Response to the Department of Health's consultation on data security, consent and opt-outs

7 September 2016

1. About the Health Foundation

1.1. The Health Foundation is an independent charity committed to bringing about better health and health care for people in the UK.

1.2. Our aim is a healthier population, supported by high quality health care that can be equitably accessed. From giving grants to those working at the front line to carrying out research and policy analysis, we shine a light on how to make successful change happen. We use what we know works on the ground to inform effective policymaking and vice versa.

We believe good health and health care are key to a flourishing society. Through sharing what we learn, collaborating with others and building people's skills and knowledge, we aim to make a difference and contribute to a healthier population.

2. Key points

2.1. **Benefits (see Section 4)**
The National Data Guardian's review of data security, consent and opt-outs (the review) argued that the public is yet to be persuaded of the case for using data and it makes recommendations for how the benefits can be assessed and communicated. In taking these recommendations forward:

- A broad approach should be adopted to assessing the benefit of processing data, recognising that the purpose of processing data is to produce information. Whether or not this information will benefit patients depends on how it is used nationally or in a local context.
- Data sharing can benefit population health in many ways, not just through improved health care. In particular, when assessing the potential benefit of new research, the potential to improve wellbeing through public services other than the NHS should be reflected.
- Several organisations are well positioned to communicate to the public about the benefits of using data to improve health and care, and this responsibility should not lie with NHS Digital or the Health Research Authority alone.

2.2. **Data security standards (see Section 5)**

The new data security standards are welcome, but it is important not to underestimate the challenge of embedding these within routine practice across health and social care. Greater clarity is also needed in places. Specifically:

- Initiatives to improve patient safety in the NHS offer lesson which could inform the approach to implementing the data security standards. Data security and patient safety, although different in nature, have many parallels.
- A set of principles is needed to guide implementation, and these must include the need to create a culture that promotes a safe approach to handling data, and simplicity for those making decisions around data sharing.
- It is unclear to which organisations and categories of data the new data security standards will apply.
- The proposed data security standards focus on people, processes and technology, but important aspects of data security are not reflected. This includes controls around what information can be released following data processing, and for which purposes data can be accessed.

2.4 **New opt-out arrangements (see Section 6)**

The proposed framework for opt-outs is welcome, but it will be important to clarify the language used and to ensure that the full range of quality improvement activity is enabled. Specifically:

- The phrase ‘anonymised data’ may be confusing to the public and alternatives should be considered.
- The review recommends that, in future, patient opt-outs will not apply to anonymised data. Such an approach would mean that it is much easier to use patient data to improve the quality and efficiency of the health care provided.
- Quality improvement is part of the day-to-day running of the NHS, but its data requirements have not been reflected in the review. Clarification is needed regarding the treatment of data used for quality improvement, so that patients do not inadvertently opt-out of data sharing for this purpose.

2.5 **Implementation (see Section 7)**

Data linkage and de-identification of patient data are key techniques that are increasingly important in safe use of patient information. For the proposed opt-out / consent model to be successful careful implementation is required. Specifically:

- Local teams should be supported to develop and maintain standards, skills, software, systems and resources to de-identify patient information.
- In the meantime, it will be necessary that health and social care data can flow in identifiable form to NHS Digital to provide linkage and de-identification. From a practical perspective, it may be necessary in the short term to exempt data collections transferred to NHS Digital from the opt-out, including those not covered by the Secretary of State’s instructions.

NHS data sets are increasingly diverse and will become more so over time, as new technologies are developed. The capacity of national organisations to receive, clean and link increasing amounts of data should be carefully assessed. It is also worth exploring alternatives, such as a federated network of trusted third parties that could fulfil this role. In transitioning to a new consent / opt-out model from the current system, a number of difficult decisions will be made. Therefore:

- Urgent research is needed regarding the characteristics of patients who have already objected to data sharing.
- In reaching a decision about the existing opt-outs, it is necessary to consider ethical requirements, technical ability, benefits and health inequalities. The Department of Health should consider 'passporting' people from the old system to the new system.

3. Introduction

3.1. We all use information as part of our day-to-day lives, and this is equally true for those people working to deliver and improve health and social care. Medicine is increasingly an information science. In making decisions about treatments, health care practitioners and patients seek to combine existing evidence with information about the particularities of the health problems presented. More generally, without information, people working in the NHS would not be able to assess the quality of the care provided, develop and test ideas for improving care, provide assurance to the public and regulators, or ensure efficiency.

3.2. One of the general features of NHS data (though not all) is that they were originally collected for the purposes of providing direct care. For example, when we visit our general practitioner, information is routinely recorded on our appointments, our health problems and symptoms, and the treatments and outcomes that follow. In addition to these clinically-generated data, the NHS produces large amounts of administrative data, often for the primary purpose to reimburse the providers of care. Finally, increasing amounts of patient-generated data are being produced, for example from apps and telemonitoring devices. We set out a taxonomy regarding NHS data and its features in our publication: '*Making sense of the shadows: priorities for creating a learning healthcare system based on routinely collected data*[1].'

3.3. While these data were originally collected for the purposes of providing direct patient care, they may have many secondary benefits. Indeed, by combining the records from multiple patients, it is possible to obtain critical information about the quality of the care provided. The NHS has been conducting secondary analysis of health care data for decades, as have researchers. But new data sets have become available, and increasingly it is possible to link multiple data sets to produce richer information. More analysis is being done, but much of this activity remains opaque to the public, who have limited understanding about how data about them are being used, and sometimes no way to find out. Part of the context of the review is that, as a society we have become more aware of the potential dangers of data sharing, and to many people health data is particularly sensitive.

[1] Deeny SR, Steventon A. Making sense of the shadows: priorities for creating a learning healthcare system based on routinely collected data. 2015. BMJ Quality and Safety 2015. Online first. – Available at http://www.health.org.uk/our-people/adam-steventon#sthash.PgVYFSfG.dpuf

3.4. We are responding to this consultation because our aim is a healthier population, supported by high quality health care that can be equitably accessed. From our work, we have insights that may be helpful in further developing the recommendations that have emerged from this review.

3.5. From our grant giving, we are familiar with many of the problems that front line teams face when seeking to use data to improve care. Moreover, we have established an in-house data analytics team, which uses de-identified patient information within a secure data processing environment. As a result, we are familiar with the potential of data analytics to inform decision-making in the NHS, and also the need to implement rigorous approaches to data security, and continuously improve them. Although, we process only de-identified data, we already exceed the ten new data security standards.

