OPEN ACCESS

Making sense of the shadows: priorities for creating a learning healthcare system based on routinely collected data

Sarah R Deeny, Adam Steventon

Data Analytics, The Health Foundation, London, UK

Correspondence to
Adam Steventon, Data Analytics, The Health Foundation, 90 Long Acre, London WC2E 9RA, UK; adam.steventon@health.org.uk

ABSTRACT

Socrates described a group of people chained up inside a cave, who mistook shadows of objects on a wall for reality. This allegory comes to mind when considering 'routinely collected data'—the massive data sets, generated as part of the routine operation of the modern healthcare service. There is keen interest in routine data and the seemingly comprehensive view of healthcare they offer, and we outline a number of examples in which they were used successfully, including the Birmingham OwnHealth study, in which routine data were used with matched control groups to assess the effect of telephone health coaching on hospital utilisation.

Routine data differ from data collected primarily for the purposes of research, and this means that analysts cannot assume that they provide the full or accurate clinical picture, let alone a full description of the health of the population. We show that major methodological challenges in using routine data arise from the difficulty of understanding the gap between patient and their 'data shadow'. Strategies to overcome this challenge include more extensive data linkage, developing analytical methods and collecting more data on a routine basis, including from the patient while away from the clinic. In addition, creating a learning health system will require greater alignment between the analysis and the decisions that will be taken; between analysts and people interested in quality improvement; and between the analysis undertaken and public attitudes regarding appropriate use of data.

In the famous cave allegory from Plato's *Republic*, Socrates described a group of people chained up inside a cave, with their heads secured in a fixed position to view the wall in front of them.[1] A fire was burning, and various people passed between this fire and the captives. These people carried objects, creating shadows on the wall that the captives mistook for the objects themselves. This is an apt allegory to keep in mind when considering 'routinely collected data'—the varied, and massive, person-level data sets that are generated as part of the routine operation of modern healthcare services.

Huge enthusiasm exists for routine data, which offer a seemingly panoramic view of healthcare.[2 3] The hope is that these data can identify ways to improve the quality and safety of healthcare, and thus lead to a learning healthcare system.[4 5] Yet, routine data provide only vague shadows of the people and activities they represent. Routine data are generated for the purposes of delivering healthcare rather than for research, and this influences what data are collected and when. Thus, analysts of routine data cannot assume that they provide the full or accurate clinical picture, let alone a full description of the health of the population.

Our research group uses routine data to improve decision making in health and social care. Many successful studies show that data shadows can be interpreted intelligently to improve care. However, the partial picture afforded by routine data means that there are considerable challenges to achieving a learning healthcare system,[6] and these may explain why 'big data' seems to have attracted as much scepticism as enthusiasm.[7–9] In this article, we discuss the need to improve the information infrastructure of healthcare systems and to understand better the factors that influenced what data were

collected. Matching the rapid growth in the availability of large data sets with knowledge about how they were generated will require supplementing big data with 'small data', for example from clinicians, patients and qualitative researchers. Creating a learning healthcare system will also require greater alignment between analysts and people interested in quality improvement, to ensure that data analytics provides supportive decision making while being sufficiently embedded within the healthcare system and maintaining public trust in the use of data.

TYPES AND USAGE OF ROUTINE DATA

It will be immediately apparent to anyone who seeks to use routine data that most healthcare systems do not contain a single comprehensive data set, but a patchwork quilt.[10] From our experience of the English National Health Service, we have found that the data that are most often used tend to be those that were generated by administrators or clinicians. However, patient-generated and machine-generated data are also important and the technologies that produce them are growing in availability (table 1).

Routine data can be contrasted with data that are collected primarily for the purposes of research, for example for a randomised controlled trial.[31] Data that are collected for a research project can be tailored to the specific populations, timescales, sampling frame and outcomes of interest. In contrast, analysis of routine data constitutes secondary use, since the data were originally collected for the purposes of delivering healthcare and not for the analytical purposes to which they are subsequently put. In our discussion of routine data, we exclude data collected as part of surveys (eg, the Health Survey for England[32]), and some clinical audits,[33] even if these are frequently conducted. The specific difference is that these have a prespecified sampling frame, whereas individuals are included in routine data by virtue of their use of specific services.

Routine data have been used successfully in comparative effectiveness studies, which attempt to estimate the relative impact of alternative treatments on outcomes.[34] For example, one recent study examined the impact of telephone health coaching in Birmingham, a large city in England. In this example, patients with chronic health conditions were paired with a care manager and worked through a series of modules to improve self-management of their conditions.[35] Although the intervention operated in community settings, the study was able to link various administrative data sets together to track the hospitalisation outcomes of 2698 participants over time. Because many of the patients had been hospitalised previously and this group may show regression to the mean, a control group was selected from similar parts of England using a matching algorithm that ensured that controls and health coached patients were similar

with respect to observed baseline variables, including age, prior rates of hospital utilisation and health conditions. This study found that the telephone health coaching did not result in reductions in hospital admissions, a finding that was robust to reasonable assumptions about potential impact of unobserved differences between the groups.[35] Another study of the same intervention, which used similar methods, uncovered improvements in glycaemic control among a subset of patients whose diabetes had previously been poorly controlled.[36]

The Birmingham study shows that routine data have some technical benefits (such as wider coverage, longitudinal nature and relatively low levels of self-report bias for some end points).[37] However, one of the major advantages of these data is simply that they are available on a routine basis. This means that, for example, matched control analyses could be repeated on a regular basis to understand how the effectiveness of services change over time. This is not the only tool based on routine data that clinicians and managers can use to measure and improve the quality and safety of healthcare in a learning healthcare system.[6 38] Other examples include:

▶ Predictive risk modelling, which uses patterns in routine data to identify cohorts of patients at high risk of future adverse events;[39 40]
▶ Analysis of the quality of healthcare, such as the identification of cases where there was opportunity to improve care;[41] and
▶ Surveillance of disease or adverse events, such as to detect outbreaks of infectious disease.[17]

In theory, all the types of routine data shown in table 1 can be used in any of these applications but a recurrent concern is whether the data are sufficient to address the questions of interest. In the next sections, we discuss some of the general limitations of routine data, first in terms of the scope of what is included, and then in terms of how the data are assembled.

THE SCOPE OF ROUTINE DATA: INTRODUCING THE DATA SHADOW

Figures 1–3 illustrate some general features of healthcare systems that influence how these data are commonly collected. We focus of administrative and clinically generated data, as these have been most commonly used to date.

