENT USAGE	HEALTH OUTCOMES
Example area	Genomics	Clinical trials	Supply chain management	Benefit-risk/HTA	Value-based commissioning	Medicines optimisation	Outcome management
COMMON TODAY	▪ Genetic testing	▪ Trial design and analysis ▪ Trial recruitment ▪ Post-trial follow-up	▪ Computerised prescribing	▪ NICE assessments	▪ Population segmentation	▪ Uptake metrics	▪ Outcomes framework ▪ Clinical model reviews
POSSIBLE TODAY	▪ Genomic analysis ▪ Automated research meta-analysis	▪ 'Real world data' studies	▪ Unique identifier on each pack	▪ Retrospective comparative effectiveness analysis	▪ Retrospective real world evidence of optimised medicines usage	▪ Linking medicine use to outcomes ▪ Adherence management technologies	▪ Self-tracking for clinical data generation ▪ Digital engagement monitoring ▪ Post-launch stratification
POSSIBLE IN FUTURE	▪ Enhanced therapeutic target identification and virtual drug design	▪ Automated stratification ▪ 'Real world' real-time analysis and pharmacovigilance	▪ Efficient inventory control systems ▪ Enhanced counterfeit prevention	▪ Predictive impact modelling	▪ Modelling stratified service and pathway redesign	▪ Elimination of prescribing errors and waste	▪ Expert diagnostic algorithms ▪ Personalised treatment

Common base of real world clinical data

3

Big data in the UK

How do we compete?

How do we compete?

From data to insight to value

Big data has potential applications across the whole value chain from drug discovery through to provision of healthcare at the front line. With so many opportunities – and issues – the challenge is to know where to begin. Not only do opportunities differ in intrinsic value across individual organisations and to the pharmaceutical industry as a whole, but practical realisation of big data opportunities relies on a wider data ecosystem of assets and services.

It is a common mistake to assume that value of big data lies only in the data itself – its volume, accuracy, accessibility, 'linkability', and so on. Despite its importance, the 'bigger' the data, the less this holds true. Even with high quality data it is not possible to leap from raw data straight to business value. All big data solutions can be approached in a similar way, from data to insight to value: capturing the data; transforming it into more easily usable formats; analysing the data; generating knowledge; and applying the knowledge to produce valuable insight. Each step requires a different investment in skills, technology, tools and techniques which broadly reflects the complexity of the data and of the question at stake.

Building UK assets in a knowledge economy

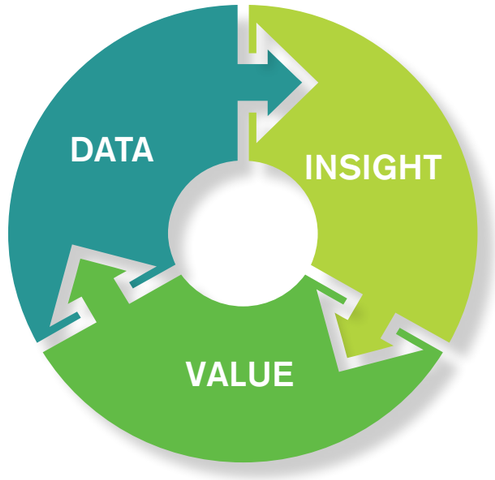
The flip side of this big data cost is the potential for competitive differentiation. Data, skills, technologies, tools and techniques are increasingly treated as assets that can be valued at the level of both companies and countries. In a complex, fast moving field it is not feasible to progress across all fronts equally quickly, so the question is – what investments and actions could be both valuable and practical in a given timeframe?

The remainder of this section gives an overview of the UK's potential assets in three stages:

1. Data

2. From data to insight

3. From insight to value



INFORMATION
HAS THE POTENTIAL TO BE THE BIGGEST DRIVER OF INNOVATION AND IMPROVEMENT IN HEALTH AND SOCIAL CARE

Kingsley Manning, Chair of the Health and Social Care Information Centre

Data

The demand strategy

The many different data types of interest to the pharmaceutical industry which are amenable to big data solutions can be considered in five data domains. In a knowledge-based business these domains, and the links between them, are a crucial driver of value. These are by no means mutually exclusive. Each domain is at a different stage of maturity in the UK and warrants attention.

1. **PEOPLE:** populations and patients, including clinical records and pathology results.

   In the UK this domain is a potentially enormously valuable asset. The NHS is one of the world's most comprehensive sources of 'deep', longitudinal patient data. It maintains 60 million patient records, across primary, secondary and tertiary care in an ethnically diverse population. In principle at least, this presents considerable analytical opportunities. The NHS is undertaking a massive programme to make this data easier to use for research purposes, although this will take time. Nevertheless, the creation of the Health and Social Care Information Centre (HSCIC), the launch in England of care.data and new Clinical Practice Research Datalink (CPRD) services offer potential breakthroughs, linking and expanding primary and secondary care data to form high volume, detailed patient journeys. Equivalent datasets in Scotland and Wales are already more closely linked.