3.6. The National Data Guardian's report acknowledges the public benefit of using data to improve health care, while trying to reduce the risk that harm results from the inappropriate disclosure of data. We agree that data security needs to be promoted and in this consultation response we suggest some ways in which the NHS could go further than the review has proposed. At the same time, it is necessary to promote the use of data analytics, since it is one of the seven ingredients for successful change. These two aims are not in conflict, since data security can be an important enabler of data analytics, provided that the approach is transparent, rigorous, and practical. We welcome the various initiatives highlighted by the review to simplify the approach to using data in the NHS, including the proposal to develop an understandable set of opt-outs that are organised around the purpose of the data processing, rather than organisational boundaries.

3.7. The structure of this consultation response is as follows. In Section 4, we comment on the benefits of using data, and how the case for data sharing can be made to the public. Then in Section 5, we comment on the proposed new data security standards. Section 6 is devoted to the new patient opt outs, while the remaining section addresses the implementation of the review.

4. Making the case for using data

4.1. The review argues that the case for using data has not yet been made to the public, and it makes recommendations for how benefit can be assessed and communicated. In taking these recommendations forward:

- A broad approach should be adopted to assessing the benefit of processing data, recognising that the purpose of processing data is to produce information. Whether or not this information will benefit patients depends on how it is used nationally or in a local context.
- Data sharing can benefit population health in many ways, not just through improved health care. In particular, when assessing the potential benefit of new research, the potential to improve wellbeing through public services other than the NHS should be reflected.
- Several organisations are well positioned to communicate to the public about the benefits of using data to improve health and care, and this responsibility should not lie with NHS Digital alone.

4.2. ***A broad approach should be adopted to assessing the benefit of processing data, recognising that the purpose of processing data is to produce information. Whether or not this information will benefit patients depends on how it is used nationally or in a local context****.*

4.3. The review rightly emphasises that communication with the public cannot be viewed as a single event, and recommends that the National Information Board conducts ongoing work to build greater public trust in data sharing for health and social care (recommendation 20). The review also contains two more specific recommendations, both aimed at making it easier for the public to access information about how data are used. The first of these (recommendation 17) is that the Health Research Authority should provide the public with an easily digestible explanation of the projects that use personal confidential data and have been approved following the advice from the Confidentiality Advisory Group. The second is that the Health and Social Care Information Centre (now NHS Digital) should develop a tool to help people understand how sharing their data has benefited other people (recommendation 18).

4.4. The focus on benefit has arisen, in part, because of protections introduced under the Care Act 2014 which mean that NHS Digital can only disseminate information for the provision of health care and adult social care, or the promotion of health. Following this legislation, a committee was established (called the Data Access Advisory Group, or DAAG) to assess whether the benefits relating to particular pieces of data analytics justify such processing. As a result, *benefit* is an important concept that underlies the processes for deciding when data sharing is appropriate. Following the National Data Guardian's review, *benefit* has emerged as an important concept underlying communication with the public. Indeed, the terms of reference for the review included 'being clear with citizens and professionals how personal health and care data should be used, and the benefits of doing so.'

4.5. Despite the focus on *benefit,* the term has not been precisely defined. It seems to have been introduced into this sphere quite recently. For example, the word *benefit* was not

used in the previous Caldicott Review, *'Information: to share or not to share[2].'* Opinions seem to differ regarding what is 'sufficient benefit' and the 'nature of benefit.' Often, a relatively narrow approach has been adopted, which relies on projecting forward the likely outputs from a piece of analysis (eg written reports), and then seeking to estimate the impact of those outputs on improved patient care, almost as one would a new drug or treatment.

4.6. In section 5 of this response, we set out a conceptual framework for information governance, relating to the Five Safes. One of the main elements is an assessment of the purpose for the data processing. For example, the data sharing must have a legal basis and be consistent with the reasonable expectations of individuals. It seems reasonable, that in this stage, there is some consideration of the likely benefit of data processing. However, data analytics is not like a new drug or treatment, and our research has shown that these benefits must not be considered too narrowly.

4.7. Our report *Constructive Comfort: accelerating change in the NHS*[3] identified seven success factors for positive change at any level of the health system, drawing upon relevant literature, testimony from professionals working on the front line and leaders in national bodies, and our own experience of funding improvement programmes. These factors are:

- Committed and respected leadership that engages staff
- A culture hospitable to, and supportive of, change
- Management practices that ensure execution and implementation
- Capabilities and skills to identify and solve problems
- Data and analytics that measure and communicate impact
- Resources and support for change
- An enabling environment which supports and drives change.

4.8. Data analytics is a fundamental part of many of these activities, not just measuring and communicating impact, but also identifying and solving problems, and ensuring execution and implementation. Nevertheless, realising the potential benefit of data analytics depends on the entirety of the change system, including for example an enabling environment to support and drive change.

4.9. This is apparent from our day-to-day lives. The output of data analytics is a type of information, but information doesn't have an impact through the mere fact of its existence – it needs to be communicated and translated into practice. Moreover, information can have effects that are not predictable from the outset, but occur a long time down the line.

[2] National Data Guardian. Information: To share or not to share? The Information Governance Review. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf (Accessed on 7 September 2016)

[3] Allcock, A. Dormon, F. Taunt, R. and Dixon, J. (2015) Constructive Comfort: accelerating change in the NHS. Available at: http://www.health.org.uk/publication/constructive-comfort-accelerating-change-nhs

4.10. When taking forward the recommendations from the review, further work should be considered to provide a conceptual framework for what is meant by 'benefit'. This should reflect the reality that the realisation of benefits depends on local context, which is developing rapidly, and is often unpredictable.

4.11. ***Data sharing can benefit population health in many ways, not just through improved health care. In particular, when assessing the potential benefit of new research, the potential to improve wellbeing through other public services than the NHS should be reflected.***

4.12. It is now widely accepted that health care systems, however good, are not the largest contributors to our overall health. Estimates range from 10% to 40% of our health being a result of our access to health care with the rest being determined by wider social, economic and environmental factors.

