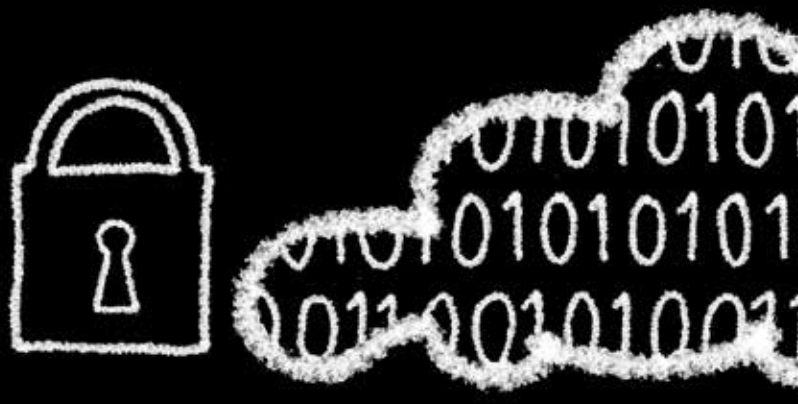
Big data has taken the business world by storm and it's coming to the NHS. While some worry about the implications for patient privacy, others believe it can only help improve care. Seamus Ward asks…



…what's the big idea?

You've probably heard the phrase – big data – and maybe dismissed it as one of those ideas that comes from commerce, then burns brightly but briefly in the minds of a select few in the NHS before making way for the next fad. But with so much information in the NHS, clinicians and managers are beginning to see that big data could be a way of understanding trends and modelling changes to improve the quality of care and transform services to patients.

Big data is used widely outside the health service. Meteorologists draw up weather forecasts; supermarkets seek to understand customer spending habits; and big data techniques are used in large hadron collider experiments and by national security agencies to analyse intelligence.

The NHS Confederation says there is no doubt that big data analytics have huge potential to better match healthcare provision with need. Industry commentators have wondered if big data could help clinicians spot the early onset of sepsis, which kills an estimated 37,000 people a year in the UK, at a cost of £2.5bn. In July this year, a blog on IBM's *Building a smarter planet* website wondered whether big data could help cure cancer.

Big data, big claims. But what is it? Information technology research and consultancy firm Gartner first defined big data in 2001 (see box). It said it had three dimensions, the three Vs – volume, velocity and variety. Some organisations have since added a fourth V – veracity – but the original definition is most widely used.

While some see big data as the management of ever larger volumes of data, many believe it is all about using large volumes of information from many sources to perform predictive analysis quickly.

In healthcare, it could be used to analyse data on how, say, a diabetic patient is cared for – such as the tests employed, drugs prescribed, number of appointments, skill mix of the clinicians, compliance with blood sugar target levels and blood pressure variations.

With this mass of information from primary, community and secondary care sources the elements of care that produce the best outcomes could be identified. Or, by changing one or more of the variables, the impact of new pathways of care could be modelled.

Big data in the NHS

Mede Analytics clinical director Mark Davies says there are many examples of big data in the NHS. 'Big data isn't just hype. It isn't just a matter of scale. It's a different relationship between the health service and the information that underpins it,' he says.

'People tend to think it's divorced from real healthcare, but it's increasingly about understanding how you do the business of healthcare. It's not going to go away. It's the building block for making effective decisions and planning of care.

'It creates all sorts of cultural challenges, but traditionally there has been a big cultural distance between finance directors, medical directors and chief nursing officers, commissioners and providers. In a more joined up world, that distance needs to close – and it's the same for data.'

Graham James, vice-president of CACI's business intelligence group, says the NHS is beginning to look at big data.

'We are starting to see a demand from the commissioning side – probably more from support units than clinical commissioning groups. To support better commissioning and for the outcomes and delivery of health services to improve, they have to have a big data project supporting their decision-making.

'The vision is there, but so far the delivery has not been there. But I always believe the NHS will have the wherewithal to deliver.'

He adds that the information must extend across care pathways, covering hospital and GP data by outcomes, for example. For planning purposes, it should include demographic details that show how a population's health needs might change over time.

Mr James believes big data could also help in the drive for cost and quality improvements. 'It offers a lot to the NHS, but there is still a long way to go before we have true big data projects and tools delivering for the NHS.'

There are examples of big data projects around the NHS. CACI currently has a bid in with NHS England to develop a big data set in one local health economy, Mr James says. This will bring together GP, hospital and social care data to allow better planning of health and care services.

There are arguably elements of big data in the Health and Social Care Information Centre

SHUTTERSTOCK

(HSCIC) secondary uses service database, which feeds the hospital episodes statistics, as well as in Derby Hospitals NHS Foundation Trust's theatres costing and stock management system (see page 12).

Liverpool collaborative

In the North West, the Liverpool Big Data Collaborative has drawn together Royal Liverpool and Broadgreen University Hospitals NHS Trust, Liverpool Clinical Commissioning Group, two of its universities and the North West Coast Academic Health Science Network.

Formally launched in March, the collaborative says a range of potential benefits have been identified, including the use of stratified medicine – where diagnosis and treatment is personalised based on an individual's genetic profile – and potentially offering solutions to problems such as preventable readmissions.

'Historically the NHS works in silos – inpatient, outpatient, never looking across the whole patient journey,' says David Hodson head of information at East and North Hertfordshire Clinical Commissioning Group, which is working on a big data project with Mede. The aim is to establish a system through which the CCG can begin to link data across different sources.

'We want to make sure this is being done in a legal way that satisfies all the requirements about how we use that information, brings patients along with us and integrates with other services, including the community trust and social care,' he says, adding that the next step is to incorporate mental health data.

