

An overview of care.data

What is care.data?

The NHS treats about one million people every 36 hours and consequently collects a vast amount of data. The care.data programme will establish a modern data service for the NHS which will improve the extraction of data, their collation and use to improve the quality, efficiency, safety, and equity of care for patients.

Everyone making healthcare decisions needs access to high quality information: clinicians need it to inform their decision making; patients need it when deciding which treatment option is best for them; and commissioners need it when making decisions about which services are right for their populations.

Safely handled and collated, data provides a much-improved, easier-to-access evidence base to support clinicians and healthcare commissioners find better treatment and ways of working, leading to improved outcomes and quality of care for patients.

Care.data is an improvement and extension of the data extraction and collation which has been routine in the NHS since 1989, when the Hospital Episode Statistics (HES) records began. Currently, identifiable patient data is automatically extracted from hospitals into the Health and Social Care Information Centre's (HSCIC) accredited safe haven for every inpatient, outpatient and A&E attendance. This data is used to generate aggregated statistics published regularly by the NHS, and has been used as the base data for countless academic research projects leading to revolutionary new treatments and ways of working.

The care.data programme will enable GP data from all surgeries in England to feed into the same system, linking GP and hospital records for the first time. This means health commissioners can better understand how the health service as a whole performs for patients as systems will then be able to track patients from their GPs into hospital and community care.

NHS England's vision for care.data is for as many specialist organisations as possible to be able to access, scrutinise and clinically assess the data the NHS holds, in a responsible way that leads to service improvement. We want charities, small research groups and responsible business to be able to access this data without prohibitive costs, so that all parts of the health and social care sector can fully contribute to healthcare improvement.

But most importantly, we also want all patients to be able to access their entire personal NHS records free of charge.

Who supports care.data?

The British Medical Association and the Royal College of General Practitioners have worked together with NHS England to produce guidance and information for patients and GPs. A number of key charities who conduct research and advocacy for patients with conditions such as cancer, heart disease and diabetes also support the care.data programme. Key supporters of the programme include:

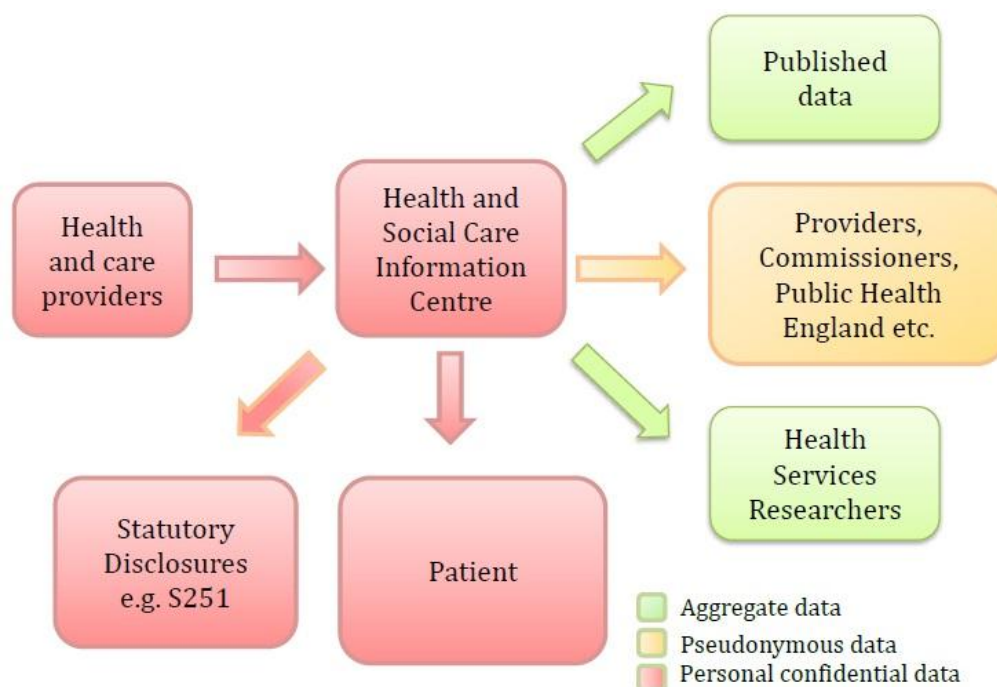
- Asthma UK
- Diabetes UK
- British Heart Foundation
- British Lung Foundation