4.13. Government and the public sector have many levers to improve health, including housing, efforts to improve our local environment, urban regeneration, the approach to disability and unemployment benefits, state pensions, and economic policy. Yet, it is often very difficult to assess the effects of these on health outcomes, since some health outcomes are recorded in NHS data sets. The Care Act 2014 identified that NHS Digital can disseminate information for the promotion of health, but it has been exceptionally difficult to link data sets from across government. Looking at NHS Digital's latest data release register (covering January – March 2016) a total of 808 datasets have been released, only 16 of these releases include data linkage. Four of these involve linkage of data to data already held by the data recipient, the other 12 involve routinely linked datasets. A further 44 datasets have been released in identifiable form to support linkage within an existing NHS system. None of the data releases relate to novel data linkage with another government department[4].

4.14. The review is recommending a sustained programme of communication with the public about data sharing. The rationale for linking NHS data to other data sets, for the purposes of research to improve population health, should be reflected in this programme. This will contribute towards establishing a firmer basis for linking NHS data sets to other administrative data to improve health through better policymaking.

4.15. ***Several organisations are well positioned to communicate to the public about the benefits of using data to improve health and care, and this responsibility should not lie with NHS Digital alone.***

4.16. NHS Digital clearly has an important role in enabling data analytics, protecting privacy, and promoting public understanding about data sharing. Yet it is important to recognise that the majority of health and social care data do not flow to NHS Digital. This includes general practice data, the operational databases used by councils with social care responsibilities, the electronic medical records of hospitals, diagnostic databases, telehealth systems, and many others. Regardless of the approach taken in future, NHS Digital will never hold the totality of NHS data, because new data will always be collected

[4] We consider linkage to the Office for National Statistics' births and death register routine linkage.

as part of small-scale, local innovative projects. Also, many government departments hold data that are relevant to efforts to improve population health. Therefore, the communication programme that is envisaged will require the active support of many organisations, and it would be helpful to clarify roles and responsibilities in this area.

Table 1: Examples of how data can be used to improve health care and population health

Aim	Examples
To deliver direct patient care	Clinical decision support tools, aiming to identify risk of events (eg heart attacks) or prompt evidence-based care (eg dangerous combinations of medicines) Coordinating care across settings (eg multidisciplinary teams)
To develop new treatments for health conditions	Research to produce information about the effectiveness of new treatments and procedures in improving health outcomes Surveillance regarding the impact of new treatments and procedures when introduced into routine clinical practice
To improve the quality of local services	Clinical audits to assess the quality of health care against objective clinical standards and promote change Local quality improvement work, which relies on data to track changes in key metrics over time Monitoring and surveillance of (avoidable) harms resulting from treatment.
To commission health care for defined populations	Making treatment more effective and efficient by targeting specific population segments Identifying the extent of and addressing health inequalities Evaluating the impact of changes to how health care is delivered (eg more integrated care) Providing assurance to local populations about the quality of care provided
To support the efficient running of services by providers of health care	Assessing the impact of new payment or service delivery models Surveillance to identify risks of poor quality care Identifying priorities for new investments/savings Planning the scale and location of new services or closure of services
To enable the efficient operation of the health care system at national level	Reimbursing providers for the care provided Allocating resources to commissioners in line with national policy
To inform broader efforts	Research to assess the impact on health of other initiatives by government

to improve population health	or the public sector (eg the sugar tax, or changes to disability benefits)

5. Data security standards

5.1. When handling patient data, it is essential to limit the potential harms that might result from these data being passed into the wrong hands, or being used for purposes that are not appropriate. Thus, controls need to be in place to ensure that data are used safely. The review proposes ten data security standards that will be applied across all health and social care organisations.

5.2. The new data security standards are welcome, but it is important not to underestimate the challenge of embedding these within routine practice across health and social care. Greater clarity is also needed in places. Specifically:

- There have been previous initiatives to improve patient safety in the NHS, and the lessons from these could inform the approach to implementing the data security standards. Data security and patient safety, although different in nature, have many parallels.
- A set of principles are needed to guide implementation, and these must include the need to create a culture that promotes a safe approach to handling data, and simplicity for those making decisions around data sharing.
- It is unclear to which organisations and categories of data the new data security standards will apply.
- The proposed data security standards focus on people, processes and technology, but important aspects of data security are not reflected. This includes controls around what information can be released following the data processing, and for which purposes data can be accessed.

5.3. ***Initiatives to improve patient safety in the NHS offer lesson which could inform the approach to implementing the data security standards. Data security and patient safety, although different in nature, have many parallels.***

5.4. The primary objective for introducing data security standards is to prevent harm to patients. This same objective is what drives initiatives on patient safety. In both instances, incidents are not acceptable and the aim is to limit their occurrence to zero. Both disciplines aim to create a blame-free culture where near misses are routinely reported, learning from incidents is used to improve processes, and staff receive support, guidance and training to ensure they can follow procedures designed to improve process reliability.

5.5. When it comes to improving the safety of systems in the NHS, there is a typical pattern: standards are set and the principles of a blame-free culture are espoused, but at the same time a comprehensive system of sanction, regulation and inspection is established to catch and punish people who fall foul of standards. Accountability is right and proper in cases of wilful negligence, but all too often latent system factors explain the events that result in harm to patients. Unfortunately, sanctions are often associated with fear, and this fear can undermine the behaviours necessary to meet the standards and principles. In taking forward the recommendations of this review, the Department of Health should bear in mind the advice set out in Don Berwick's review of the safety of patient in England in 2013, *A promise to learn – a commitment to act*, commissioned by the Secretary of State for Health:

- "Fear is toxic to both safety and improvement: Fear impedes improvement in complex human systems. Time and again, we see the harvest of fear in the Mid Staffordshire story, a vicious cycle of over-riding goals, misallocation of resources, distracted attention, consequent failures and hazards, reproach for goals not met, more misallocation and growing opacity as dark rooms with no data came to look safer than ones with light. "Better not to know" became the order of the day."[5]

5.6. Evidence provided by the NHS staff survey illustrates the increasing sense of blame being felt by people working in the NHS. While on the one hand NHS staff feel more confident that their organisation takes action to ensure the safety incidents do not reoccur (rising from 55% in 2010 to 63% in 2014), the number of staff who feel their organisation blames or punishes people involved in errors has risen during the same period (from 10% to 13%).[6] *While these figures relate to patient safety broadly, and to our knowledge there has been no equivalent research to assess views in relation to data security, the presence of a blame-culture raises serious concerns about how the new data security standards are implemented. If enforced too rigidly such standards could undermine the safe, secure and effective use of data rather than support it.*

5.7. ***A set of principles are needed to guide implementation, and these must include the need to create a culture that promotes a safe approach to handling data, and simplicity for those making decisions around data sharing.***

5.8. The review points to a number of problems with the existing NHS Information Governance Toolkit (the IG Toolkit). We agree that the IG Toolkit has provided a valuable framework, but it is currently very complex, and that there are concerns with the self-assessed nature. We welcome the suggestion that the Toolkit should be redesigned to embed the ten new standards. There is also an opportunity to make the Toolkit into a portal to access information governance training and other resources for a wider range of organisations.