Figure 1 begins by taking the viewpoint of a typical healthcare system, and shows the information that might theoretically be exchanged as part of its day-to-day operation. At the centre of the diagram is the clinical encounter, during which symptoms, observations, treatment options, preferences and outcomes may be relevant. Outside of this encounter, there is generally one ongoing dialogue between patients and their support network (the left hand side of the diagram), and another between the clinician and the wider care team (right hand side). A range of factors,

Table 1 Types of routine data in healthcare

Data type	Definition	Characteristics	Examples
Administrative data	Data collected as part of the routine administration of healthcare, for example reimbursement and contracting. Secondary uses include the assessment of health outcomes and quality of care.	Records of attendances, procedures and diagnoses entered manually into the administration system for a hospital or other healthcare organisation and then collated at regional or national level. Little or no patient or clinician review; no data on severity of illness.	*Hospital episode statistics (England):* Clinical coders review patients' notes, and assign and input codes following discharge. These codes are used within a grouper algorithm to calculate the payment owed to the care provider.[11] HES data have been used to generate quality metrics, including hospital standardised mortality indicators.[8] Hospital episode statistics (HES) data have also been used to build predictive risk models, for example to allow clinicians to identify cohorts at risk of readmissions,[12] or to allocate scarce resources in real time.[13]
Clinically generated data	Data collected by healthcare workers to provide diagnosis and treatment as part of clinical care. These data might arise from the patient (for example, reports of symptoms) but are recorded by the clinician. Secondary uses include the surveillance of disease incidence and prevalence.	Electronic medical record of patient diagnoses and treatment. Results of laboratory tests. Compared with administrative data, less standardised in terms of the codes used and less likely to be collated at regional and national levels.	*Electronic medical record:* More than 90% of primary care doctors reported using the Electronic Medical Record (EMR) in Australia, the Netherlands, New Zealand, Norway and the UK in 2012.[14] Linked EMR data have been used in Scotland to create a prospective cohort of patients with diabetes. In addition to being used to integrate patient care, they have been used in research to estimate life expectancy for the patient cohort.[15] An evolution of the EMR (an electronic physiological surveillance system including improved recording of patients' vital signs) was used to calculate early warning scores that led to a reduction in mortality as part of an advanced predictive risk model.[16] *National and regional microbiological surveillance system (UK):* Results of clinical tests ordered by clinicians are recorded at a laboratory level before being reported regionally and nationally. There is mandatory reporting of certain infections and organisms (eg, *Clostridium difficile*) and voluntary reporting of others. These data are interpreted using automatic statistical algorithms to detect outbreaks of infectious disease.[17]
Patient-generated data (type 1: clinically directed)	Data requested by the clinician or healthcare system and reported directly by the patient to monitor patient health.	Data collected by the patient on clinical metrics (eg, blood pressure), symptoms, or patient reported outcomes. Choice of data directed by the healthcare system.	*Swedish rheumatology quality registry:* Uses patient reported data as a decision support tool to optimise treatment during routine clinic visits and for comparative effectiveness studies. These data have also been used to examine the impact of multiple genetic, lifestyle and other factors on the health of patients.[18] *Telehealth:* For example, patients with heart failure are asked to supply information on weight or symptoms on a regular basis, using either the telephone[19] or Bluetooth-enabled devices[20]
Patient-generated data (type 2: individually directed)	Data that the individual decides to record autonomously without the direct involvement of a health care practitioner, for personal monitoring of symptoms, social networking or peer support.	Symptoms and treatment recorded by the patient. Recorded outside the 'traditional' healthcare system structures.	*Patients like me:* An online (http://www.patientslikeme.com) quantitative personal research platform for patients with life-changing illnesses. A cross-sectional online survey showed that patients perceived benefit from using these networks to learn about a symptom they had experienced and to understand the side effects of their treatments.[21] Similar platforms exist for mental health[22] and cardiology.[23] *Individual and patient activity on social media*: Analysis of key terms on Twitter has been used to monitor patient outcomes and perception of care. No clear relationship between Twitter sentiment and other measures of quality has been shown.[24 25] There has also been an attempt to use search engine usage (Google) to track and predict flu outbreaks but to date there has been no demonstrated public health benefit.[26]

Continued

Table 1 Continued

| Data type	Definition	Characteristics	Examples
Machine-generated data	Data automatically generated by a computer process, sensor, etc, to monitor staff or patient behaviour passively.	Record of individual behaviour as generated by interaction with machines. The nature of the data recorded is determined by the technology used and substantial processing is typically required to interpret it.	*Indoor positioning technologies*: Sensors have been used to record the movement of healthcare workers within out-of-hours care.[27] A recent study used sensors on healthcare workers and hand hygiene dispensers to show that healthcare workers were three times more likely to use the gel dispensers when they could see the auditor.[28] *Telecare sensors*: Telecare aims for remote, passive and automatic monitoring of behaviour within the home, for example for frail older people.[29] A Cochrane review on 'smart home' technology found no studies that fulfilled the quality criteria,[29] and a larger, more recent randomised study has failed to demonstrate a positive impact of this approach on healthcare usage.[30]

including clinical practice, influence what information is shared and recorded.[42]

Analysts of routine data face particular problems because, in most healthcare systems, the information infrastructure is better developed on the right hand side of figure 1 than on the left hand side. Thus, information systems gather data from the clinician during the encounter (eg, their observations from the patient) and from the wider care team (eg, results of diagnostic tests) but, for most patients, there is no systematic way of communicating what happens outside of the clinical encounter, including symptoms, response to treatment, and the priorities of patients and their families,[43] despite the potential value of this information for both health and healthcare.[42 44] As a result, analyses of routine data often focus on clinical or process outcomes, and there is a risk that the perspective of patients might be forgotten or undervalued. However, there are examples of healthcare systems that combine patient-generated data with administrative and clinically generated data into research databases or registries.[18]

Another feature of healthcare systems is that they are more likely to interact with people who have concerns about their health than those who do not. Moreover, the lens of data collection is typically focused on the patient, rather than on other groups who might be instrumental to improving health outcomes, such as informal carers.[45] This feature leads to another limitation to the scope of routine data: from a population perspective, some groups are described in much more detail than others (figure 2). For example, in England young men visit a GP (General Practice, denotes primary care practice in the UK) much less often than women and older men,[46] meaning that less information is available on this group.