   The ownership of patient data is slowly shifting from the clinician toward the patient, creating new opportunities for patients to engage in clinical research and use online forums and telehealth products.

2. **DISEASES:** the biology of disease, including genetic and genomic data, tissue banks and research outputs.

   This data domain is rapidly globalising and there is a strong push for global standardisation to support analysis across massive datasets. In the UK our deep research expertise and government and research funder investment, for example in the 100K genome project and UK Biobank, will position us well for future links between genomic, proteomic data and clinical records enabling the UK to continue to play a leading role in the understanding of the pathophysiology of important diseases and the effects of treatment.

3. **CLINICAL TRIALS:** from registers to results.

   The pharmaceutical industry is responding to moves toward greater transparency in the reporting of clinical trial results and wants to see this implemented in a manner that supports good research, with data being accessed, analysed and interpreted appropriately. This will allow researchers to supplement companies' own analyses of trial data, particularly through the pooling of data from multiple studies. Such systematic reviews and meta-analyses can provide clinical decision-makers and health policy-makers with valuable information about the effectiveness and safety of medicines. The Cochrane Collaboration is at the forefront of such research.

   As the NHS begins to use more 'real world' data to make local decisions about healthcare delivery, there is also an opportunity for the pharmaceutical industry to focus research on demonstrating the value of medicines beyond the efficacy and safety demonstrated in randomised clinical trials. Working with key stakeholders, the UK pharmaceutical industry has the potential to be a world leader in the conduct of real world studies.

4. **MEDICINES:** from compounds to commercial products, including prescribing, sales and safety data.

   As the scale and complexity of precompetitive collaborations increase, the UK research community is actively engaged in building platforms to enable pre-clinical data to be shared with a view to accelerating the pace of discovery of new medicines.

   Public data releases have already had a significant impact, with the release in England of prescribing data by GP practice. However, better integrating NHS prescribing and adverse incident data with clinical records has the potential not only to provide strong commercial insights to both commissioners and pharmaceutical companies, but also to strengthen post-marketing safety signal surveillance.

5. **TREATMENTS:** healthcare system operational data including NHS pathways, commissioning models, performance, clinical support systems and patient feedback.

   As the NHS shifts from volume to outcome-based commissioning, it will become increasingly important for commissioners, regulators and suppliers to develop a common understanding of NHS operational data. Treatment underpins real world impact and the associated data and models form part of a shared assessment of value throughout the life cycle of a medicine. There is also significant potential for these models to assist commissioners in making better use of resources in the NHS.

Achievements to date

The UK boasts 'cradle-to-grave' healthcare with associated health data and has pioneered projects such as the General Practice Research Database (GPRD) now the Clinical Practice Research Datalink (CPRD). The December 2011 launch of the Government's 10-year Strategy for UK Life Sciences has led to a number of investments and initiatives to build capabilities for research based on electronic health records, demonstrating a clear commitment to leveraging electronic health data for improved patient outcomes as well as efficiency in the delivery of services.

Important developments have included: the establishment of the Clinical Practice Research Datalink (CPRD); the enhanced role of the Health and Social Care Information Centre (HSCIC) and the recent launch of the Farr Health Informatics Research Institute (which include the E-Health Informatics Research Centres). The CPRD is considered by many as a gold standard for its observational data and interventional research service and its usage has already resulted in over 1400 clinical reviews and papers. Importantly for researchers CPRD holds a range of over-arching governance approvals that make undertaking research using data far simpler. Additionally CPRD has started releasing powerful and fast new web-based software that is able, using real world data, to undertake complex inclusion/exclusion queries and actually enable potential site and patient identification.

The HSCIC has also provided access to some 293 data sets either directly or indirectly, with more available through the NHS Business Services Authority.

The recent Caldicott review has also established a clear framework under which data should be used and governed. There are also commitments in the NHS Constitution on access to information, informed choice and opportunities to participate in research for all patients.

Challenges that remain

Despite ambitious plans and real opportunities for the UK, barriers remain to exploiting the full breadth and depth of big data across the developing ecosystem and maximising the UK opportunities:

- Simplification of how permission to access some datasets can be obtained: lack of overview of permissions for different uses of available data, who 'owns' it and how to access it
- Linkage and interoperability limitations within and between healthcare, academia and private sector datasets
- Limited data services and an immature intermediary market to deliver rapid business intelligence solutions in the innovation space
- Issues of breadth and depth of data coverage and accessibility across primary and secondary care and related services such as care homes, mental health and ambulance services
- The need for extending the governance framework that further protects patient confidentiality and addresses issues such as intellectual property, enhanced data security and data quality assurance for new 'big data' sources
- Limited experience amongst our industry members and across healthcare generally of the full extent of big data opportunities and need for a clearer articulation of the UK offerings.