5.9. The review mentions that the IG Toolkit has often been seen as a 'tick-box' exercise. Given this conclusion, the review contains some welcome reflections on how to embed the new standards into practice in a meaningful way. Its recommendations include:

- strengthening leadership (through the Senior Information Risk Owner Role and by quicker reporting to senior management about near misses);
- changing processes so that it easier to do the right thing;
- providing a simple explanation to health and social care professionals about what they can and cannot do with data;
- amending CQC's inspection framework for providers of registered health and care services to include assurance that appropriate internal and external validation against the new data security standards have been carried out;
- the use of information from the redesigned Information Governance Toolkit to target action by CQC;
- tougher sanctions for malicious or intentional data breaches.

[5] 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. London: Crown; 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/226703/Berwick_Report.pdf (accessed 7 September March2016).

[6] Illingworth, J (2015) Is the NHS getting safer? Available at: http://www.health.org.uk/publication/nhs-getting-safer

5.10. The review acknowledges that more detailed design work is needed, both to develop its recommendations and consider how it is implemented. The *how* of implementation can be as important as the *what.* Thus, it would be very valuable to articulate a set of principles regarding implementation.

5.11. One of the most compelling findings to emerge from the review is the importance of creating a culture of continuous improvement, where near misses, hazards or risky behaviours can be reported without fear of recrimination, and are seen as an opportunity to improve care rather than a reason to assign blame. In our own work on data security at the Health Foundation, we have found that a safe and supportive culture, focussed on learning, is essential to encourage individuals to come forward to report 'near misses' and to do the deep analysis required to understand the root causes.

5.12. Another guiding principle might be simplification. The review presented convincing evidence that, in health and social care organisations, the 'rules' about what can and cannot be done with data are very unclear. This does not create an environment conducive to data security. For example, it is hard to imagine that people can be effectively held to account by their peers for behaviours relating to data security, when opinions differ regarding which behaviours are appropriate. The danger is that, if the system is opaque, then it may be hard to challenge colleagues or even to suggest better ways of working.

5.13. We give one example of how these principles may be helpful when developing the approach to implementing the recommendations of the review. One of the recommendations of the review is that performance information from the redesigned Information Governance Toolkit is collected by NHS Digital, and that this information is shared with CQC to target inspection activity. While there is a rationale for doing so, it will be important to establish an environment that is conducive to open and honest assessment of data security standards.

5.14. We welcome the strong focus of the review on an engaged board that demonstrates clear ownership and responsibility for data security. However, we also think that, in parallel with these efforts, more support might be given to senior and middle management. In our own work, we have found that efforts to improve data security require close collaboration between a wide range of functions, including the end-users of data and people working in facilities, information technology, communications and human resources. As such it has been important to establish an engaged and diverse management group. At the Health Foundation, such a group is responsible for running our secure data environment, and for implementing and continuously improving our processes.

5.15. ***It is unclear to which organisations and categories of data the new data security standards will apply.***

5.16. Page 2 of the review states that the new standards will apply to all health and social care organisations, while page 4 recommends that they apply to every organisation handling health and social care data. These two sets of organisations are not the same (for example, the Health Foundation would fall under the second category but not the first). It will be important to develop a precise scope.

5.17. The review refers throughout to 'personal confidential data' but this term is not defined in the review and it is not in common use. Although we cannot be certain, one possible reading of the review is that 'personal confidential data' refers to data that contain patient

identifiers (*ie*, fields such as name, full date of birth or NHS Number). If this is the case, then the proposed data security standards would not apply to de-identified data. However, there is still a risk of re-identification even with de-identified data, which has had all patient identifiers removed. This risk is not negligible, so it seems appropriate that these data are also protected by a data security standard (although these data may not necessarily require the same set of controls as fully identifiable data).

5.18. In going forward, it will be important to be clear regarding the scope of the new standards. Without more clarity, the risks involved in processing de-identified data might not be made visible to leaders or the public, or adequately managed.

5.19. ***The proposed data security standards focus on people, processes and technology, but important aspects of data security are not reflected. This includes controls around what information can be released following the data processing, and for which purposes data can be accessed.***

5.20. The proposed data security standards focus on people, processes and technology. Although this is an intuitive approach, we are concerned that some important aspects of data security are not covered. These include:

- Controls put in place to protect patients should be proportionate to the sensitivity of the data (eg to the nature of the data concerned and the risk of re-identification). The current set of standards do not take in to account the various sensitivity levels that patient data may take, and how these might affect the approach required.
- The proposed standards do not account for the risk that outputs from the data processing might pose to patients. Where appropriate, a process called statistical disclosure control could be implemented to safeguard the non-disclosive nature of analysis results (eg making sure that small numbers are adequately supressed).
- Although the proposed standards would control who has access to patient data (eg people who have completed information governance training), an assessment of the purpose of data use is not currently proposed. For instance, an academic evaluation of the effectiveness of certain medication could lead to changes in prescriptions and benefit future patients. The same benefit might be realised if the research is carried out by a pharmaceutical company, however, the purpose of the work might be different (eg to identify gaps in the 'market'). Depending on the purpose of the work, additional controls may or may not be warranted.

5.21. In further developing the data security standards, it may be helpful to use the Five Safes Framework (the framework).[7] This framework is already used by a number of organisations including the Office for National Statistics (for their Virtual Microdata Laboratory), Her Majesty's Revenue and Customs' Data Lab, the UK Data Archive and the Administrative Data Research Network. Although these organisations primarily support secondary uses of data, the framework applies to all uses of data. However, the Framework was not referred to in the evidence review presented in the supporting document published by CQC.

[7] Desai, T., Ritchie, F. and Welpton, R. (2016) ***Five Safes***: ***Designing data access for research***. Working Paper. University of the West of England, available at http://eprints.uwe.ac.uk/28124/

5.22. The Five Safes are:

- Safe Data: What are the characteristics of the data to be processed?
- Safe People: Who will be allowed to access the data?
- Safe Projects: What will the data be used for?
- Safe Settings: How should a secure environment be designed?
- Safe Outputs: What information can be released from the environment?