These two features mean that the picture shown by routine data can differ markedly from the actual experience of an individual.

Figure 3 (bottom half) shows the information that is typically obtainable from routine data for a hypothetical individual, and would be traditionally included in a theogram.[47] Underlying each of the contacts shown, information might be available about the person's recorded diagnoses, test results and treatments. This is what we call the *data shadow*[48] of the person, because it is an imperfect representation of the person's lived experience (top half of figure 3). The data shadow can omit many aspects of a life that the person considers significant, including some that may influence health outcomes, such as: whether the person lives alone; their drug, tobacco and alcohol intake; the breakdown of a relationship; financial hardship; employment status; violence; and other stressors over the life course.[49] These observations are not merely theoretical and can pose specific challenges to the analysis. For example, it can be difficult to assess long-term outcomes following surgery because,
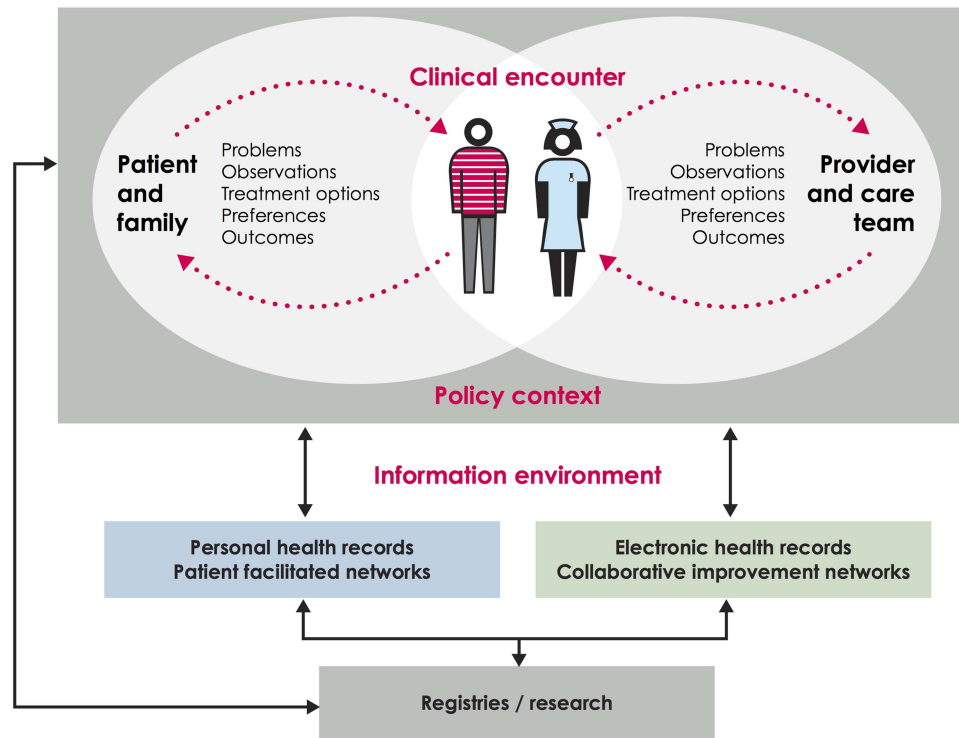
Figure 1 The first boundary of routinely collected data. This illustrates the relationship and distance between data recorded by the patient or family and data recorded by clinician and administrative health records. Cross referencing with table 1, administrative and clinically-generated data generally lie on the right hand side of this figure, while the two types of patient-generated data lie on the left hand side. Based on work by the Dartmouth Institute and Karolinska Institutet.

as people recover, they interact with the system less often and so less information is recorded.

Figure 3 shows that data linkage may address some of the limitations of routine data.[47] While valuable information might be missing from healthcare data, in some cases this might be obtained from other governmental, social and commercial data sets.[50] For example, approximately five million (as reported May 2014 51) working-age people claimed some kind of social security benefit in Great Britain, and thus records are theoretically available on their socioeconomic situation, housing and previous taxable earnings.[51] Innovative examples of data linkage include combining mobile phone usage with police records to show the impact of phone calls on collisions while driving,[52] and linking social care with hospital data to predict future usage of social care.[40] Although data linkage may shed light on some of the social determinants of health, including employment status, housing, education, socioeconomic deprivation and social capital,[53] not all life events will be recorded. Thus, regardless of their sophistication, routine data will only ever represent shadows of individuals' lives.

UNDERSTANDING THE DATA GENERATING PROCESS

In our opinion, many of the major methodological challenges in using routine data have at their root the difficulty of understanding the data generating process. We define this broadly to consist of the relationships between the factors that influenced why some data were recorded and not others.[54] This relates to the choice of diagnoses codes used, and to the factors that influenced whether a person was seen by the healthcare system and which treatments were received. We now give two examples of how a lack of clarity about the data generating process can result in biased analysis,[54] before describing in the following section how this can be overcome.

The telephone health coaching studies have already illustrated the risk of selection bias that can occur if, for example, the patients enrolled into a programme have poorer glycaemic control than those who are not,[36] when glycaemic control is also associated with outcomes.[55] Thus, a straightforward comparison of the outcomes of people receiving health coaching versus not receiving it will yield a biased estimate of the treatment effect. Although matching methods can adjust for observed differences in the characteristics of people receiving health coaching and usual care, the important characteristics must be identified, so successful application of these methods requires an understanding of why some patients received the intervention and others did not.[54]

