Chapter 2

Taking stock of the evidence – from data use to health system improvement

Many countries are benefiting from the linkage and analysis of personal health data to provide the evidence needed for health policy decisions to improve the quality and efficiency of health care. Examples range from reporting on the cost-effectiveness and clinical appropriateness of care in Finland, Korea and Singapore; to assessments of the quality and efficiency of clinical guidelines in Sweden; to evaluating the safety of patient screening in Germany; to evaluating the quality of surgical outcomes in Israel and the United Kingdom; to examining care transitions in Australia and Canada.

This chapter summarises 29 within-country projects and 10 multi-country projects deemed by country respondents to be policy relevant and to exemplify good practices in data protection. Among them, 14 study leaders were interviewed to provide additional information about their project and its relevance to health policy, as well as the steps taken to ensure privacy-respectful data use. For these 14 projects, a detailed case study summary is presented.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

There is a very large and growing body of evidence of the importance of the collection, analysis, linkage and reporting of results from personal health data assets for health care quality monitoring and improvement, population health policy, and health system performance measurement and evaluation. This OECD study asked respondents to identify up to three studies conducted in their country over the past five years that were relevant to policy makers and that demonstrated best practices in the protection of data confidentiality, respect for patient privacy and privacy legislations, excellent data security, high quality data, and a sound study methodology. Further, this study also asked respondents to identify a recent multi-country project involving the analysis of personal health data from health care administrative databases, disease registries or electronic medical record databases.

There were several very important examples provided by countries of the linkage of personal health data to follow the pathway of care and understand the health outcomes of care in order to evaluate the quality and effectiveness of health care treatments. The PERFECT study (Table 2.5) in Finland monitors the content, quality and cost-effectiveness of treatment episodes in specialised medical care and thus contributes to monitoring health system performance. The methodology developed for PERFECT is now having an impact on monitoring among other countries throughout Europe via the EUROhope study (see below). Korea's quality assessment of medical services includes assessment of the clinical appropriateness and cost effectiveness of health care by reporting on quality and inducing service providers to make improvements in response to the evidence (Table 2.7). It aims to identify underuse, overuse and misuse of therapies and to reduce variation in care practices through the regular reporting of quality indicators. There are also quality and efficiency assessments of clinical care guidelines in Sweden (Table 2.8). For areas of care subject to national guidelines, such as cardiac and stroke care, care for selected cancers, dental care, diabetes care and mental health care, data linkages are undertaken to develop indicators to evaluate the effectiveness of recommended therapies and the evidence contributes to revisions of the care guidelines. To monitor and study health care consumption and expenditures, Belgium has developed a permanent sample of socially insured persons via the linkage of health care reimbursement invoice data to create longitudinal histories of health care encounters (Table 2.1). Results inform policy decisions to manage health care expenditures.

In Germany there have been projects to evaluate the effectiveness and safety of breast cancer screening (Table 2.6). A project examined the quality of breast-cancer mammography as a diagnostic tool and involved a follow-up of women who had experienced breast pain or a suspicious lump through subsequent health care encounters and cancer outcomes. A second project involves an evaluation of early detection guidelines for mammography screening where patients who participated in a clinical trial and those who did not will be followed up for health outcomes. In so doing, the benefits and the potential adverse effects of exposure to mammography screening can be evaluated and the evidence used to develop policy.

Two data linkage projects are underway in the United Kingdom to improve understanding of infant health (Table 2.11). These involve overcoming gaps in existing information to provide a more comprehensive and consistent picture of maternity outcomes and to enable statistics of births and infant deaths by key characteristics, such as gestational age and ethnicity. To extend the information available about pathways of stroke care beyond the acute care setting, a pilot data linkage project is underway in Canada (Table 2.2).

In Switzerland, a linkage of population Census data and mortality data is enabling a better understanding of the socio-economic and socio-demographic characteristics of mortality and life expectancy and forms a base cohort from which additional data may be linked for specific, approved, studies (Table 2.9). In the United States, a platform has been developed to support health and health services studies, including a repository of surveys that have been readied to support linkage projects and two key linkages: the linkage of population health survey data to mortality data; and the linkage of population health survey data to data on health care encounters for Medicare and Medicaid insurance beneficiaries (Table 2.13). In the United Kingdom, there is an initiative to facilitate research involving personal health data that is in the public's interest. The service can both produce tabulations and conduct data linkages on behalf of clients with approved projects (Table 2.10).

A care trust in England has a new project to link records across health and social care databases to produce pathways of services and associated costs on an on-going basis (Table 2.12). A Canadian province has established a university-based research centre to conduct linkage of individual-level data to inform on health system performance, patient safety, population health, diagnostic services and primary care (Table 2.3). A health care maintenance organisation in the United States has accumulated 50 years of experience in research and monitoring involving personal health data and linkages to improve health care services and patient outcomes and is now analysing data from an electronic medical record system (Table 2.14).