5.23. The framework is a useful guide in implementing secure data processing environments, and it mitigates the risk that important aspects of data access are overlooked. By making sure that controls are appropriate for the nature of the data, implementation of overly restrictive controls can be mitigated. As the review mentions, controls that are too strict or impractical can add to the risk to patients rather than reduce it, as users of the data may try to circumvent the controls, thereby exposing the data.

5.24. The framework can also be used in ongoing governance of a processing environment. At the Health Foundation this framework is used by the management group looking after our secure data environment. In mitigating risks, controls are considered for each of the five aspects of data security. Consider a hypothetical case where, a novel data linkage is proposed that would increase the re-identification risk to patients above our 'business as usual' (safe data). In order to continue to process data safely, we would want to make sure the purpose of the analysis is in line with our charitable goals (safe projects) and we would consider additional training for analysts involved in the project (safe people). We also would need to make sure our technical and physical infrastructure is adequately secure (safe setting) and make sure that our analysis results are scrutinised by applying statistical disclosure control by two rather than one colleague that is not related to the project (safe outputs). From a governance perspective, we would also want to make sure that all data controllers involved are aware of the added re-identification risk that follows from the data linkage (this could for instance be reflected in a privacy impact assessment)."

5.25. In Table 2, we map the ten proposed data security standards to the Five Safes framework. As can be seen, the standards do not refer to the characteristics of the data to be processed, or acknowledge that some data processing is riskier than others. Moreover, none of the proposed standards address the results and findings (or outputs) of the data processing.

Table 2: The ten proposed data security standards mapped against the Five Safes

Five Safes	Proposed Data Security Standards
Safe Data	*This is missing from the proposed standards.*
Safe People	(1) All staff ensure that personal confidential data is handled, stored and transmitted securely, whether in electronic or paper for. Personal confidential data is only shared for lawful and appropriate purposes. (2) All staff understand their responsibilities under the National Data Guardian's Data Security standards including their obligation to handle information responsibly and their personal accountability for deliberate or avoidable breaches. (3) All staff complete appropriate annual data security training and pass a mandatory test, provided through the revised Information Governance Toolkit.
Safe Projects	Only the legal aspect of this is covered in (1).
Safe Settings	(4) Personal confidential data is only accessible to staff who need it for their current role and access is removed as soon as it is no longer required. All access to personal confidential data on IT systems can be attributed to individuals. (6) Cyber-attacks against services are identified and resisted and CareCERT security advise is responded to. Action is taken immediately following a data breach or near miss, with a report made to senior management within 12 hours of detection. (8) No unsupported operating systems, software or internet browsers are used within the IT estate. (9) A strategy is in place for protecting IT systems form cyber threats which is based on a proven cyber security framework such as Cyber Essentials. This is reviewed at least annually.
Safe Outputs	*This is missing from the proposed standards.*
Governance	(5) Processes are reviewed at least annually to identify and improve processes which have caused breaches or near misses, or which force staff to use workarounds which compromise data security. (7) A continuity plan is in place to respond to threats to data security, including significant data breaches or near misses, and it is tested once a year as a minimum, with a report to senior management. (10) IT suppliers are held accountable via contracts for protecting the personal confidential data they process and meeting the National Data Guardian's Data Security Standard.

6. New opt out arrangements

6.1. We echo the findings of the review that the existing system of opt-outs is confusing. Currently, there are two opt outs, one applying to personal confidential data leaving the general practice for purposes beyond direct care (type 1), and the other to personal confidential data being disseminated from NHS Digital aimed at purposes beyond their direct care (type 2). This system presupposes a public understanding about how data flow across the NHS, which is unrealistic. We welcome the development of a new approach that focusses on the purposes for which data are shared.

6.2. The proposed framework for opt-outs is welcome, but it will be important to clarify the language used and to ensure that the full range of quality improvement activity is enabled. Specifically:

- The phrase 'anonymised data' may be confusing to the public and alternatives should be considered.
- The review recommends that, in future, patient opt-outs will not apply to anonymised data. Such an approach would mean that it is much easier to use patient data to improve the quality and efficiency of the health care provided. Quality improvement is part of the day-to-day running of the NHS, but its data requirements have not been reflected in the review. Clarification is needed regarding its treatment, so that patients do not inadvertently opt out of data sharing for this purpose.

6.3. ***The phrase 'anonymised data' may be confusing to the public and alternatives should be considered.***

6.4. Although these data do not contain strongly identifiable fields (such as name, date of birth, or address), by necessity they do still include some patient information, such as age, gender, ethnicity or area of residence. As a result, there is still a possibility that these data could be combined with other information to identify individual patients ('re-identification risk').

6.5. In its approach to the opt-outs, the review uses the phrase 'anonymised data'. This term is described in the Information Commissioner's (ICO) *Anonymisation: managing data protection risk code of practic*e and in the ICO's code, there is an assumption that controls are in place to mitigate the residual risk of de-identification. This is important because, although data sets can be structured to minimise the risk of re-identification (for example by using age rather than full date of birth), often it is not possible to entirely remove the risk of re-identification purely by changing the nature of the data set. As a result, in many sectors it is usual practice to implement a set of controls to reduce the risk of re-identification. These include many of those recommended by the review (although as we point out above, the review only recommends these controls will apply to personal confidential data).

6.6. The review is recommending that the opt-outs will not apply to anonymised data, but at least in terms of the ICO's code, this implies a level of protection around context and purpose which is not apparent from the review. Moreover, it may not be apparent for non-experts that anonymous and anonymised datasets are very different in nature. Rather than refer to anonymised data, it may be preferable to use the phrase 'de-identified' data. This

term more accurately describes the process followed to protect patient privacy and the residual re-identification risk that exists. Moreover, it has already been communicated to the general public by the socio-economic research community.[8]

6.7. As opt-outs will not be applied to anonymised data, it will be essential to maintain public trust that they are being processed safely. Thus, it will be important to ensure that data security standards are developed to cover these data sets as well. At the moment, there is a logical inconsistency in the proposals. The review recommends that the good practice advice contained in the Information Commissioner's Office *code of practic*e should be used to safeguard all de-identified data, and it recognises in paragraph 3.2.32 that this code mitigates the residual re-identification risk by the use of contracts and other controls. Yet, as we mention above (section 5), our interpretation of the review is that it only recommends that the ten standards are applied to identifiable data. In our view, an approach to data security should be developed that covers both de-identified and identifiable data, even though the nature of the controls might differ between these cases.