The second example relates to postoperative venous thromboembolism (VTE), a potentially preventable cause of postoperative morbidity and mortality. VTE
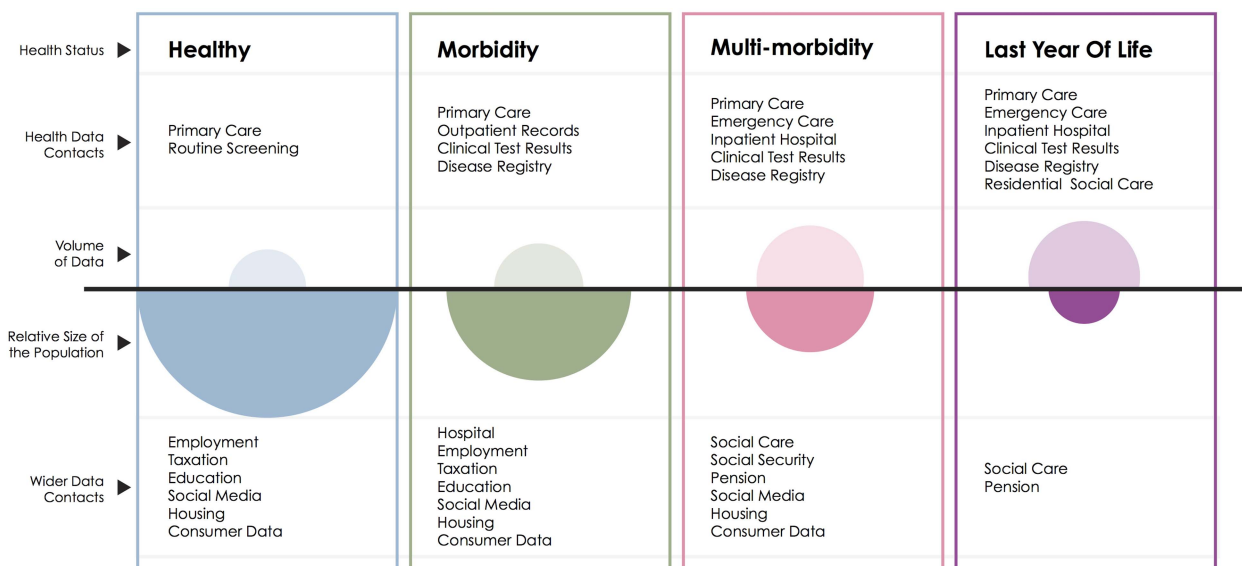
Figure 2 The second boundary of routinely collected data. This schematic illustrates the volume of data available (open ovals) from health datasets for each group of the population defined by health status in contrast with the size of that population (filled ovals).

rate is often used in pay-for-performance programmes,[56] but some clinicians have a lower threshold in ordering tests for VTE than others, while some hospitals routinely screen asymptomatic patients for VTE. Therefore, if one hospital has a higher rate of VTE than another, this does not necessarily mean that they offer less safe care, as they might just be more vigilant.[56] Unless the analyst understands this feature of the data, a particular type of bias will result, called surveillance bias.[57 58]

Information about the data generating process is arguably less important in other applications. For example, when developing a predictive risk model, the primary aim might be to make accurate predictions of future risk, and the data generating process might be considered less important as long as the predictions are valid.[59] However, in general, robust analysis of routine data depends on a good understanding of how the data were put together. This has implications for both research methods and the future
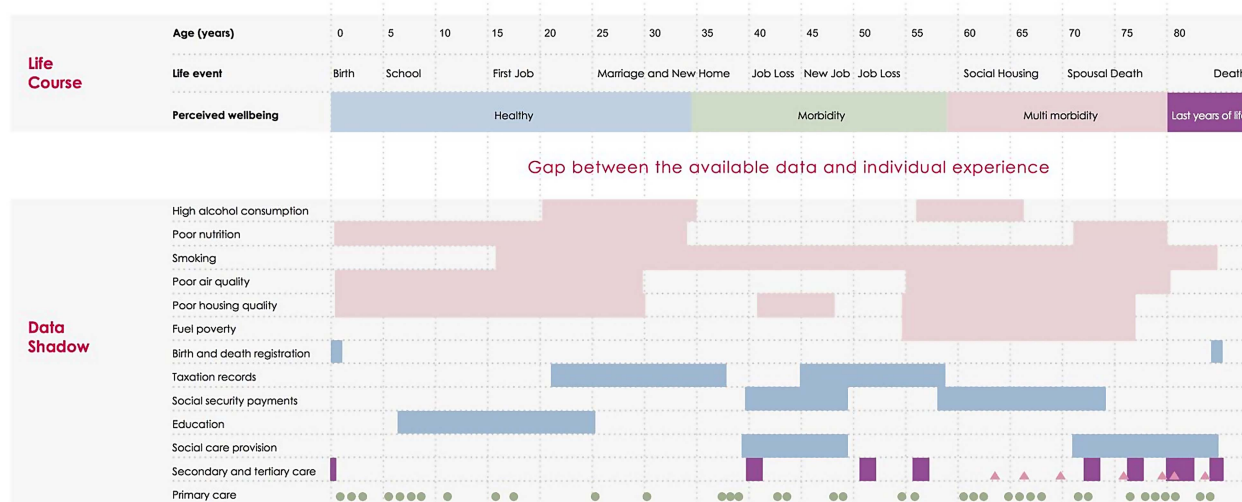
Figure 3 Comparison between the individual experience and the data shadow: We map the life course of a hypothetical individual, showing age in 5-year increments, important personal events and their personal perception of health. In the first six rows of the 'Data Shadow' section we indicate behavioural and environmental factors that may be damaging to health, with periods of time of exposure illustrated with solid rectangles. In the following five sections we indicate governmental data sources that contain further information relevant to the health of the individual, with periods of data collection indicated by rectangles. The final three categories give a timeline of patient contact with care services (shown as a theograph[47]); this uses triangles, circles and lines to indicate different types of health care (primary, secondary and tertiary care) or social care provision. Small circles indicate a GP visit, rectangles a period in secondary, tertiary or social care and triangles an attendance at an emergency department.

development of the information infrastructure in healthcare systems, as we now discuss.

MAKING SENSE OF THE SHADOWS

Many of the efforts to address the limitations of routine data have focused on developing and applying more sophisticated analytical methods. These don't attempt to improve or enhance the data but, seek to make inferences with lower levels of bias compared with other methods. For example, a variety of analytical methods exist for dealing with selection bias[60–64] and surveillance bias.[58 65] For example, a recent study[58] proposed a proxy for 'observation intensity' when dealing with surveillance bias, namely the number of physician visits in the last 6 months of life. It offset the effect of variation in this proxy from the variation in the recorded outcomes. Likewise, many methods have been developed to deal with selection bias, including machine-learning methods such as genetic matching,[66] and classification and regression trees.[67] Multiple imputation is now an established technique for analysing data sets where there are incomplete observations,[68] while Bayesian inference,[69] hidden Markov models[70] and partially observed Markov process[71] have been used to overcome deficiencies in data collection when fitting epidemiological models to routine data. Many other examples exist.