Other examples of projects involving data linkages summarised in this chapter include health care quality monitoring through understanding care pathways and outcomes for chronic disease patients, for cancer patients, for patients suffering a heart attack, and for patients after key surgeries; studies of the health effects of radiation exposure; the development of disease registries; monitoring pregnancy outcomes; and monitoring health care use and expenditures.

Multi-country projects

The European Best Information through Regional Outcomes in Diabetes (EUBIROD) project (Table 2.4) is a public health project funded by the European Union that aims to implement a sustainable European diabetes register to monitor diabetes complications and the health of diabetes patients (EUBIROD, 2011). EUBIROD is amalgamating aggregate data from 18 diabetes registries across Europe and it was challenging for the participants to find common ground where the local requirements for data security and privacy would be respected. The solution was the Best Information for Regional Outcomes or BIRO system (Di Iorio et al., 2009). In BIRO, each disease registry provides aggregated data for their region with very little to no re-identification risk using an on-line data transfer system. In working with participating countries, the conclusion of the EUBIROD team is that the sharing of

de-identified person-level data from diabetes registers would not be possible and still succeed in securing the participation of a large set of countries.

The Nelson trial is a randomised trial of the potential to use low-dose CT scans to screen at risk patients for lung cancer (van Klaveren et al., 2009). Data are from Belgium and the Netherlands. The trial began in 2004 and is continuing until 2015. The world is waiting for the trial results because this is the only study where patients were recruited from population registries where it could be certain that those in the no-screening group indeed had not been screened. Results will have worldwide implications for health system policy regarding the uptake of and guidelines for lung-cancer screening.

EuroHOPE, the European Health Care Outcomes, Performance and Efficiency project, is a new initiative funded by the European Union and co-ordinated by the National Institute for Health and Welfare in Finland to evaluate the performance of European health care systems in terms of outcomes, quality, use of resources and costs through data linkages. Participating countries all have the necessary health information infrastructure and legal framework to undertake the data linkages and include Norway, Sweden, Scotland, regions in Italy and the Netherlands. For EuroHOPE, each participating country will link health care administrative databases for in-patient hospitalisations, pharmaceutical data, and cancer registry and mortality data in order to begin to generate indicators of the quality of hospital-based treatments across the whole cycle of care that would be comparable across the participating countries. The five focus areas for the development of these health care quality indicators are acute myocardial infarction, stroke, hip fracture, breast cancer and low birth-weight infants.

EuroHOPE aims to develop indicators that could be recommended to the European Union for routine reporting, develop methods for international comparative health services research based on data linkages of person-level data; and inform about the policy-relevant drivers of health care quality, including treatment practices, use of medicines and new medical technologies, waiting times, financing, and the organisation of care. EUROHOPE is following the analytical model established by National Institute for Health and Welfare in Finland (Table 2.5).

EURO-PERISTAT is a European project to monitor and evaluate perinatal health in the European Union by establishing a sustainable system for reporting perinatal health indicators (EURO-PERISTAT, 2011). The Deepening our Understanding of Quality Improvement in Europe (DUQuE) project is funded by the European Union to study the effectiveness of quality improvement systems in European hospitals by assessing the relationship between hospitals quality improvement systems, management and culture and the quality of hospital care, such as clinical effectiveness, patient safety and patient involvement (DUQuE, 2011).

There are numerous multi-country projects exploring dimensions of cancer incidence, treatment and survival. The *Cancer Incidence in Five Continents* series, published by the World Health Organisation is a reference for international comparison of cancer incidence (IARC, 2011). EUROCARE is a study of cancer survival across Europe (EUROCARE, 2011). Australia, Canada, Denmark, Norway, Sweden, and the United Kingdom participated in an International Cancer Benchmarking Project to better understand both how and why cancer survival varies among countries (Coleman et al., 2011).

Denmark, Finland and Sweden collaborated in a study of mortality and life-expectancy trends from 1987 to 2006. This study required each country to link hospital

discharge registers to cause of death registers to examine the excess mortality and life expectancy gaps for people hospitalised with severe mental health disorders. Results helped to inform about the quality of psychiatric services (Wahlbeck et al., 2011). Canada and the United States collaborated to conduct the Joint Canada-US Survey of Health (JCUSH) to compare access to and use of health care services and population health between the two countries (Gulley and Altman, 2008; Altman and Gulley, 2009).