Where is the data?

From static metadata catalogues to a dynamic metadata platform

Paradoxically, one of the UK's strengths is also one of the barriers. In general, there are too many data sources and it is difficult to know where to begin. Much data consequently remains underused. The IT industry sometimes describes this type of transparency issue as 'dark data'. Across the life sciences and healthcare sectors its scale is vast. The five data domains described above cover many tens of thousands of datasets held by thousands of organisations across the public, private and third sectors globally. All potential data users want to understand not only where the 'best' data is, but its content, accessibility, accuracy, interoperability and likely relevance to a given task.

This type of knowledge is known as metadata – data about data. Ultimately, metadata is needed to judge the potential value of the associated data in a given context as quickly and reliably as possible. In general, the quality of metadata is improving. For example, the Health and Social Care Information Centre (HSCIC) provides standardised metadata for open datasets and will build a form of 'data asset catalogue' to help give a clear view of the NHS data landscape.

Nevertheless, today most metadata is still of limited value and held with, or near, the data at source. It still requires initial effort by each potential user to locate and interpret, and that knowledge cannot be easily shared. This project initially considered creating a simple metadata catalogue with a reach beyond the NHS, a high level view of the data landscape of owners and their data for newcomers to the field (see illustration). The limitations of such a product were clear from the outset – the landscape is large and fast changing and a high-level snapshot can only give a passable sense of the potential value of the data.

These challenges suggest an opportunity to consider a different solution. An application more akin to a dynamic metadata platform, for example, could be created and updated by stakeholders, who would form communities of interest around specific data domains, related datasets or sources. It could include linkage experiences and examples of use in addition to standardised metadata.

Ultimately this drive for greater metadata transparency is part of a wider shift in the basis of competitive differentiation between companies and countries in the life sciences. This shift is an opportunity for the UK to leverage its data assets through the creation of a framework to manage this kind of dynamic metadata platform.

Extract from illustrative metadata catalogue

WHAT TYPE OF DATA SOURCE IS IT?			WHAT TYPE OF CONTENT?					CAN IT BE USED?				WHAT COULD IT BE USED FOR?					HOW GOOD IS THE DATA?							
SOURCE			DATA CONTENT					ACCESSIBILITY				RELEVANCE					DATA QUALITY					DATA SERVICES			
Generic/Eg/Family/Specific	Custodians	Aggregated/ Pseudonomised/Anonymised	People	Disease	Drug	Trials	Treatment	Open	Subscription	Defined agreement process	Bespoke agreement required	Observational studies	Interventional trials	Pharmacovigilance	Comparative effectiveness	Genomic/biomarker analysis	Volume	Completeness	Accuracy	Use of standards, e.g. coding	Timelines	Metadata quality	Related services	Speed of access	Linkage opportunity
	SOURCE 1	AGG						✓	✗	✗	✗														
F	SOURCE 2	AGG						✓	✗	✗	✓														
F	SOURCE 3	AGG						✓	✗	✓	✗														
F	SOURCE 4	AGG						✓	✗	✗	✗														

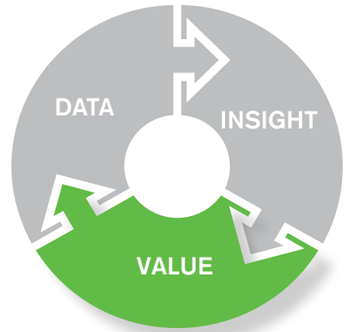
From data to insight

Converting big data into insight in order to improve healthcare is not fundamentally about amassing ever-larger quantities of data. It requires a combination of technical, analytical and clinical skills and experience. It also needs a dynamic market for data solutions and services, and collaboration across the ecosystem.

This is potentially good news for the UK. As data becomes less resource intensive and easier to collect, the value will increasingly lie in the reputation of its services. Although these specific skills are hard to come by, the UK boasts three competitive strengths in addition to data itself: deep research expertise; a strong heritage of collaboration between academia, industry and the public sector, soon to be supported by Academic Health Science Networks (AHSNs); and recent government and research funder investments including a network of centres focused on health informatics. Training conducted at the Farr Health Informatics Research Institute addresses one of the key challenges associated with big data: not enough well-trained data scientists capable of advanced data manipulation and analytics.