While these developments are helpful, any analytical method will be biased if the underlying assumptions are not met, so it is still important to address deficiencies in the underlying data.[54] Relatively few studies have examined the relative strengths of the different types of data listed in table 1 in terms of dealing with common sources of bias.[39 43] As we have already seen, it may be possible to obtain additional information through data linkage, and thus reduce the gap between a person and their data shadow.[15 72–75] However, data linkage increases the volume of data available, and so may compound rather than alleviate problems associated with understanding the data.

Although data linkage and advanced analytical techniques are valuable, it seems that, in parallel with these efforts, the information infrastructure of many healthcare systems could also be improved to capture more data relevant to current and future health outcomes (figure 1) and to identify more populations of interest (figure 2). Many outcome measures are available to choose from, but key questions are about how these should be used within the care pathway, where and when they should be collected, and how to put in place a system to collect them.[44] There is also a need to balance the benefits of standardising data collection with the heterogeneity in patient and provider preference about what is recorded and when.[76] Although some patients might prefer to submit data electronically, others might prefer to use more traditional methods. This was illustrated in a recent telehealth trial, in which patients were given home-based technology to relay medical information to clinicians over Bluetooth, but some still took their blood sugar information to the GP on paper.[76] Showing the need to work with both clinicians and patients to develop a method of data collection that serves their needs as well as providing a source of useful data for analysis.

Until we have data that more directly measure what is important, the debate will continue about how to interpret the available metrics.[77 78] For example, many interventions in primary and community care aim to prevent hospital admissions.[2] There are some good reasons to reduce hospital admissions, including: their cost; the geographical variation in admission rates, suggesting that reductions are possible; and findings that a significant proportion of admissions are preventable through improvements in primary and community care.[79 80] What is less clear is the relationship between hospital admissions and patient preferences for how care is delivered.[81]

In parallel with efforts to improve the information infrastructure, we must match the growth in the use of routine data with growth in understanding about how these data were generated.[82] This is a considerable challenge, particularly as the nature of the data collected is changing over time. For example, the Quality Outcomes Framework linked payments for primary care to indices derived from the electronic medical record,[83] and in doing so both standardised codes and altered clinical practice in the UK.[84] The nature of the data recorded may also change as healthcare organisations attempt to become more person-centred, for example, as more patients are given access to their visit notes.[85 86]

Ideally, routine data collection would be extended to include variables that describe the data generating process. In some instances this may be feasible, for example by recording which diagnostic test was used alongside test results.[87] However, it is not always realistic, and instead 'big' data can more useful if used to create 'wide data'[88] where 'big data' is supplemented with qualitative and 'small' quantitative datasets, providing information on the social and technical processes underlying data collection and processing.[7] Examples within healthcare include audits of clinical practice to identify heterogeneity or changes in testing procedures, such as might occur in screening for healthcare associated infections.[87 89] Also valuable is qualitative work into patient and provider decision making,[76] or patient and clinical involvement in the design and interpretation of analysis.[90]

DIFFUSING DATA ANALYTICS

Even if the distance between a person and their data shadow is reduced or at least better understood, this will by no means create a learning health system. Routine data can be used to develop tools to improve the quality and safety of healthcare, but these need to be taken up and used to deliver the anticipated benefits.[91] In our opinion, this requires closer alignment between:

1. The questions addressed by the analysis and the decisions that will be taken;
2. Analysts and people interested in quality improvement; and
3. The range of analysis undertaken and public attitudes regarding the use of data.

Supporting decision-making

In the midst of the current hype around big data analytics, it can be easy to forget that the analytical products need to be developed so that they better meet the needs of their end users (eg, managers, clinicians, patients and policy makers) and, moreover, that it is important for analysts to demonstrate value. The timeliness of the analysis is important, but there are more subtle questions about whether the analytics is targeted at the decisions that clinicians and managers must make.[92] For example, although predictive risk models identify cohorts of patients at increased risk of adverse events,[39] by itself this does not provide information about which patients are most suitable for a given intervention.[93]

Another example is that comparative effectiveness studies often report the probability that the sample treatment effect could have occurred by chance (the p value), but this quantity might have limited applicability within the context of a particular quality improvement project.[94] When developing complex interventions,[95] the decision is not between offering the intervention or not, but many other options including spreading the intervention, modifying the eligibility criteria, improving fidelity to the original intervention design or refining the theory of change.[96 97] Recent developments in analytical methods aim to assess the relative costs and benefits of alternative decisions, and relate these to the mechanism of action of the underlying interventions for the target population.[98 99]

Embedded analytical teams[99a]

Statisticians and data scientists in academia have an essential role to play in developing new methods for health informatics, data linkage and modelling. Unfortunately, these groups have a limited ability to provide routine support and evaluation within healthcare systems, with some notable exceptions.[100 101] Data analysts working closely with clinicians at a local level have the advantage of local knowledge, which may make it easier to understand the care pathway and data generating process. Some healthcare providers appear to have successfully embedded data analytics, including the US Veterans Health Administration,[102] Intermountain[103] and University Hospitals Birmingham NHS (National Health Service) Foundation Trust.[104] However, in other health systems it is more challenging to assess the number, job descriptions and skills of data analysts,[105] and so plan for future workforce needs.

Ultimately, data analytics will not reach its full potential unless its importance is recognised by healthcare leaders. Unfortunately, in some instances there is uncertainty about how best to establish thriving analytical teams within healthcare organisations. Ideally, skilled analysts would be able to adapt to local needs and bring their expertise in the design, analysis and interpretation of data with an incentive to produce work that is clinically relevant and benefits patients and the public.[102]

Attitudes towards using data

Public trust in healthcare providers and research must not be undermined by real or perceived abuses of data.[3 106] Given data protection legislation that is generally formulated in terms of consent, system of collecting explicit patient consent for the secondary use of data would give the greatest certainty to both patients and the users of data, and has been signposted as an aim in England.[3] However, the primary focus when individuals interact with the health system is delivering healthcare, often with several points of access for patients, and so consent is not always sought for the secondary use of data that are generated as part of this interaction.