- Cancer Research UK
- Genetic Alliance UK
- MIND
- Association of Medical Research Charities

How will patients' data be presented and published?

There are three broad categories of data sensitivity (red, amber and green), which require different levels of approval for access. The ways in which this data is handled, and the ways in which it can be accessed, remain exactly the same as those covering HES data.

- The "green" data flow contains completely anonymous, aggregated data which will be routinely published and easily-accessible. This includes statistics on the number of people with conditions like diabetes, attendances at A&E etc. It will be presented in a similar way to current HES data, and published on the new care.data website. NHS England and HSCIC comply with the ICO code of practice on anonymisation
- The "amber" data flow is not aggregated, meaning it is possible to identify individual patients' pathways of care. No individual is identifiable from this data, as it is securely pseudonymised in line with guidance from the Information Commissioners Office. This data will be available only to those who apply to the HSCIC fulfilling all legal requirements under the Data Protection Act and the guidance set down by the independent Information Advisory Group.
- The "red" data flow contains fully identifiable information. Fully identifiable information will be made available only to the patient themselves or where there is special legal approval under section 251 of the NHS Act 2006. In very rare circumstances it will be made available where there is an overriding public interest e.g. in the event of a civil emergency



How is data protected throughout the extraction and collation process?

Data is coded and encrypted as it is extracted from GP surgeries and hospitals. No free text flows from providers to the HSCIC's statutory "safe haven" secure system – there are no names or addresses.

The HSCIC's automated secure system extracts four pieces of identifying information linking to every item of data so that records from different parts of the NHS system can be accurately linked to one another to track full patient care pathways.

The system then strips the linked sets of data, either by completely anonymising and aggregating them (green data flow – see above) or by securely pseudonymising and scrambling them (amber data flow – see above), and converts them into a range of formats.

These systems are the same as those used for many years in the production of HES data both in anonymised and aggregated format for publication and in pseudonymised format for research and improvement organisations.

The pseudonymisation system used by the HSCIC involves complex algorithms, specifically designed so that they are extremely complicated. The key is changed for different data sent to different organisations for different purposes in order to protect the integrity of the data, the keys and individuals.

Whilst pseudonymising data does not completely remove all risk of re-identification, there remains a very small risk of re-identification in the event that data is put together with other very detailed data sources. The HSCIC mitigates that risk by only sharing data with organisations that have signed a robust data sharing agreement which restricts how they store, share, use and destroy the data. The agreement includes commitments that no attempt to re-identify patients is made.

Reactive Q&A

The content of this document will inform the briefing and form the basis of the FAQ offered to media.

1. What is care.data?

Care.data aims to improve the collection, collation and use of NHS data to ensure all NHS treatments and services are designed using the best possible evidence and improve quality of care for patients.

The NHS treats about one million people every 36 hours; a vast amount of data on the types of illness and injury we suffer, how quickly and effectively we are treated, and using which medicines and procedures, is held across the healthcare system.

2. Why is it important?

Anyone making healthcare decisions needs access to high quality information: doctors need it to inform their clinical decision making; patients need it when deciding which treatment is best for them; and commissioners need it when making decisions about which services are right for their populations.

Safely handled and collated, the data will provide an evidence base to support the research, learning and development of life scientists, clinicians and healthcare commissioners to develop better treatment and ways of workings leading to improved outcomes and quality of care for patients.