Notwithstanding these strengths, there remain a number of challenges to attracting global investment on the basis of the UK's skill assets. These include:

- Fragmented service providers across healthcare and academia with embryonic industry-facing services, risking duplication of investment, assets and services
- Immature methodologies, algorithms, platforms and models and associated governance
- Ad-hoc mechanisms to build innovative partnerships and shared products
- A serious shortage of data scientists, bioinformaticians and chemoinformaticians with the appropriate skills and understanding to handle the complexities of big data and all its applications



From insight to value

The growth of research and development (R&D) in the UK requires clear value propositions for global investors – practical, proven commercial offers which do not require detailed knowledge of the data and service landscape behind them. Based on the UK's emerging data and service assets, there are a few high priority areas for the pharmaceutical industry which could benefit from further attention:

- Interventional and adaptive clinical trials requiring large, diverse populations: This is a much awaited UK offer requiring linked patient journey data.
- Early stratification: Linking a wide range of genomic, phenotypic, prescribing and self-generated data for precise targeting of medicines. This is a breakthrough opportunity, though longer term.
- NHS commissioning process: The relationship between company executives and the newly-formed NHS commissioning bodies is still developing. There is an opportunity to standardise frameworks and models to demonstrate efficacy, safety and clinical utility of medicines to underpin improved relationships between commissioners, clinicians and the industry.
- Near-real-time pharmacovigilance: Greater use of electronic clinical and prescribing records could enhance patient safety and support regulatory innovation, such as the Early Access Scheme and adaptive licensing, by ensuring appropriate early delivery of medicines where there is a high unmet medical need.

CASE STUDY:
The Salford Lung Study

The purpose of the Salford Lung Study is to test the safety and effectiveness of a new treatment for asthma and Chronic Obstructive Pulmonary Disease (COPD), compared with standard medications used for these conditions. The study is sponsored by GlaxoSmithKline (GSK).

The study is a collaboration between GSK, North West e-Health (NWeH), the University of Manchester, Salford Royal NHS Foundation Trust, NHS Salford's local general practitioners, and local community pharmacists. Collectively these organisations' involvement in the project has been unique and is a recognised world first for the use of such data.

The initiative draws on Salford's e-Health records infrastructure, a clinical information system that provides a single, integrated electronic patient record across primary and secondary care. This will ensure patients are closely monitored over the course of the study, yet with minimal intrusion into their everyday lives.

What are the challenges?

"The Salford Lung Study has seen excellent collaboration between a number of partners. As a 'first' of its kind, there have been many challenges and lessons learnt on how to set up and run such studies better in the future. The most important of these is the need to develop an efficient integrated multidisciplinary team who all come from different working cultures – for example, the GSK clinical operations team, the community nurse team and NW e-health technical team – to ensure a common understanding and ways of working. In addition, we should never underestimate the amount of training and support required when setting up research naive sites, in this case, GP practices and pharmacies; this resulted in a significant increase in resource required."

"Another significant challenge has been that primary and secondary data and indeed pharmacy data are not linked. North West eHealth has had to work hard to link the data and this is one of the few places in the UK to have the capability. If the UK is to become a key player in the generation of real world data, we will need the ability to link these data sources in real time."

"We are confident that we have been able to adapt and address the issues and will be able to deliver results from this important study."

GSK, 2013

CASE STUDY:
Systemic Anti-Cancer Therapy Dataset (SACT)

The national collection of all cancer chemotherapy information in the NHS in England began in April 2012, in line with the requirements of the Department of Health's policy document *Improving Outcomes: A Strategy for Cancer* (January 2011).

With the advent of electronic recording of treatment, and in particular electronic prescribing systems, it has now become realistic to operate a national collection and analysis system for cancer chemotherapy provided within the NHS. The data collection covers all patients receiving cancer chemotherapy in or funded by the NHS in England, both adult and paediatric, in acute inpatient, daycase, outpatient settings and delivery in the community. It covers chemotherapy treatment for all solid tumour and haematological malignancies and those in clinical trials.

The SACT team has focused on data collection, standardisation, quality and completeness. Regular reports now highlight changing patterns and variation in clinical practice across the country.

How could it be used in future?

As the data quantity and quality increases, the SACT will build a complete picture of the cancer chemotherapy treatment pathway for each patient and will be able to satisfy new demands for information in clinical expert groups, academics, commissioners and pharmaceutical research.

The SACT team has reported preliminary discussions on how the SACT data model and the CIU delivery mechanism could be expanded to other disease groups, stroke and musculo-skeletal, for example. It is also considering plans for linking to related datasets, for example the Radiotherapy Dataset (RTDS) or National Audit Data such as DAHNO (Head and Neck), and the development of similar systems in the devolved nations, This would add a new dimension to how the clinical and commissioning communities are able to view treatment and outcome data at a national level.

Current examples of big data related investments in UK healthcare

Clinical Practice Research Datalink (CPRD) – The new English NHS data research service launched in 2012 and is jointly funded by the NHS National Institute for Health Research (NIHR) and the Medicines and Healthcare products Regulatory Agency (MHRA). CPRD provides an observational data and interventional research service using extensively linked clinical information from GPs, hospitals, audit datasets, disease registries, and holds the resulting data in an anonymised form. This presents an array of opportunities for academia and industry to access patient cohorts to run many types of pharmacoepidemiology studies as well as improve the efficiency in running clinical trials.