In the absence of consent, legal frameworks are necessary that strike the right balance between privacy and the public good,[3 107] and policy development in this area must be informed by an understanding of complex public attitudes about the use of data,[108 109] which requires an ongoing dialogue about the uses to which the data are put. Although most discussions about routine data are formulated in terms of identifying the subset of people who would prefer that data about them were not used in analysis, it is also worth bearing in mind that the 'data donation' movement shows that many people are prepared to contribute a wide range of data to improve healthcare for other people, as well as to further medical and scientific understanding.[110 111]

CONCLUSIONS

Ongoing feedback of insights from data to patients, clinicians, managers and policymakers can be a powerful motivator for change as well as provide an evidence base for action. Many studies and systems have demonstrated that routine data can be a powerful tool when used appropriately to improve the quality of care. A learning healthcare system may address the challenges faced by our health systems,[2] but for routinely collected data to be used optimally within such a system, simultaneous development is needed in several areas, including analytical methods, data linkage, information infrastructures and ways to understand how the data were generated.

Those familiar with the final part of Plato's allegory will know that one captive escaped his chains and gained full knowledge of his world. In that instance, the experience was not a happy one—on returning to his peers, they were not receptive to his offers to free

them from their chains.[1] While there is huge enthusiasm for routine data, diffusing data analytics throughout a healthcare system requires more and more sophisticated ways of analysing the shadows on the wall, and ongoing communication between different groups of people with very different perspectives. Only in this way will we achieve closer alignment between data analytics and the needs of decision makers and the public.

REFERENCES

1. Plato. The myth of the cave: from book VII. In: Lee D. ed. *The Republic*. Penguin Classics, 2007:312–60.
2. NHS England. Five year forward view. 2014. http://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf
3. National Information Board. Personalised Health and Care 2020, Using Data and Technology to Transform Outcomes for Patients and Citizens, A Framework for Action. 2014. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384650/NIB_Report.pdf
4. National Advisory Group on the, Safety of Patients in England. *A promise to learn—a commitment to act: Improving the Safety of Patients in England*. Leeds: Department of Health, 2013.
5. OECD. Strengthening Health information infrastructure for health care quality governance: good practices, new opportunities, and data privacy protection challenges. *OECD Heal Policy Stud* 2013;9:1–188.
6. Institute of Medicine. The Learning Healthcare System: Workshop Summary (IOM Roundtable on Evidence-Based Medicine). Washington, DC: The National Academies Press 2007:37–80.
7. Pope C, Halford S, Tinati R, *et al*. What's the big fuss about 'big data'? *J Health Serv Res Policy* 2014;19:67–8.
8. Black N. Assessing the quality of hospitals. *BMJ* 2010;340: c2066.
9. Shahian DM, Normand S-LT. What is a performance outlier? *BMJ Qual Saf* 2015;24:95–9.
10. Gitelman L. ed. *'Raw data' is an oxymoron*. Cambridge, MA: MIT Press, 2013.
11. Department of Health. Payment by results guidance for 2013–14. 2013.
12. Billings J, Blunt I, Steventon A, *et al*. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;2:1–10.
13. Amarasingham R, Patel PC, Toto K, *et al*. Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. *BMJ Qual Saf* 2013;22:998–1005.
14. Schoen C, Osborn R, Squires D, *et al*. A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. *Health Aff* 2012;31:2805–16.
15. Livingstone SJ, Levin D, Looker HC, *et al*. Estimated life expectancy in a Scottish cohort with type 1 diabetes, 2008–2010. *JAMA Intern Med* 2015;313:37–44.
16. Schmidt PE, Meredith P, Prytherch DR, *et al*. Impact of introducing an electronic physiological surveillance system on hospital mortality. *BMJ Qual Saf* 2014;24:10–20.
17. Noufaily A, Enki DG, Farrington P, *et al*. An improved algorithm for outbreak detection in multiple surveillance systems. *Stat Med* 2013;32:1206–22.
18. Ovretveit J, Keller C, Forsberg HH, *et al*. Continuous innovation: developing and using a clinical database with new technology for patient-centred care-the case of the swedish quality register for arthritis. *Int J Qual Heal Care* 2013;25:118–24.
19. Chaudhry SI, Mattera JA, Curtis JP, *et al*. Telemonitoring in patients with heart failure. *N Engl J Med* 2010;363:2301–9.
20. Henderson C, Knapp M, Fernandez J-L, *et al*. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. *BMJ* 2013;346:f1035.
21. Wicks P, Massagli M, Frost J, *et al*. Sharing health data for better outcomes on PatientsLikeMe. *J Med Internet Res* 2010;12:e19.
22. Drake G, Csipke E, Wykes T. Assessing your mood online: acceptability and use of Moodscope. *Psychol Med* 2013;43:1455–64.
23. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M, *et al*. Mobile apps in cardiology: review. *JMIR mHealth uHealth* 2013;1:e15. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4114428&tool=pmcentrez&rendertype=abstract (accessed 12 Feb 2015).
24. Greaves F, Laverty AA, Ramirez Cano D, *et al*. Tweets about hospital quality: a mixed methods study. *BMJ Qual Saf* 2014;23:838–46.
25. Bardach NS, Asteria-Peñaloza R, Boscardin WJ, *et al*. The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ Qual Saf* 2013;22:194–202.
26. Butler D. When Google got flu wrong. *Nature* 2013;494:155–6.
27. Brown M, Pinchin J, Blum J, *et al*. Exploring the relationship between location and behaviour in out of hours hospital care. *Commun Comput Inf Sci* 2014;435:395–400.
28. Srigley JA, Furness CD, Baker GR, *et al*. Quantification of the Hawthorne effect in hand hygiene compliance monitoring using an electronic monitoring system: a retrospective cohort study. *BMJ Qual Saf* 2014;23:974–80.
29. Currell R, Urquhart C, Wainwright P, *et al*. Telemedicine versus face to face patient care: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2000;(2): CD002098. http://dx.doi.org/10.1002/14651858.CD002098
30. Steventon A, Bardsley M, Billings J, *et al*. Effect of telecare on use of health and social care services: findings from the

Whole Systems Demonstrator cluster randomised trial. *Age Ageing* 2013;42:501–8.

31 Ryan M, Kinghorn P, Entwistle VA, *et al*. Valuing patients' experiences of healthcare processes: Towards broader applications of existing methodsTitle. *Soc Sci Med* 2014;106:194–203. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3988932/?report=reader http://dx.doi.org/10.1016/j.socscimed.2014.01.013

32 Craig R, Mindell J. *Health Survey for England 2013: health, social care and lifestyles*. Leeds: 2014. http://www.hscic.gov.uk/catalogue/PUB16076/HSE2013-Sum-bklet.pdf

33 Black N, Barker M, Payne M. Cross sectional survey of multicentre clinical databases in the United. *BMJ* 2004;328:1478.