"Big data isn't just hype. It isn't just a matter of scale. It's a different relationship between the health service and the information that underpins it"

Mark Davies, Mede Analytics

Work has been under way since the start of the financial year, though a group of local health and social care organisations have been discussing integrated data for about a year. The area has been working with Mede for about three years on linking different data sources.

The project will cover around 1.1 million people, all with patient records. Although he is unsure as yet how many social care records the system will have to deal with, Mr Hodson says it will be integrating data from one million outpatient records each year, together with information from 300,000 A&E attendances.

Having collected the data, the CCG can then interrogate it using risk stratification techniques – for example, to identify the most vulnerable patients.

Mr Hodson continues: 'Ideally, we'd like to link to GPs so they could start to see what's happening to their patients with long-term conditions such as diabetes or chronic obstructive pulmonary disease and start redesigning the services they are getting.

'We will be able to see who the high users are and whether we can deliver better services for those patients. That's why we are doing this. It's not to create a huge dataset – the whole reason is to see clinical benefits,' he adds.

Mede director of product development Kevin Truby says big data could show the NHS the way to harness the power of the information it collects.

'We are not short of data. The NHS is awash with data. But with big data we are looking at how you start to explore and use this data for meaningful purposes,' he says.

'Historically, finance teams used their data for a single purpose. Commissioning teams used similar data for a different purpose and clinicians used another set of data for a clinical purpose. In reality, all that data was pretty common, so we are bringing that together with patient records and making it accessible to a wider range of users,' Mr Truby continues.

'End users do not have to resort to sending

a request to the information department for a report on trauma and orthopaedic costs in the last six months. Effectively, this is self-service.'

Big data could also facilitate the integrated care agenda. 'With integrated care you need integrated information,' says Mr Truby. A good example of this is to link patient records from all sources. 'It means that we know a patient went to their GP surgery on Monday morning and attended A&E and was treated as an emergency admission on Tuesday afternoon. We know they stayed in hospital for 17 days, the drugs they were given, their next outpatient appointment and the social care package.

'That sort of information is useful to a GP or to someone doing contract modelling and thinking of setting up a new clinic.'

Dr Davies believes big data could give NHS organisations a broader, more patient-centred view. 'With the acceleration of NHS and social care costs, we cannot continue to do things the way they have always been done. Any exercise around service redesign or reconfiguration needs effective business intelligence to make the right decisions and to bring the public with us.'

Further benefits will be gained when clinical outcome measures are added, Hertfordshire's Mr Hodson says. 'We've held initial discussions with patient groups about this; what they think would be useful measures and can we start to collect some of that information.

'Both the county council and health organisations are keen to deliver this. We realise we can't work in these silos any more.'

Privacy concerns

The elephant in the room when it comes to big data, particularly for patients, is privacy and data security. This was clear from reactions to care.data, NHS England's ambitious project to give the NHS a complete picture of health and social care by linking patient records from a variety of sources. The plan is to create a new care record from the information gathered and although NHS England insists it will not include information that identifies an individual, many patients and professional bodies have been concerned that pharmaceutical companies or even insurers would be able to access identifiable data.

In the face of public distrust, NHS England delayed the scheme by six months until this autumn and has launched a governance review.

Local projects are aware of the sensitivity over patient data. Mr Hodson says the East and North Hertfordshire initiative will be clear that personal patient information is not identifiable and is pseudonomised.

'From my perspective and from talking to others, the indications are that we don't need identifiable data,' he says. 'Only the clinicians involved directly in the care of a patient needs access to identifiable data.'

The HSCIC, which will manage care.data, is also aware that many patients need to be convinced about data security. In a statement, it says: 'The valuable work that the HSCIC does for the health and social care system needs the endorsement of the public if it is to be effective. Our chief executive, Andy Williams, recently stated how important it is that people are certain their choices will be followed, that clinicians feel supported in their roles and for data users to know where they are with us.'

It continues: 'The HSCIC board has agreed a programme of work that includes active communication with the public and patients. Public representatives will also be part of a new membership of the HSCIC's data oversight committee, the Data Access Advisory Group.

'The HSCIC is absolutely committed to keeping all data it handles safe and secure and applies the same principle to any data that is shared outside of the organisation. The new provisions in the recently passed *Care Act* clarify that the HSCIC may only share data where it would be for the provision of health or adult social care or the promotion of health.'

Security is a big issue, but if NHS organisations can ensure patient privacy and assure the public, big data could play a role in transforming the delivery of NHS services.

A brief history of big data

While humans have probably been worried about how they will be able to understand and use all the information they have gathered and written down since they built the first bookshelves, big data is a very modern phenomenon.

Initially, it referred to the quantity of information. With knowledge growing exponentially, by the mid-1980s scientists and IT specialists have been concerned with storing, searching and analysing the information. By the late 1990s electronic storage became cost-effective and at that time the concept of big data emerged – probably just in time, as internet use took off. In 2001 Doug Laney, an analyst with the Meta Group (now Gartner) put forward a definition of big data, known as the three Vs:

- Volume – the large amount of data
- Velocity – how quickly the information is generated and analysed

- Variety – it should come from a number of different sources.

Some organisations have since added a further V for veracity (the accuracy or reliability of the data).

Now, big data is used in many areas, from science – it reportedly helped reduce the time taken to decode the human genome from the original 10 years to less than a day – to commerce, where supermarkets and online retailers analyse customer habits to inform marketing campaigns or even to decide which products to sell in individual stores.