Farr Health Informatics Research Institute (including the E-Health Informatics Research Centres) – Supported by a 10-funder consortium, co-ordinated by the Medical Research Council (MRC), it comprises four major centres in London, Dundee, Manchester and Swansea and will link research in 19 universities across the UK. The Farr Institute will pursue cutting-edge research, deliver innovative linkage and analysis of health-related data sets, build research capacity and expertise and act as an interface with industry, clinical practitioners and policy makers with the aim of building and sustaining a vibrant health informatics research capability in the UK.

The Health and Social Care Information Centre (HSCIC) – Established in April 2013 under the Health and Social Care Act (2012), the HSCIC provides a range of technology and information services and products that are used by patients and service users, the public at large, health and care professionals, and by research, industry and commercial organisations. Its services and products are used extensively to support the commissioning, design and delivery of health and care services and to provide information and statistics that are used to inform decision-making and choice.

The HSCIC has legal powers to hold and process identifiable data, and offers a data linkage service which combines many different sets of data at an individual record level in a secure environment – for example, hospital episode statistics, mental health and mortality data. The data available via this service supports, among others, commissioners, local healthcare providers, and researchers from academia and industry, the latter usually via CPRD. It has also led on the release of healthcare open data, such as prescribing data by GP practice.

Care.data – In Spring 2013, NHS England received approval to commission care.data from the HSCIC. This significant programme will capture and link data from primary and secondary care in order to increase transparency and improve patient outcomes through better design of integrated services, and will incorporate robust information governance controls. Primary and secondary care records will be matched using identifiers such as NHS number, date of birth, gender and postcode. Users other than NHS commissioners or approved researchers will only have access to data in anonymised, pseudonymised or aggregate form. A clear process has been developed for patients to register an objection to the release of their personal confidential data.

Where is the economic value in the UK?

The UK already has some of the world's deepest expertise in research, clinical development, health regulation, and health economics. The UK is also a world leader in releasing open data, in particular in relation to healthcare, creating a unique data-rich environment for innovation and economic growth. These factors provide the UK with potential competitive advantages. However the global competitive landscape is changing fast. Breakthroughs in data and analytics are accelerating change by making the invisible visible. As big data approaches become more mature, datasets, applications and services are expected to multiply, perhaps exponentially. This could create challenges such as diverging data ecosystems in some areas (eg clinical trial services) and converging data ecosystems in others (eg genomic data analysis). The data vision will take time to become reality. The size and complexity of the NHS are both strengths and weaknesses it will take many years to achieve the ultimate vision of easily linkable NHS data within and across the five data domains (described on page 16), and for that to be fully exploited.

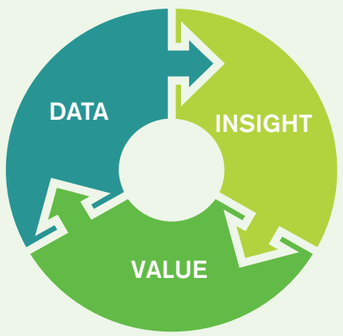
This report is a first step to defining how the UK could maximise the potential for increasing global pharmaceutical industry investment in the UK by leveraging its big data assets. The road map in the next section is underpinned by a number of considerations:

1. The UK has a number of valuable assets which could form the basis for successful long-term competitive differentiation.

The UK could differentiate itself in specific aspects within data, insight and value propositions. In particular the UK could compete globally in analytics services: the market for expert data analytics services is valuable, and services which address complex questions will continue to be in demand regardless of the state of the data. Investing in service standardisation and communication now could keep the UK in a leading global position in the transformation required to shift business models from volume to value.

DATA

- Large volume of deep, longitudinal data covering a diverse population
- Large public investments in genomics, biobanks and integration of disease registries



INSIGHT

- Deep research expertise
- Specialised health data analytics centres

VALUE

- Large scale trials for diverse populations
- Models for commissioning
- Early stratification
- Pharmacovigilance

2. To sustain a competitive advantage the industry needs a stronger, shared data ecosystem across health and life sciences.

Big data needs more than data – a supporting ecosystem also includes skills, services, technical platforms, standards, legal and governance frameworks and financing mechanisms. These components are partly in place now but there is scope for further development and greater visibility to researchers. Skills development at various levels must be an early priority. The maturity of the data ecosystem will also be reflected in a general shift from ad-hoc collaboration toward shared assets and efficient markets for services. This will require a parallel shift in culture across health and life sciences based on open dialogue and common values, grounded in improving patient outcomes. A cultural shift will also benefit individual companies by presenting opportunities to change relationships with the NHS from supplier to partner.

3. The pharmaceutical industry would benefit from focusing on a few specific opportunities.

There are many opportunities for applying big data approaches within and across the pharmaceutical industry and it will not be possible to pursue them all. A joint approach to capability building and demonstrator projects may benefit the UK ecosystem as a whole.