6.8. ***The review recommends that, in future, patient opt-outs will not apply to anonymised data. Such an approach would mean that it is much easier to use patient data to improve the quality and efficiency of the health care provided.***

6.9. Although the language needs to be clarified, and the data security standards developed, we welcome the suggestion that in future, patient opt-outs will not apply to anonymised information. There are some pragmatic reasons to follow this approach. It is common practice, both in the NHS and in other industries, to use de-identified data, and the approach taken by the review greatly clarifies what can be done and under what circumstances. As the review mentions, many health and social care professionals lack confidence in what they are allowed to do with data, and tackling this problem is important to help promote data security, for the reasons outlined in section 5.

6.10. Communication to the public is crucial. Patients need to be fully aware of what opting-out would mean for their data. Patients should not be surprised about what their data (in de-identified form) are being used for, therefore ongoing patient engagement will be important to the successful implementation of this recommendation.

6.11. ***Quality improvement is part of the day-to-day running of the NHS, but its data requirements have not been reflected in the review. Clarification is needed regarding its treatment, so that patients do not inadvertently opt out of data sharing for this purpose.***

6.12. The review uses a taxonomy about how data are used for direct care (pages 24-26) and for purposes beyond direct care (pages 26-28). We are concerned that quality improvement was not reflected in this taxonomy, and as a result, its treatment within the proposed opt-out remains unclear.

6.13. Quality improvement relies on data to construct metrics regarding the quality of the health care provided. These metrics vary between settings, but might include the number of

[8] The Administrative Data Research Network used the term de-identified (https://adrn.ac.uk/protecting-privacy/de-identified-data/)

wrong-site surgeries, the number of delayed discharges from hospital, or the number of prescribing errors within general practice. Data are often assembled within clinical microsystems (eg, individual operating theatres or wards) but may also be assembled across organisations (eg all operating theatres in a hospital) or across clinical pathways (eg maternity care). These data are used by local teams to identify opportunities to improve care, and to monitor the effectiveness of any changes made.

6.14. Quality improvement differs from each of the purposes outlined in pages 24-28 of the review. For example, quality improvement would not be classified as 'direct patient care' because, although it aims for a high level of clinical engagement, it concerns the care offered to multiple individuals. Nor could quality improvement be classified as 'monitoring health and social care services', since the examples given by the review relate only to national bodies (eg, the CQC's responsibility to monitor, inspect and regulate services and the work of NHS Improvement), or clinical audit. Finally, quality improvement is not research in the sense defined in pages 27-28. Unlike many research projects, quality improvement projects are typically conducted by people working in the NHS, and there is a very explicit link between the outputs of the analysis and steps taken to improve patient care.

6.15. This distinction is important because on page 40, the review describes a two-part opt-out model:

- **"**Allow my information to be used to support research to improve treatment and care."
- "Allow my information to be used to run the NHS and social care system."

6.16. We argue that quality improvement is part of the everyday work of the NHS. A series of reviews, including the Berwick report 'A promise to learn – a commitment to act', emphasised how quality improvement is a fundamental role of a health care system. The General Medical Council considers quality improvement an important factor in appraisals and revalidation[9].

6.17. Some further clarification is needed that quality improvement falls under the second purpose (ie about the use of information to run the NHS and social care system). While this may seem obvious, some difficulties emerge because quality improvement and research share a common goal to improve treatment and care (a goal that is prominent in the other opt out). Thus, it will be important to be clear about the distinction between quality improvement and research. Thankfully, there is experience to draw on from the approach to research ethics[10] (which distinguishes between service improvement, research and audit).

6.18. Unless this is addressed at the subsequent design stages, there is a risk that people working in the NHS remain uncertain about how data can be used for quality improvement, and this uncertainty might act as a barrier to critical work to improve the quality of care

[9] General Medical Council. Supporting information for appraisal and revalidation. (2012) Available at: http://www.gmcuk.org/RT___Supporting_information_for_appraisal_and_revalidation___DC5485.pdf_55024594.pdf

[10] Health Research Authority. Research Governance Frameworks. Accessed on 7 September 2016. Available at:http://www.hra.nhs.uk/resources/research-legislation-and-governance/research-governance-frameworks/

provided to patients.

7. Implementation (see Section 5)

7.1. Data linkage and de-identification of patient data are key techniques that are increasingly important in safe use of patient information. For the proposed opt-out / consent model to be successful careful implementation is required. Specifically:

- Local teams should be supported to develop and maintain standards, skills, software, systems and resources to de-identify patient information.
- In the meantime, it will be necessary that health and social care data can flow in identifiable form to NHS Digital to provide linkage and de-identification. From a practical perspective, it may be necessary in the short-term to exempt data collections transferred to NHS Digital from the opt-out, including those not covered by the Secretary of State's instructions.
- NHS data sets are increasingly diverse and will become more so over time, as new technologies are developed. The capacity of national organisations to receive, clean and link increasing amounts of data should be carefully assessed, and alternatives considered, such as a federated network of trusted third parties.
- The review does not recommend any changes to the existing arrangements until there has been a full consultation of the proposed new consent opt-out model. This seems quite sensible, as clearly there is a risk of causing further confusion if too many changes are made in this area. However, the Department of Health consultation document asks for views regarding what needs to be done to move from the current opt-out system to a new model, and we have shared some preliminary thoughts below**.**

7.2. Between the introduction of the opt-outs in 2014 and July 2016, approximately 1.5m type-1 objections were registered, along with almost 1.3m type-2 objections.[11] This represents a very significant concern about data security among sections of the population.

7.3. In transitioning to a new consent / opt-out model from the current system, a number of difficult decisions will be made. Therefore:

- Urgent research is needed regarding the characteristics of patients who have already objected to data sharing.
- In reaching a decision about the existing opt-outs, it is necessary to consider ethical requirements, technical ability, benefits and health inequalities.
  The Department of Health should consider 'passporting' people from the old system to the new system.

