

EXECUTIVE SUMMARY

The UK can significantly enhance its clinical research capability by using, strictly within the bounds of patient confidentiality, the electronic patient data that the UK's National Programmes for IT in the NHS have the potential to allow. This will have enormous benefits for all types of clinical, public health and health services research and for many aspects of patient care.

The UK Clinical Research Collaboration's (UKCRC) Research & Development Advisory Group to Connecting for Health therefore commissioned a series of simulations in October 2006 to provide the Department of Health Directorate of Research and Development, and NHS Connecting for Health (NHS CfH), with detailed specifications for a range of possible research applications. The objective was to:

- ▶ Inform future development of the NHS Care Records Service (NHS CRS)
- ▶ Highlight technical, regulatory and governance issues
- ▶ Inform plans for any further simulations and full pilots to test the capacity of the infrastructure, using real patient data with appropriate safeguards when this becomes feasible.

Four simulations were commissioned, based on a range of clinical research applications. These were: interventional clinical trials; surveillance; prospective tracking of an identified cohort; and observational epidemiological research. Detailed reports and key findings were presented to the Advisory Group in February 2007 and form part of this report.

The simulation leads worked as a team over this period and there was strong consensus in relation to both the high level and more detailed messages

emerging from their work. They have identified a number of key data, regulatory and governance issues that need to be addressed for future development:

- ▶ Clinical services and research share the same mission of improving patient care and patient safety: **research is integral to patient benefit**
- ▶ Research makes a very important contribution to assessing the completeness and quality of data used for clinical care and health services
- ▶ Leadership is needed to create the sustainable and governance infrastructure required to exploit the research opportunities afforded by routine patient and other data
- ▶ Solutions should be addressed from a UK-wide perspective and build on the extensive experience with record linkage already in place
- ▶ Much of this research involves information about groups of patients rather than *individuals* and hence requires anonymised rather than identifiable data. However there will be occasions where data needs to be linkable (possibly by an 'honest broker') and comprehensive at the individual patient level in order to have maximum value and to allow quality and completeness to be validated
- ▶ Where data are required at individual patient level, such data access will need to be to pseudonymised data. Where identifiers need to be retained, appropriate consent must be gained as part of enabling access to those data
- ▶ Existing UK strengths in the use of routine and other patient data for research will be significantly enhanced by the mandated use of a unique identifier (for example NHS number) in all

key NHS records and activities and by ensuring that new and existing data sets are person-based

- ▶ The data made available must cover the whole population, be up-to-date, and be retrospective over a number of years to give a rich historical picture of patients' health and care. They must also be accurate and based on high-quality input
- ▶ Further work to confirm the detailed requirements for data, which have been spelled out in each of the individual simulation reports, will need to be finalised. Much of the same data required for purposes such as safety monitoring and clinical trial research is of interest for public health and NHS management activity including monitoring service delivery. So there is a high degree of commonality in the data needs
- ▶ The breadth of data needed for the potential research applications explored in the simulations supports the concept of a data switchboard, with potential to link NHS Care Record data widely to other data sources. Thus future strategic developments should be based on this premise, rather than that of a single data warehouse
- ▶ A federated structure of data sources rather than a single data warehouse would also provide an effective infrastructure with optimal governance systems in place. This could be an honest broker with responsibility for removing identifiers, linkage of data and data quality checks

Tackling regulatory and governance issues successfully will be key to ensuring appropriate access and use of the data for research purposes.

- ▶ Ensuring patient confidentiality is critical. Data governance must be robust and at the same time capable of facilitating research

- ▶ Although for some research applications fully anonymised data will suffice, for many research applications pseudonymised data is required to enable linking of data sets or elimination of duplicate records. However, for other research purposes it will be important for patients to be contactable in an appropriate manner. Appropriate approaches to consent will need to be built into access mechanisms for information which might be capable of being linked to a specific patient
- ▶ In order to satisfy regulatory requirements for purposes such as pharmacovigilance and for clinical trials research, there are specific data quality and access requirements that need to be addressed

The dual role of the honest broker in ensuring patient data confidentiality and security as well as scientific integrity of data delivered to the research community will be key to engendering trust amongst patient, clinical professional and research communities.

The potential benefits for research will be lost unless these issues can be addressed.

It is critical that the needs of research be formally prioritised so that both individual healthcare and public health can reap the full benefits of this NHS resource. The recommendations are summarised below.

Key recommendations of the UKCRC simulations

Quick Wins

Recommendation 1: Mandate a common patient identifier

To enable linkage of sources of data at patient level a unique patient identifier will be required: use of the NHS Number should be mandated in all key NHS records and activities, including laboratory records.

Recommendation 2: Communicate the relevance of research to healthcare

There should be formal recognition that research is a core, not secondary, component of the development of the NHS Care Records Service as it benefits patients directly. Objectives, strategy and resources need to be committed or endorsed at the highest level of NHS Connecting for Health and reflected in its literature including website content.

Short Term Deliverables

Recommendation 3: Federate existing databases

A federated structure of data sources is required for research. A high-level strategy to support such an infrastructure needs to be developed together with a roadmap for its delivery. This strategy should ensure that the data made available cover the whole population, are up-to-date, person-based and of high quality, and extend back over a number of years to give a rich historical picture of a patients' health and care.

Recommendation 4: Improve data quality

Data quality is of paramount importance both in the clinical setting and for research. Data should be accurate (relying on high quality input) and based on a set of standards for recording and processing data. Ongoing processes will need to be developed to improve data completeness and quality which could involve development of incentives.

Recommendation 5: Initiate governance discussions

Tackling regulatory and governance issues successfully will be key to ensuring appropriate access and use of the data for research purposes. Data governance must be robust and at the same

time capable of facilitating research.

Recommendation 6: Engage key stakeholders

It is essential to engage professional audiences who are key to implementation, particularly for the enhancement of data quality and improving data access. Patient safety is of importance to all audiences and should be at the forefront when communicating the value of research. A communication strategy regarding the joint benefits of using patient data for research and clinical care needs to be developed. The responsibility for this development and those in recommendation 5 above will be with the Care Record Development Board and may subsequently transfer to the National Information Governance Board upon its formation.

UK-Wide Strategy: Next Steps

In informing plans for next steps, the outcomes of the simulations suggest that more extensive data are required to enable research than those currently available through the Secondary Uses Service.

We recommend that an approach that relies on a federated system of databases should be based on a UK-wide strategy.

This will require:

- ▶ Initiation of pilots to link datasets, on the basis of existing successful examples within the UK;
- ▶ Definition of methods of access to the different sources housing the data. This should include access to detailed patient-level data from primary care, pathology services, disease registers and key private sector services;
- ▶ Future development which learns from, and build upon, existing skills, knowledge, databases and systems that have been developed in the UK over many years;
- ▶ Adoption of a UK-wide approach: not only are there examples of good practice beyond England that can be built upon, but the future development should ensure compatibility across the UK;
- ▶ An organisation capable of managing the specification and delivery of the required infrastructure and providing the linkage and definition of research support services.