Accelerating the pace of change

The UK healthcare data and analytics ecosystem will play an increasingly important underpinning role for healthcare services and the UK life sciences industries. More needs to be done to accelerate the pace of change, and this will require a co-ordinated approach across multiple stakeholders in the healthcare ecosystem.

4 Road map

Seizing future opportunities

A four-point plan for four years of progress

How is this different?
Similar actions have been identified in the past, but progress has been slow. Any large scale, cross-sector change needs a clear focus and a coherent programme of actions. Slow progress has been due in part to missing 'pieces of the puzzle' – a lack of NHS focus on data for further analysis beyond its primary use, under-developed technical and analytical capabilities or weak mechanisms for cross-industry working. Today key stakeholders are more closely aligned, providing the foundations for a shared programme for big data across healthcare, academia and the pharmaceutical industry.

	Short term (0-1 years)	Medium term (1-2 years)	Long term (2-4 years)
INCREASE AWARENESS	[1] Leadership capability in big data developed		
BUILD CAPABILITY AND CAPACITY		[2] Skills development programmes [3] Virtual 'Big Data Academy' launched	
CREATE SUSTAINABLE DATA ECOSYSTEM	[4] Joint work with NHS data and analytics service providers begun [5] Feasibility of creating a dynamic metadata platform explored		[6] Research tools more widely standardised and shared
ACCELERATE HIGH-VALUE OPPORTUNITIES		[7] Demonstrator projects launched	

DELIVER BETTER PATIENT HEALTH

ENHANCE COST-EFFECTIVENESS WITHIN THE NHS

DRIVE THE INDUSTRY'S CONTRIBUTION TO THE UK ECONOMY

Road map

Increase awareness

Accelerating awareness and understanding of big data is critical to leverage public and private investment in the short-term. The pharmaceutical industry has a wealth of experience and insight that will be key to realising the potential of big data. Proactive initiatives to grow awareness across the ecosystem will assist future engagement between differing organisations.

Build capability and capacity

The emergence of big data analytics and the diversity of the types of data available across the pharmaceutical industry, academia and healthcare services require new capabilities. Technical skills are needed to design and perform the complex analyses inherent in big data applications – but these skills alone will not unlock the full potential of big data. They need to be complemented with business and informatics skills to translate the information into business value. A new generation of informatics/business analysts are required – data scientists who are able to extract and analyse information from large data sets and then present value-added knowledge and insight to non-technical experts.

Develop a strategic big data leadership capability within the pharmaceutical industry

Develop the ABPI's role in fostering increased awareness and understanding of big data across all stakeholders.

Responsibilities could include:

- Co-ordinating initiatives for promoting awareness and knowledge exchange, for example regular multi-stakeholder conferences or workshops involving industry, government, the NHS, academia, research funders, patients and the public
- Publishing a series of big data thought leadership papers, working together with academia, research funders, technology firms, pharmaceutical companies and the NHS to stimulate debate by exploring big data themes in more depth
- Regular dialogue with ABPI members and across ABPI groups to maintain awareness of big data innovations and a platform for industry to access guidance and advice on data governance, frameworks, transparency and openness
- Signposting users to national and international sources of data, including use of a potential metadata dynamic platform

Broaden and accelerate skills development programmes

Data scientists, bioinformaticians, computational biologists, chemoinformaticians, statisticians and business informatics specialists are all in demand. The industry should engage proactively with the Farr Institute, higher education institutes and the research councils to adjust or expand relevant MSc and PhD programmes, accredited undergraduate courses and industry placements. The industry should also work with the HSCIC and Health Education England as they contribute to the professional development of staff within the NHS and social care. The life science sector-specific skills council, Cogent should play a leading role in establishing standards and commissioning high quality training solutions in this area. The ABPI can play a role in facilitating dialogue with industry.

Establish a cross-sector virtual 'Big Data Academy'

A cross-sector task force should explore the establishment of jointly run modules with the NHS, academia and the third sector which form a virtual centre of excellence for developing the common skills and culture needed by industry, health care professionals and managers to harness big data via appropriate training. A range of public and third sector institutions may have an interest in supporting or operating such an initiative and should be invited to join the task force. The aim of the virtual 'Big Data Academy' would be to ensure that individual development within participating organisations aligns to the collection and appropriate use of big data to meet future needs of patients.

Core offerings could include:

- Virtual classrooms for introductions to big data, uses, case studies of collaborative projects that could be built around existing modules currently offered by key stakeholders such as CPRD.
- Workshops (and/or classroom courses) to improve ways of working in a collaborative environment where academia, the NHS and life sciences will have equal interest.