7.4. ***Local teams should be supported to develop and maintain standards, skills, software, systems and resources to de-identify patient information.***

[11] Data from NHS Digital (Care Information Choices, available at http://digital.nhs.uk/careinfochoices). The nature of type-1 objections means that individual-level data cannot be transferred to NHS Digital for analysis. As a result, it has not been possible for NHS Digital to de-duplicate the data, and the number of type-1 objections is likely to be an overestimate. Some individuals might have registered more than one type-1 objection, for example if they moved general practice. The data for type-2 objections have been de-duplicated and hence are more reliable.

7.5. The proposed opt-out / consent model has de-identification of patient data at its heart. In order to allow local teams to maximise benefit from quality improvement methods, and work with small evaluation teams available to them, it is crucial that de-identification can take place locally.

7.6. It is crucial that local teams are supported to develop and maintain standards, skills, software, systems and resources to de-identify patient information as the technical challenges are very significant. Health and social care datasets are already very diverse, and more datasets will emerge as the NHS continues to experiment with new ways of delivering health care (including the use of technology to diagnose health conditions and monitor them from a distance). Moreover, as the review acknowledges, initiatives aimed at more integrated care will produce new challenges regarding information sharing.

7.7. There are risks that, if the new approach is implemented before the technical infrastructure is operational, this will raise a barrier to innovation as data cannot be de-identified or opt-outs cannot be honoured. This is not a dilemma that anybody will want to face – a choice between delivering high quality health care or undermining public trust. Therefore, very high levels of assurance around technical infrastructure will be required before the new system is introduced.

7.8. ***In the meantime, it will be necessary that health and social care data can flow in identifiable form to NHS Digital to provide linkage and de-identification. From a practical perspective, it may be necessary in the short-term to exempt data collections transferred to NHS Digital from the opt-out, including those not covered by the Secretary of State's instructions.***

7.9. NHS Digital could play an important role in a transitional period where the proposed model is being implemented and while the technical infrastructure is being upgraded. We note that in paragraph 3.2.31, the review proposes that 'personal confidential data should be passed to NHS Digital, as the statutory safe haven of the health and social care system, to de-identify or anonymise and then share with those who need to use it." The review also states that patient data (in identifiable form) should flow to NHS Digital for patients who have opted-out, when instructed by the Secretary of State.

7.10. Consideration should be given, in this initial transition period, to whether all patient data should be allowed to flow to NHS Digital, irrespective of the Secretary of State's instructions. Failing this could stop new and innovative data linkages from happening, and limit the ability of the NHS to operate normally whilst improving the quality of care provided to patients.

7.11. ***NHS data sets are increasingly diverse and will become more so over time, as new technologies are developed. The capacity of national organisations to receive, clean and link increasing amounts of data should be carefully assessed, and alternatives considered, such as a federated network of trusted third parties****.*

7.12. It is clear that NHS Digital will have an important role in rendering certain data sets de-identified. However, at present NHS Digital has access to only a proportion of NHS data – perhaps less than 20% of all observations. The vast majority of data are used for clinical purposes and stored within primary, secondary, community or mental health care settings.

7.13. At present the review seems to have assumed quite a centralised model for de-identification, but this might have negative consequences for the ability of NHS teams to innovate. For example, many health care providers are introducing new technologies that collect information about the diagnosis or ongoing management of health conditions. In many instances, it will not be cost effective for NHS Digital to collect this data (particularly when the approach is at pilot stage and only used by a small number of providers). But the new framework for patient opt-outs will require the data to be de-identified to monitor the effectiveness of this service, when this is not forming part of a research project. Quality improvement work, as described above, which is an essential part of running the NHS cannot take place if no local capacity to de-identify data exists.

7.14. Although we would not want to pre-empt the detailed work about technical implementation, it seems likely that the approach will need to be de-centralised. In other words, the implementation team should consider an option that involves all health and social care organisations developing the ability to de-identify patient records. It is our experience that many information teams in health care providers and social care organisations are already able to do this, and software have been developed for this purpose. There may also be an argument, in certain circumstances, for other organisations to act as trusted third parties (for example, the Administrative Data Research Centres or the Office for National Statistics), this could be part of a federated system.

7.15. ***Urgent research is needed regarding the characteristics of patients who have already objected to data sharing.***

7.16. Although in some cases patients will be asked if they would like their data to be shared (eg when registering with some GP practices), in most cases patients have to actively engage with their practice to opt-out from their data being shared under the current model. This seems to suggest that those patients who opt-out have serious concerns related to data security. The way that options regarding opt-outs or consent are presented to the patient also differs as there is no standardised template (eg some practice provide forms with multiple opt-out options[12], where other forms only provide a single choice[13]). If only some patients are actively engaged on opt-outs, and methods available to patients to opt-out vary, datasets with opt-outs applied are likely to be biased, and affect analysis results.

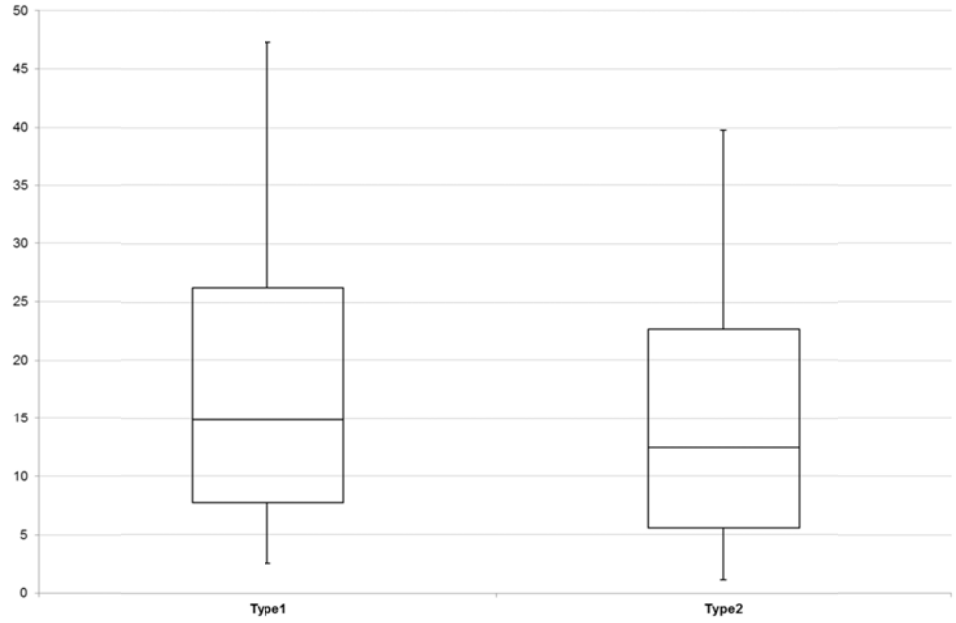
7.17. Unfortunately, no information has been published about the characteristics of people who have opted out – for example we do not know if they are concentrated in certain age or racial groups. It seems that this information will not be forthcoming, even in aggregated form, particularly for type 1 objections which are registered locally with general practices. However, what we do know is that there is considerable variation in the opt-outs by general practice (see Figure 1). As can be seen, there are some parts of England where a substantial proportion of patients have opted out of data sharing – upwards of 30%. To