34 Stuart EA, DuGoff E, Abrams M, *et al*. Estimating causal effects in observational studies using electronic health data: challenges and (some) solutions. *eGEMs* 2013;1:4.

35 Steventon A, Tunkel S, Blunt I, *et al*. Effect of telephone health coaching (Birmingham OwnHealth) on hospital use and associated costs: cohort study with matched controls. *BMJ* 2013;347:f4585.

36 Jordan RE, Lancashire RJ, Adab P. An evaluation of Birmingham Own Health telephone care management service among patients with poorly controlled diabetes. A retrospective comparison with the General Practice Research Database. *BMC Public Health* 2011;11:707.

37 Van Dalen MT, Suijker JJ, Macneil-Vroomen J, *et al*. Self-report of healthcare utilization among community-dwelling older persons: a prospective cohort study. *PLoS ONE* 2014;9:e93372.

38 Etheredge LM. A rapid-learning health system. *Health Aff* 2007;26:w107–18.

39 Billings J, Georghiou T, Blunt I, *et al*. Choosing a model to predict hospital admission: an observational study of new variants of predictive models for case finding. *BMJ Open* 2013;3:e003352.

40 Bardsley M, Billings J, Dixon J, *et al*. Predicting who will use intensive social care: case finding tools based on linked health and social care data. *Age Ageing* 2011;40:265–70.

41 Spencer R, Bell B, Avery AJ, *et al*. Identification of an updated set of prescribing-safety indicators for GPs. *Br J Gen Pract* 2014;64:181–90.

42 Collins A. Measuring what really matters: towards a coherent measurement system to support person-centred care. 2014. http://www.health.org.uk/public/cms/75/76/313/4772/Measuringwhatreallymatters.pdf?realName=7HWr4a.pdf

43 Austin PC, Mamdani MM, Stukel TA, *et al*. The use of the propensity score for estimating treatment effects: administrative versus clinical data. *Stat Med* 2005;24:1563–78.

44 Nelson E, Eftimovska E, Lind C, *et al*. Patient reported outcome measures in practice. *BMJ* 2015;350:g7818.

45 Pickard L, King D, Knapp M. The 'visibility' of unpaid care in England. *J Soc Work* 2015;41:404–9.

46 Hippisley-Cox J, Fenty J, Heaps M. *Trends in Consultation Rates in General Practice 1995 to 2006: Analysis of the QRESEARCH database*. Final Report to the Information Centre and Department of Health. 2007. http://www.qresearch.org/Public_Documents/Trendsinconsultationratesingeneralpractice1995to2006.pdf

47 Bardsley M, Billings J, Chassin LJ, *et al*. *Predicting social care costs: a feasibility study*. London: Nuffield Trust, 2011.

48 Graham M, Shelton T. Geography and the future of big data, big data and the future of geography. *Dialogues Hum Geogr* 2013;3:255–61.

49 Krieger N. Epidemiology and the web of causation: has anyone seen the spider? *Soc Sci Med* 1994;39:887–903.

50 Acheson ED. the Oxford record linkage study: a review of the method with some preliminary results. *Proc R Soc Med* 1964;57:269–74.

51 ONS. Labour Market Profile—Great Britain. Nomis Off. labour Mark. Stat. 2015. 2015. https://www.nomisweb.co.uk/reports/lmp/gor/2092957698/report.aspx (accessed 24 Mar 2015).

52 Redelmeier DA, Tibshirani RJ. Association between cellular-telephone calls and motor vehicle collisions. *N Engl J Med* 1997;336:453–8.

53 Marmot M, Wilkinson RG. eds. *Social determinants of health*. 2nd edn. Oxford: Oxford University Press, 2009.

54 Rubin DB. On the limitations of comparative effectiveness research. *Stat Med* 2010;29:1991–5.

55 Currie CJ, Peters JR, Tynan A, *et al*. Survival as a function of HbA(1c) in people with type 2 diabetes: a retrospective cohort study. *Lancet* 2010;375:481–9.

56 Bilimoria KY, Chung J, Ju MH, *et al*. Evaluation of surveillance bias and the validity of the venous thromboembolism quality measure. *JAMA* 2013;310:1482–9.

57 Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. *JAMA* 2011;305:2462–3.

58 Wennberg JE, Staiger DO, Sharp SM, *et al*. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. *BMJ* 2013;346:f549.

59 Steyerberg EW, Harrell FE, Borsboom GJJM, *et al*. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.

60 Stukel TA, Fisher ES, Wennberg DE, *et al*. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA* 2007;297:278–85.

61 Imbens GW, Lemieux T. Regression discontinuity designs: a guide to practice. *J Econom* 2008;142:615–35.

62 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.

63 Steventon A, Bardsley M, Billings J, *et al*. The role of matched controls in building an evidence base for hospital-avoidance schemes: a retrospective evaluation. *Health Serv Res* 2012;47:1679–98.

64 Stuart EA, Huskamp HA, Duckworth K, *et al*. Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Heal Serv Outcomes Res Methodol* 2014;14:166–82.

65 Wennberg DE, Sharp SM, Bevan G, *et al*. A population health approach to reducing observational intensity bias in health risk adjustment: cross sectional analysis of insurance claims. *BMJ* 2014;348:g2392.

66 Diamond A, Sekhon J. Genetic matching for estimating causal effects. *Rev Econ Stat* 2012;95:932–45. http://sekhon.polisci.berkeley.edu/papers/GenMatch.pdf

67 Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Stat Med* 2010;29:337–46.

68 Carpenter JR, Kenward MG. A comparison of multiple imputation and doubly robust estimation for analyses with missing data. *J R Stat Soc A* 2006;3:571–84.

69 O'Neill PD, Roberts G. Bayesian inference for partially observed stochastic epidemics. *J R Stat Soc A* 1999;162:121–9.

70 Cooper B, Lipsitch M. The analysis of hospital infection data using hidden Markov models. *Biostatistics* 2004;5:223–37.

71 He D, Ionides EL, King AA. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *J R Soc Interface* 2010;7:271–83.

72 Curtis LH, Brown J, Platt R. Four health data networks illustrate the potential for a shared national multipurpose big-data network. *Health Aff* 2014;33:1178–86.

73 Chitnis X, Steventon A, Glaser A, *et al*. Use of health and social care by people with cancer. London: Nuffield Trust, 2014.