Create a sustainable data ecosystem

The multitude of new information sources remains an underutilised asset. Making data accessible and interoperable is critical to generating greater economic value and to realising personalised medicine. A collective of NHS data and service providers and a wider data service market is fast emerging in the UK and helping to address these challenges, within a robust governance framework to protect patient privacy. There is an opportunity for the industry to become a more strategically engaged customer of NHS data services and to help build a stronger collaborative culture – one that involves stakeholders aligning around the common aim of delivering better value for patients.

Improve 'data and analytics services' in healthcare

As these services are refined and co-ordinated, it will be important for policy-makers, national as well as local data and informatics service providers and research funders to engage with industry as the 'voice of the customer'. This group has the capability to build data linkages, improve services and develop standard methodologies while protecting the public interest.

The potential role of the industry as the 'voice of the customer' could include:

- Agreeing processes and guidelines for industry engagement such as the role of the ABPI's Pharmaceutical Industry Health Information Group
- Providing input for prioritisation of investments in data, linkages and services that build upon the work of the Research Capability Programme.
- Exploring options for emerging governance frameworks, standards, methodologies, principles and services
- Sharing knowledge and experience
- Exploring collaborations and co-investment models.

Explore the feasibility of creating a dynamic metadata platform

A cross-sector initiative should assess the business case and feasibility of establishing a permanent dynamic metadata platform. Seed funding could be sought from the government and research funders. This action could potentially fall within the brief of the data and analytics service providers group mentioned in Action 4 above. The platform would be an evolving product focused on metadata, (data about the data) with the functionality to enable users to identify and assess the data available in the context of their business and clinical questions, as well as the associated governance arrangements to maximise its value. A pilot application could consider exploring a relatively mature data domain in the pre-clinical or clinical space, such as translational medicine research where significant investment has already been made.

Collectively the pharmaceutical industry has massive repositories of data on the biology and natural history of disease, clinical trials, market research and sales. In the world of big data, new research tools are fast being added to this list of assets, such as statistical analysis, tested data and probabilistic models. In theory these could be used to uncover actionable insights about patients, treatment choices by healthcare professionals or the behaviour of payers and providers that could be invaluable across the value chain. Some pre-competitive consortia share research data but not always in a structured way that manages risks and builds industry assets. In general opportunities are constrained by limited common informatics platforms and data standards. Sharing data and information will become increasingly essential to realising the value of big data, for example, the UK has a real opportunity to be a world leader in stratified medicine by shaping the long-term R&D pipeline together with the NHS and academia.

Enhance collaboration in the pre-competitive space

The ABPI and the Farr Health Informatics Research Institute should explore how industry and academic data – as well as emerging analytical research tools such as statistical methods, algorithms and technical platforms – could be shared at greater scale in a clearly defined pre-competitive space. It is important that this is carried out in partnership in order to accelerate learning, support data transparency and enable collaboration within and across sectors.

Accelerate high-value opportunities

There are many potential applications for big data but both the pharmaceutical industry and wider healthcare organisations are still in the infancy of leveraging these for business and clinical use. Evidence of which applications generate the most value is currently limited. Demonstrator projects are needed to explore options and demonstrate proof of principle.

7 ACTION

Establish a Joint Demonstrator Projects Portfolio

The ABPI intends to facilitate proof of concept projects for further joint investigation of the application of big data techniques in high value areas such as genomics, interventional trials, observational studies, research into patient reported outcome measures and pharmacovigilance. Industry oversight of these projects would be provided by the ABPI's Pharmaceutical Industry Health Information Group and emphasis will be placed on cross-industry learning and development of project governance and methodologies. It is hoped that these projects would attract funding from across sectors.

Two suggested projects are described on the next page.

Typical questions during project definition could include:

- Which big data challenges are we seeking to explore?
- Which key strategic questions are we asking?
- Which conditions or patient cohort(s) should we consider?
- Which datasets are needed?
- Which datasets do we already have?
- How can we get the data and interpret it correctly?
- Where are the gaps in the data and how can these be filled?
- Who should be involved?
- What skills and methodologies do we need?
- What are the legal and ethical principles and constraints?

Candidates for joint demonstrator projects

INCREASE SAFETY:
Enhancing pharmacovigilance and facilitating regulatory innovation

Investigate the use of expanded and integrated NHS data (eg care.data), linked to additional sources of data such as disease registries and international databases, for near real time monitoring of patients in selected situations such as the Early Access scheme or adaptive licensing. A successful pilot project could open up options for the Early Access scheme and/or adaptive licensing by enhancing monitoring of patients involved in these potential new routes for providing earlier access to innovative medicines. It would also create a world leading tool for pharmacoepidemiology to assess potential safety signals. The feasibility of the project depends on increased coverage of GP data, links to hospital data and/ or joint working with disease registries in specific clinical areas, as appropriate.