The data will be linked in the highly-secure accredited 'safe-haven' environment of the Health and Social Care Information Centre (HSCIC), which operates to the highest technical and security standards. At the Health and Social Care Information Centre, information is linked together with data from other parts of the NHS, stripped of identifiable information before it is compiled into dataset reports.

3. What has happened so far?

The Health and Social Care Information Centre has developed a highly-secure environment in which patients' data from their GPs, hospitals and other clinics can be linked together to show their whole pathway. Identifiable data is fed into the system, but no identifiable data is fed out of it, unless there is special legal approval.

4. What is happening now?

GP Practices are working with their patients to raise awareness of the fact that their data will be automatically extracted, via the GPES system (though GPs can use other electronic systems as approved by the HSCIC), to the HSCIC safe haven as part of the ongoing care.data project.

The timetable for extraction has not been finalised as we want to ensure the public and patients know about care.data and have time to discuss any concerns they have with their GP.

GP Practices are working hard to inform patients of how they can register objections which to stop their personal and confidential GP data being collected if they so wish.

5. What constitutes personal confidential data?

In the context of health and social care data, personal confidential information is any information which identifies a patient together with information collected from them in the course of their care.

This could include medical history, current conditions and treatment, and costs of treatment.

6. Can patients object? How can patients prevent their data being used in this way?

Patients can talk to their GP to register two types of objection.

First, patients can object to information containing data that identifies them from leaving their GP practice. This type of objection will prevent the identifiable information held in their GP record from being sent to the HSCIC secure environment. It will also prevent those who have gained special legal approval from using identifiable health information held in their GP record for research.

Patients can also object to any information containing data that identifies them from leaving the HSCIC secure environment. This includes information from all places they receive NHS care, such as hospitals. Data that may identify an individual will only leave the HSCIC where there is special legal approval (under section 251 of the NHS Act) for medical purposes e.g. for research. However if a patient objects their confidential data will not be shared for these purposes except in very rare circumstances for example in the event of a civil emergency.

This is a step forward as it will make it easier for patients to make choices about how their information is shared.

7. The GP FAQs say that patients who register an objection will not be able to stop their data being used in “anonymised data flows”. What does this mean? What constitutes anonymised data?

Anonymised data are data that do not identify individuals. Anonymised data showing raw numbers, for example numbers of appointments attended or numbers prescribed a particular medicine, has flowed throughout the NHS for some time. This will not change.

In order to be considered anonymised, data must have all identifiers including, for example name, address, date of birth, NHS number and postcode removed. Data will only ever be published under the “green” data flow - anonymised to the standards set out by the Information Commissioner’s Office.

8. How can identification be avoided where there are only small numbers of people with similar case histories, in both the anonymised and pseudonymised data flows?

The HSCIC has effective, long-standing techniques for the protection of individuals from identification in these circumstances.

In anonymised datasets for publication (“green” data flow), any information showing small numbers of people affected is restricted, with the numbers removed from the published information. For example, information about numbers of patients with extremely rare diseases would be restricted in this way. This is in line with guidance from the Information Commissioner's Office.

It is true that pseudonymising data does not remove all risk of re-identification. There remains a risk of re-identification if data is put together with other data sources. The HSCIC mitigates that risk by only sharing data with organisations that have signed a robust data sharing agreement which restricts how they store, share, use and eventually destroy the data, and includes commitments that no attempt will be made to identify individuals from the data.

9. If sensitive patient data has been flowing through the system for so long, why are patients only now being told about it and invited to register their objection?

Routine collection of data from hospitals predates the Data Protection Act, but all patients can object to their data being used in this way, and currently some patients do so.

NHS England is committed to transparency and openness, and is working with GPs, patient groups and key stakeholders to ensure patients are aware of how their data can be used for health and care improvement, and how they can object to it being used in this way. We are making it easier for patients to object to secondary care data flows.

We are committed to making sure all patients and the public are aware this is happening and know how to object they wish. We will be launching a high-profile, national patient and public information programme to make sure that every household in the country knows what is happening, when, and how to find out more.