[12] http://www.howdenmedicalcentre.nhs.uk/hmcoof.pdf
[13] https://medconfidential.org/wp-content/uploads/2015/05/caredata_trifold.pdf

inform our consultation response and try to understand why there is such variation, we assessed whether opt-outs were more common in socioeconomically deprived areas: this does not seem to be the case (see Figure 2).

7.18. In order to safely and correctly use patient data (with opt-out applied), and to better understand what needs to be done to transition from one opt-out model to the next, we argue that urgent research into patients who have currently objected to their data being shared is necessary. In the future, if the proposed model is implemented, NHS Digital can do this work, as they would be exempt from patient opt-outs.

Figure 1:[14] The variation in the number of objections per 1,000 registered patients, across all general practices in England (as at July 2016)



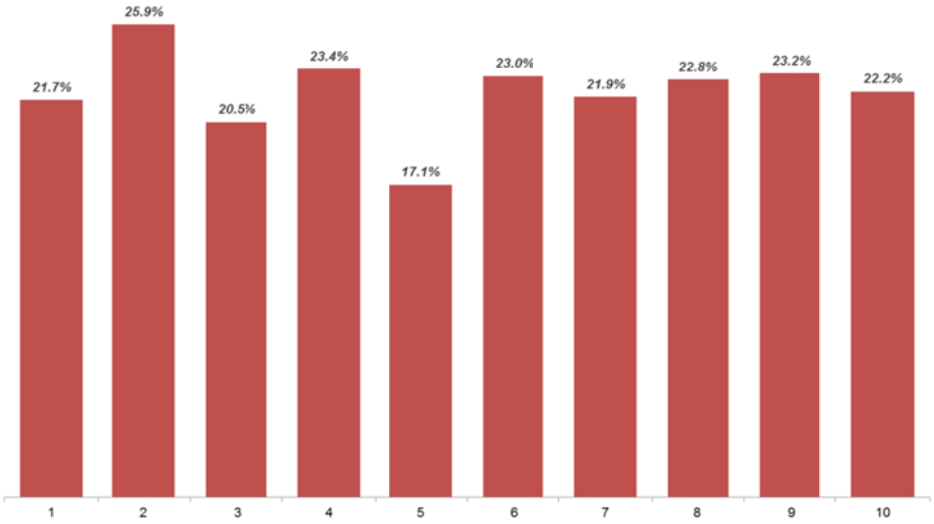
Over 1000 patients	Type1	Type2
10th	2.55	1.14
25th	7.74	5.55
50th	14.95	12.55
75th	26.28	22.60
90th	47.29	39.83

Note: In these box-and-whisker plots, the central horizontal line shows the median general practice (ie the 50th percentile). The upper and lower horizontal lines show the 75th and 25th percentiles of general practices, while the top and the bottom points in the vertical line show the 90th and 10th percentiles of general practices.

Figure 2:[15] Number of type-2 objections per 1,000 people, by decile of socioeconomic deprivation (as at July 2016)

[14] Health Foundation analysis based on data produced by NHS Digital

Where 1=most deprived and 10=least deprived



7.19. In reaching a decision about the existing opt-outs, it is necessary to consider ethical requirements, technical ability, benefits and health inequalities.

7.20. Following this review, there is a difficult decision about whether person-level data should be made available on people who have opted-out, if suitably de-identified, for purposes beyond direct care. There are four considerations, which must be weighed against each other:

- Ethical: the question here is whether sharing these data in de-identified form is consistent with the reasonable expectations that people had when opting out. As the review mentions in paragraph 3.2.9, there should be no surprises. If this data sharing would not be in line with these expectations, then much further communication will be needed to explain the change in approach, or otherwise public trust might be undermined.

- The technical ability to implement opt-outs: if opt-outs are only recorded once, but need to be consistently applied across all providers, commissioners, NHS Digital and other data controllers within the NHS, technical infrastructure is required to facilitate this. Implementation of the current opt-outs has proven difficult. Implementing the proposed opt-out model will likely prove equally challenging.

- The benefits of making de-identified data available for purposes beyond direct care: benefits of using data should outweigh potential harm to patients. Judging this benefit is difficult, especially when the benefit is realised outside the direct

[15] Health Foundation analysis based on data produced by NHS Digital and Department for Communities and Local Governmentt. Socioeconomic deprivation score is the 2015 Index of Multiple Deprivation, and was assigned to patients according to Clinical Commissioning Group. Equivalent analysis is not possible for type-1 objections.

care of patients. Data analysis provides information, which in turn can benefit patients depending on how the information is used.

- Potential impacts on health inequalities: opt-out rates can be related to certain patient characteristics. When processing data with opt-outs applied, analysts needs to be mindful of potential health inequalities. With very little information available on patient characteristics for the cohort of patients who have opted-out, it is very difficult to judge the potential impact of this.

7.21. The Department of Health should consider ‘passporting’ people from the old system to the new system.

7.22. When moving from one opt-out model to the next, it is very important to carefully consider the position of patients who have previously opted-out. In particular, consideration needs to be given to:

- How registered opt-outs in the current model are translated to the new model.
- If opt-outs on previously collected data will be honoured, or whether all data extracts (including retrospective ones) will have the new opt-outs applied.

7.23. Ongoing patient engagement will be important to explain data flows within the NHS, and in particular the role of NHS Digital in that flow. This engagement would also provide the necessary transparency, so reasonable expectations of patients who opted-out can be met. And a case needs to be made to patients why opting-out from data flowing to NHS Digital is no longer possible under the new model.

For further information:

Lewis Pickett
Public Affairs Officer
020 7257 8017
lewis.pickett@health.org.uk
www.health.org.uk