74 Saunders MK. People and places: in Denmark, big data goes to work. *Health Aff* 2014;33:1245.

75 Carlsen K, Harling H, Pedersen J, *et al*. The transition between work, sickness absence and pension in a cohort of Danish colorectal cancer survivors. *Eur J Epidemiol* 2012;27: S29–30.

76 Sanders C, Rogers A, Bowen R, *et al*. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. *BMC Health Serv Res* 2012;12:220.

77 Bardsley M, Steventon A, Smith J, *et al*. Evaluating integrated and community based care. London: Nuffield Trust, 2013.

78 Patton MQ. *Utilization-focused evaluation*. 4th edn. Thousand Oaks: SAGE Publications, Inc, 2008.

79 Bardsley M, Blunt I, Davies S, *et al*. Is secondary preventive care improving? Observational study of 10-year trends in emergency admissions for conditions amenable to ambulatory care. *BMJ Open* 2013;3:e002007.

80 Wennberg JE. Forty years of unwarranted variation-And still counting. *Health Policy* 2014;114:1–2.

81 Guo J, Konetzka RT, Magett E, *et al*. Quantifying long-term care preferences. *Med Decis Mak* 2015;35:106–13.

82 Brown B, Williams R, Ainsworth J, *et al*. Missed Opportunities Mapping: Computable Healthcare Quality Improvement.

83 Quality and Outcomes Framework (QOF). http://qof.hscic. gov.uk/ (accessed 18 Feb 2015).

84 Gillam SJ, Siriwardena AN, Steel N. Pay-for-performance in the United Kingdom: impact of the quality and outcomes framework — a systematic review. *Ann Fam Med* 2012;10:461–8.

85 Morris L, Milne B. Enabling Patients to Access Electronic Health Records. 2010. http://www.rcgp.org.uk/Clinical-and-research/Practice-management-resources/~/media/Files/Informatics/Health_Informatics_Enabling_Patient_Access.ashx

86 Cresswell KM, Sheikh A. Health information technology in hospitals: current issues. *Futur Hosp* 2015;2:50–6.

87 Goldenberg SD, French GL. Diagnostic testing for Clostridium difficile: a comprehensive survey of laboratories in England. *J Hosp Infect* 2011;79:4–7.

88 Tinati R, Halford S, Carr L, *et al*. Big data: methodological challenges and approaches for sociological analysis. *Sociology* 2014;48:663–81.

89 Fuller C, Robotham J, Savage J, *et al*. The national one week prevalence audit of universal meticillin-resistant Staphylococcus aureus (MRSA) admission screening 2012. *PLoS ONE* 2013;8:e74219.

90 Boland MR, Hripcsak G, Shen Y, *et al*. Defining a comprehensive verotype using electronic health records for personalized medicine. *J Am Med Inform Assoc* 2013;20: e232–8.

91 Rogers EM. *Diffusion of innovations*. 5th edn. New York: Free Press, 2003. doi:citeulike-article-id:126680

92 Roski J, Bo-Linn GW, Andrews TA. Creating value in health care through big data: Opportunities and policy implications. *Health Aff* 2014;33:1115–22.

93 Lewis GH. 'Impactibility models': identifying the subgroup of high-risk patients most amenable to hospital-avoidance programs. *Milbank Q* 2010;88:240–55.

94 Forster M, Pertile P. Optimal decision rules for HTA under uncertainty: a wider, dynamic perspective. *Health Econ* 2013;22:1507–14.

95 Craig P, Dieppe P, Macintyre S, *et al*. *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council, 2008.

96 Parry GJ, Carson-Stevens A, Luff DF, *et al*. Recommendations for evaluation of health care improvement initiatives. *Acad Pediatr* 2013;13:S23–30.

97 De Silva MJ, Breuer E, Lee L, *et al*. Theory of Change: a theory-driven approach to enhance the Medical Research Councils' framework for complex interventions. *Trials* 2014;15:267.

98 Frangakis CE. The calibration of treatment effects from clinical trials to target populations. *Clin Trials* 2009;6:136–40.

99 De Stavola BL, Daniel RM, Ploubidis GB, *et al*. Mediation analysis with intermediate confounding: structural equation modeling viewed through the causal inference lens. *Am J Epidemiol* 2015;181:64–80.

99a Hartman E, Grieve R, Ramsahai, R, *et al*. From sample average treatment effect to population average treatment effect on the treated: combining experimental with observational studies to estimate population treatment effects. *J R Stat Soc*. Series A (Statistics in Society). 2015. [epub ahead of print]. doi:10.1111/rssa.12094

100 Academic Health Science Networks. http://www.england.nhs.uk/ourwork/part-rel/ahsn/ (accessed 23 Feb 2015).

101 Department of Health. Innovation health and wealth: accelerating adoption and diffusion in the NHS. *Proc R Soc Med* 2011;55:1–35.

102 Fihn SD, Francis J, Clancy C, *et al*. Insights from advanced analytics at the veterans health administration. *Health Aff* 2014;33:1203–11.

103 James BC, Savitz LA. How intermountain trimmed health care costs through robust quality improvement efforts. *Health Aff* 2011;30:1185–91.

104 Healthcare Evaluation Data. https://www.hed.nhs.uk/ (accessed 23 Feb 2015).

105 Allcock C, Dormon F, Taunt R, *et al*. *Constructive comfort: accelerating change in the NHS*. The Health Foundation, 2015. http://www.health.org.uk/public/cms/75/76/313/5504/Constructivecomfort-acceleratingchangeintheNHS.pdf?realName=Dz5g23.pdf

106 Carter P, Laurie GT, Dixon-woods M. The social licence for research: why care.data ran into trouble. *J Med Ethics* 2015.

107 Clark S, Weale A. *Information Governance in Health: an analysis of the social values involved in data linkage studies*. London: Nuffield Trust, 2011:2–39.

108 Cameron D, Pope S, Clemences M. Exploring the public's views on using administrative data for research purposes. 2014. http://www.esrc.ac.uk/_images/Dialogue_on_Data_report_tcm8-30270.pdf

109 Department of Business Skills & Innovation. *Seizing the data opportunity: a strategy for UK data capability*. London, 2013.

110 Wikilife Foundation. Data Donors. http://datadonors.org/ (accessed 25 Mar 2015).

111 Open SNP. https://opensnp.org/ (accessed 23 Mar 2015).