We have already started sending materials to GP practices to support them in raising awareness with their patients. The materials make it clear that patients have a right to object, and that their objection will be respected. In addition to current awareness activity through the regions and locally, NHS England and the HSCIC have worked together to develop the care.data website, under which patients and clinicians can find out more about how data is used, as well as about how to request extracts from the system once it is fully up and running.

The HSCIC and NHS England continue to work closely with the Information Commissioner's Office to ensure plans for the rollout of the programme are shared and that the ICO are content with the approach.

NHS England and the HSCIC have been working together with Dame Fiona Caldicott and the National Information Governance Board for Health and Social Care to ensure that all patient data is thoroughly protected and that patients understand how their data is used and shared.

NHS England and HSCIC have asked the Information Advisory Group to review the care.data programme on an on-going basis to ensure it meets the highest possible standards of information governance.

10. Are you having a national campaign because the ICO has demanded it?

No. NHS England and the HSCIC are committed to ensuring that all patients and members of the public are aware of how their healthcare data are used and shared and their right to object.

There have been a number of activities at national level for some time (e.g. Info. On nhs choices etc) and we are pleased to announce a household leaflet drop as part of these awareness raising activities. We have been liaising closely with the ICO to ensure that they continue to be happy with our plans

11. Why are you doing a leaflet drop? Isn't it a bit 'old-fashioned'?

We want to make sure everyone is aware of what is happening. Office of National Statistics research about internet access shows usage is particularly low in over 65s, who are the heaviest users of the NHS.

According to ONS figures, in 2013

- 4 million or 17 per cent of households had no internet connection
- In single adult households where the adult is aged 65 or over, only 40 per cent had an internet connection

While you can obviously still access the internet on phones and at public libraries for example, there is a very strong argument that doing this digitally would risk leaving members of the public out.

12. Will personal data be sold to private companies?

No. There will be no commercial sale of data, and there never has been. The HSCIC does not charge for data. Currently, the HSCIC may charge an administrative fee (for example, to link the data) but there is no commercial sale of NHS data.

NHS England and the HSCIC believe that pseudonymised information should be made available in future to approved organisations outside the NHS, on a case-by-case basis for a range of purposes that will benefit patients, such as new service and other innovations in healthcare, developing predictive models, and detecting unwarranted variations in care.

NHS England and the HSCIC believe that linked primary care and secondary care data is useful for a wide range of purposes and that it would be irresponsible not to make the maximum use of this data. Such purposes include designing, facilitating and evaluating integrated care; identifying where improvements and efficiencies can be made; reducing variation and improving the equity of services; identifying patterns of diseases and outcomes; and developing sophisticated tools such as predictive models, impact models, dashboards, and data visualisations – all with the aim of improving healthcare and outcomes for patients.

NHS England's vision for care.data is for as many specialist organisations as possible to be able to access, scrutinise and clinically assess the data the NHS holds, in a responsible way that leads to service improvement. We want charities, small research groups and responsible businesses to be able to access this data without prohibitive costs, so that all parts of the health and social care sector can fully contribute to healthcare improvement.

We also want all patients to be able to access their entire personal NHS records free of charge.

At present, the care.data GP extract will only be used by NHS commissioning organisations, in anonymised and aggregated form to help them plan healthcare services on behalf of their local populations.

The HSCIC will only share this type of “amber-level” data with organisations that have signed a robust data sharing agreement which restricts how they store, share, use and ultimately destroy the data, and includes commitments that they make no attempt to re-identify patients.

13. How the whole programme is being funded and by whom?

Please refer this question to Geraint Lewis.

14. What legislation is underpins the care.data programme?

The care.data programme is underpinned by three key pieces of legislation: Health and Social Care Act 2012, common law duty of confidentiality and the Data Protection Act 1998.