ACCELERATE STRATIFIED MEDICINE:
Modelling outcomes

Develop a risk-based model for a chronic or complex condition to identify cohorts of patients who might benefit from different combinations of treatment and so shift the understanding of treatment value and identify opportunities for drug discovery. The project may require a wide range of 'big' and 'small' clinical and genomic data, for example laboratory tests and tissue samples, which would provide an opportunity to develop and test more efficient processes and governance arrangements for such complex, targeted studies.

whilst exploring how to...

- Improve the interoperability of disparate datasets which lack standardisation
- Develop the methodologies for searching and linking multiple data sets
- Leverage existing data governance arrangements and develop new frameworks
- Develop appropriate governance to protect personal data in smaller patient populations.

Road map:
Case study

A global alliance for a data-sharing future

In June 2013 over 70 leading healthcare, research, and disease advocacy organisations across 40 countries announced that together they have taken the first steps to form a not-for-profit "global alliance to enable responsible sharing of genomic and clinical data".

The idea of an alliance was inspired by the World Wide Web Consortium which has played a critical role in the development of the global open technology standards that underpin the World Wide Web. The alliance is based on the belief that the public interest will be best served if institutions and governments work together to develop and promulgate standards (both technical and regulatory) that make it possible to share and interpret this wealth of information in a manner that is both effective and responsible – that it must respect patient and participant autonomy, enable collaboration and comply with regulatory, governance and ethical frameworks. Initial signatories in the UK included the Wellcome Trust Sanger Institute, NIHR, Cancer Research UK and Oxford University.

The vision is that ultimately data can be stored in platforms built on interoperable standards. Whether participants put all or some data into these or other platforms will be their own decision, but the common vision for the platforms is not that they will be a shared repository for data; rather, the platforms will enable sharing and learning from data wherever it is stored.

Overcoming challenges

The alliance needs to overcome barriers and mitigate risks related to regulation, ethics, technology and public opinion, and maximise engagement with prospective partners.

Public attitudes toward sharing

The project recognises a need to maintain a diversity of approaches to engaging the public both nationally and internationally in order to encourage public and patient participation. The signatories committed to establish a framework where participants will have the right to share genomic and clinical information as broadly or narrowly as they are comfortable with, including not at all.

Regulation and ethics

The alliance will identify which legal frameworks and ethical considerations must be understood by all members. As such, an interoperable platform has been proposed that will support storage and control of data that encompasses informed consent, patient privacy, protection of research data and electronic data privacy. The primary stakeholders of the alliance propose to engage relevant government authorities to improve policy harmonisation, taking advantage of its own stakeholder knowledge and best practices to achieve this.

Technology

The alliance recognises a need to provide open standards and a shared platform in order to facilitate comparisons and meta-usage. The platform itself will be designed to work via more than one technology vendor (eg Google, Microsoft, Amazon) and allow for interfaces to develop third party apps. US NASDAQ financial exchange cloud storage and credit card industry cloud storage have been investigated as examples of maintaining data security with multiple stakeholders. The alliance also recognises that it is easier to contribute to other efforts rather than duplicate them for some future uses of this data.

Value creation

The alliance foresees potential benefits both for human health and the wider biomedical ecosystem.

Impact on human health

At present it is generally not possible to predict which changes to DNA sequence lead to clinical consequences. Interpretation of individual genome sequences could be enabled by comparison to extensive data on variation in genome sequence and phenotype. In its published White Paper, the alliance defines new opportunities to gain insight into disease, improve prevention and detection, define diagnostic categories, streamline clinical trials and match patient to therapy.

Impact on the wider biomedical ecosystem

The alliance expects a positive impact beyond patients; scientific researchers, hospitals, life sciences, governments and disease advocacy organisations should all benefit from trusted, turn-key solutions for searching, storing and analysing a wide range of data sources.

5 Conclusion

Next steps

This document highlights some of the key issues facing the healthcare and pharmaceutical sectors, and argues that big data can play an important role in helping to overcome them. The road map lays out a set of actions for the next one to four years, requiring productive collaborations and partnerships for success. Some of the recommendations will require more financial investment over the longer term, but the payoff for both industry and the public is likely to be significantly higher.

We also need to continue to scan the big data horizon, identifying new opportunities where innovative technologies may play an emerging role. This will ensure focused investment and contribute to a sustainable cycle – identifying and prioritising next-generation high-value opportunities.

The key to generating long-term value is honest and open dialogue and collaboration between all stakeholders. This goes beyond bilateral or multilateral initiatives. A sustainable data ecosystem across health and life sciences is a requirement for the success of the UK in a global knowledge economy: at stake are the economic health of the UK pharmaceutical industry, the quality of healthcare services being delivered and ultimately the health of the public.

Association of the British Pharmaceutical Industry
7th Floor, Southside, 105 Victoria Street, London SW1E 6QT
t +44 (0)870 890 4333 abpi@abpi.org.uk

abpi.org.uk