

Data Sharing Review Secretariat
5.26 Steel House
11 Tothill Street
London
SW1H 9LH

15th February 2008

Cancer Research UK is the world's largest independent organisation dedicated to cancer research. Last year we spent £315 million on research, which included research based on the using and sharing of data. We welcome this review of data sharing and look forward to review recommendations that will support the use of patient data in research.

Access to patient data is especially important in cancer, because epidemiological research is pivotal to our understanding of cancer as a disease in a way that is not true for many other conditions. As two-thirds of cancers are potentially preventable, patient data have been especially important to help identify some of the factors that put people at risk of cancer in the first place. Registry data were crucial in defining the causal relationships between tobacco and cancer, between asbestos and mesothelioma, and in indicating the role of diet in cancer prevention. These data are also important in uncovering trends in cancer, such as evidence that skin cancer rates have been increasing year on year; and that some cancer drugs can themselves increase the risk of developing leukaemia or other cancers. Data from cancer registries can also be used to track and assess variations in treatment that cancer patients are receiving around the country.

Cancer Research UK supports the submission made to this review by the UKCRC. The submission highlights key issues that influence data sharing for research purposes such as the disproportionate time, effort and cost required to obtain access to data, and the attitude of the Patient Information Advisory Group (PIAG) to benefits of health research.

In addition, we would like to share with this review two examples of data sharing that we feel would be valuable additions to your review. The first example is the Comparison Arm for ProtecT (CAP) study, and the second is the Health Information Research Unit (HIRU) for Wales and the work it is undertaking. Full details of both of these examples are attached.

Cancer Research UK examples of Data Sharing in Practice

Comparison Arm for the ProtecT Study

The Comparison Arm for the ProtecT (CAP) Study compares the risks and benefits of population based prostate-specific antigen (PSA) testing to those of current NHS care for the management of prostate cancer. The study involved 50-69 year old men at 800 general practices. Many of the men involved in this study had given written consent through their involvement in the ProtecT Trial. However investigators concluded that requesting consent from men used in the comparison arm would result in a study sample size that would not be adequate for statistical analysis. In addition, it would not be possible to anonymise the data because of the need to link the records to men diagnosed with prostate cancer.

In July 2003 Cancer Research UK confirmed funding for the trial. However it took until March 2006 for the Patient Information Advisory Group (PIAG) to make its final decision on the research applications for the study. This meant that £560,000 of public and charity funding was spent on research staff who were unable to do a large part of their work until the final year of the three year grant. The experiences of Dr Chris Metcalfe, and the other researchers involved in this application process are discussed in full in a paper published earlier this year in the *Journal of Medical Ethics* (*J Med Ethics* 2008;**34**:37-40, attached).

It is becoming increasingly apparent that at times PIAG is resistant and unsupportive of research. PIAG also appears insensitive to the timeframes that the research community works within.

Health Information Research Unit for Wales

The Health Information Research Unit for Wales (HIRU) was established to harness the potential of routinely collected electronic data to support and undertake research. HIRU is working with organisations such as GP practices and NHS trusts to bring electronic data together.

In order to achieve this, HIRU is continually:

- developing new methods for accessing and combining routine data in ways which do not breach data confidentiality rules;
- devising ways to use routinely collected data to support large scale multi-site intervention, cohort and policy relevant studies;
- developing innovative approaches to knowledge discovery from the analysis of large and combined datasets, using supercomputing facilities; and
- developing methods for research data capture to common standards and definitions in multiple and remote locations.

As part of their work HIRU is also building a large-scale data warehouse of anonymised but linked information about individuals.

Secure Anonymous Information Linkage (SAIL)

HIRU has established the Secure Anonymous Information Linkage (SAIL) Data Bank project. The SAIL Data Bank has several purposes including developing data extraction methods, assessing data quality issues, and simulating research questions.

In total, 105 million records have been loaded into the SAIL Data Bank, with 2.4 million individuals assigned a unique key, enabling any individual to be anonymously linked across various datasets. This information includes data from GPs, hospital admissions and outpatient appointments, with a 99% success rate in unique key allocation.

HIRU will be taking forward the learnings from the development of the SAIL Data Bank in order to roll out the data bank across Wales.

Working in partnership

It is necessary for HIRU to work with organisations in order to gain access to the patient data. However HIRU recognises the value of partnership and has established relationships with many organisations who have provided both data and support. One relationship of note is that with Health Solutions Wales (HSW). HSW have been central to the success of HIRU playing an integral role in the design of the anonymisation process in addition to their role as a data provider.

Other data-providers contributing to the SAIL Data Bank include Swansea Local Health Board (SLHB), the Department of Social Services from the City and County of Swansea (SSS), and Swansea NHS Trust. Recognising the value that the SAIL Data Bank can deliver in providing insights into the way health and social care is provided in the locality, these organisations deliver their information on a voluntary basis.

Supporting Research Questions

The primary purpose of HIRU is to enable research. One way in which HIRU achieves this is as part of the Clinical Research Collaboration Cymru, through which it works with the Thematic Research Networks and support services. As part of the research infrastructure HIRU has supported many research questions including the following up of survival rates for specific electronic cohorts, a retrospective study of intentional self harm behaviour by gender and antibiotic prescribing versus rates of complications due to infection.

Providing Benefits to the Data Provider

Whilst the primary role of HIRU is to support research there are also benefits for the data provider organisations. In order to maximise these a SAIL Project Benefits Group has been established. The group has successfully contributed to providing more informed incidence and prevalence data for chronic diseases and profiling health and social events for care pathways.