Health and Social Care Act 2012

The Health and Social Care Act 2012 provides the legal basis for care.data. The Act introduces legislation, which empowers the Health and Social Care Information Centre to require organisations to submit data to it when directed by the Secretary of State for Health or NHS England. The first use of this power, under the Act, is for care.data. Directions for the extraction of primary care data have been approved by NHS England and the HSCIC Boards and we are in the process of formally issuing them to the HSCIC.

Common law duty of confidentiality

The common law is based on previous court judgments, which have established that patient consent is generally required for the disclosure of confidential health information. The Health and Social Care Act 2012 however, explicitly overrides the common law duty of confidentiality and the requirement to seek patient consent when the HSCIC has been directed to collect the data as described above. This law was passed because data collected as part of care are so valuable and necessary for the NHS to function and it was considered there was a greater good in ensuring this data can be linked by the HSCIC. Whilst the common law duty of confidentiality has been overridden for the collection of data by the HSCIC, it applies in relation to the disclosure of confidential data from the HSCIC. The HSCIC can only therefore disclose confidential data where there is a legal basis e.g. where there is special approval for research under section 251 of the NHS Act or patient consent.

Data Protection Act 1998

The Data Protection Act 1998 regulates the processing of confidential information about living individuals and covers paper and computer records. It sets out eight principles or rules to protect data. Particularly relevant to care.data is the requirement for data to be processed fairly, which in practice means that a person's

personal data is handled in ways that they would reasonable expect. Gps, as data controllers, are therefore required under the DPA to ensure that patients are aware of how their information is used and shared. We are working with the ICO to ensure that they are content with our awareness raising plans

15. What legal approvals do we still need to negotiate in order to access patient data?

The powers in the Health and Social Care Act 2012 allow for data to be extracted.

16. What do the legal definitions of terms like “pseudonymous”, “anonymous”, “accredited safe haven” and “personal, confidential data” mean?

The Information Commissioner (ICO) provides definitions of personal data and sensitive data:

http://www.ico.org.uk/for_organisations/data_protection/the_guide/key_definitions

The ICO has also produced a code of practice on anonymisation, which both NHS England and HSCIC comply with when referring to anonymised data.

There isn't a legal definition of pseudonymous but for care.data, NHS England and HSCIC will comply with the ICO's guidance.

'Safe haven' uses the authoritative Caldicott review definition which is where 'specialist, well-governed, independently scrutinised and accredited environments are called 'accredited 'safe havens'.

17. Are there any aspects of the programme are currently subject to review? If so, who is carrying out the reviews and what are their terms of reference?

An Independent Advisory Group considers changes to the scope of the GP extraction. We are in the process of seeking approval for analysts such as researchers to have access to pseudonymised data so that the health service can benefit from the value they can bring.

18. What is the position regarding continued extensions of s251 agreements and plans for future legal agreements?

Section 251 is relevant to care.data in relation to the outputs from the HSCIC, for example confidential data can only be disclosed by the HSCIC where there is a legal basis with section 251 approval.

Approval has been sought for other uses but this is being considered by CAG so we cannot provide further detail.

19. How does Fiona Caldicott's view that commissioners do not need to access patient data impact on care.data?

NHS England is clear that commissioners of services can only access pseudonymised data for care.data.

20. What about making the data available more widely than commissioners – have you got the necessary approvals?

See FAQ 17.

21. Didn't the IAG reject an application for phase two of the care.data programme?

No, they requested further information which we are providing.

22. When do we expect to be able to make the data available to researchers and business?

Once we have responded to the IAG recommendations - we are hopeful that this will be soon.

23. Will there be further awareness-raising work at this stage to make sure patients and the public know about this?

Yes, there is awareness raising taking place e.g. we are cascading messages via stakeholders and using social and digital media, there will be regional awareness raising using regional communication channels e.g. radio and regional newspapers in addition to the awareness raising at GP practices.

24. Isn't this just an example of 'scope creep'?

No. It has always been our intention for analysts such as researchers to access pseudonymised data.