Case studies

Table 2.1. **Belgium: Permanent sample of socially insured persons**

Study title	Permanent sample of socially insured persons
Lead organisation	National Institute for Health and Disease Insurance (INAMI-RIZIV), Belgium.
Project description	Belgium has developed a permanent sample of socially insured persons involving the collection and linkage of health care reimbursement invoice data from the seven Belgian health insurance organisations (Commission Technique d'Échantillon Permanent au Conseil Général de l'INAMI, 2011). Data are linked to create longitudinal histories of health care encounters to study health care consumption and expenditures. The seven health care insurance organisations entered into a partnership for this project and the data is collected and linked by a trusted third party operating on behalf of the insurers called the Intermutualist Agency or IMA-AIM. Other partners in the project are government departments including the Belgian Health Care Knowledge Centre (KCE), National Institute for Health and Disease Insurance (INAMI-RIZIV), and the cancer registry. The management committee of the project consists of the partners and representatives of the Belgian Privacy Protection Commission.
Project approval	The permanent sample is authorised by law. The law establishes the IMA-AIM as the party that would select from the universe of social security numbers in Belgium a representative sample of the population of one in 40 persons and one in 20 persons aged 65 and older. The law requires that the data only be used for statistical purposes related to management and research and forbids participating partners from undertaking operations that might directly or indirectly identify an individual. The management committee is authorised to approve studies using the permanent database. This includes the approval of any extension of the database involving linkages to other databases within the custody of the partners (Committee for Social Affairs). Extensions of the database involving linkages to databases outside of the participating partners, such as linkages to databases in the custody of the Committee for Public Health, are not part of the legal authorisation and would require all of the steps for approval of the Privacy Protection Commission of any new project and may require the legislation authorising the permanent sample to be amended.
Data and data linkage	Included in the permanent database are records from primary health care, a subset of hospital data and information on reimbursed medications. Data is composed of reimbursement codes by procedure, service, admission, and drug delivery that include dates, providers, institutions and costs. Virtually all health care encounters have an associated Social Security Number. Missing from the data would be a small number of cases of foreign persons receiving health care and infants born in hospital who have not yet received a Social Security Number. According to the law authorising the permanent database, Social Security Numbers within the micro data are re-coded by the seven health insurance mutualities before being sent to an intermediate party which also re-codes the social security number a second time and then transmits the micro data to the IMA-AIM. The insurers also further de-identify the data by removing other direct identifiers including names, exact birth dates and addresses. The IMA-AIM link the records using the coded Social Security Number to create a longitudinal view of each selected insured person in the sample. Partners are provided access to micro data that includes the coded Social Security Number, the year of birth and a city code. IMA-AIM provides partners with access to the permanent sample using a secure electronic data transfer. Data for the sampled insured persons is linked for a maximum of ten years before it is destroyed. Each year a new sample is drawn and thus the database itself is maintained permanently.
Protection of data privacy	There is no prior consent requested of insured persons to be selected for inclusion in the database. However, Belgian health interview surveys can only be approved to be linked to the permanent sample database if the survey respondents have consented to this linkage. Government partners receiving micro data are required by law to control access to the permanent database. Only a small number of individuals (4-5 persons) are permitted access to detailed micro data within each organisation. More aggregated data views where, for example, year of birth is grouped to five year intervals, may be accessible to a broader number of employees. Government partners are required to consult with experts in data security and privacy protection as well as a health care practitioner to ensure that internal practices conform to the intent of the law. Further, access to data and uses of data must be tracked and may be verified through an external audit by the Privacy Protection Commission. Academic partners can obtain access to aggregated results from the permanent database from IMA-AIM that have been screened to protect data confidentiality.

Table 2.1. **Belgium: Permanent sample of socially insured persons** *(cont.)*

Study title	Permanent sample of socially insured persons
Study results and future directions	The permanent database was developed to better inform policy decisions and it is directly used by policy makers, particularly in the area of managing health care expenditures. Work includes monitoring the costs of treating patients with chronic diseases, monitoring out-of-pocket payments for health insurance, monitoring the use of health technologies, and monitoring the recurrence of diagnostic exams and biological tests, such as blood tests. Also examined has been exposure of patients to radiation from medical imaging. Results influence the development of clinical guidelines. Overall the database helps the government to explore new ways of financing health care and to explore policy scenarios. The permanent database is focused on insurance transactions and therefore does not have information on the results of lab tests nor diagnosis. Diagnosis can, for some conditions, be inferred from prescription medicines. There are new proposals to link the permanent database to the hospital clinical minimum dataset to enable the study of readmission to hospital by the reason for the readmission. This request falls outside of the authorising legislation for the permanent database and will require the approval of the Privacy Protection Commission and a legal amendment is noted as required. A second proposed project is to understand the degree to which cancer patients may have a caregiver in their household. This will also require a linkage that is outside of the legislation and a separate approval process. There is a new research requirement to examine the histories of health care use of sampled persons for a period of up to 30 years. This change to the permanent database would require a legal amendment. Overall the governance of the permanent database with the establishment of the management committee has increased the ability of project partners to generate evidence for the management of the health system by reducing the heavy burden of documentation and the time lag required that would occur if each study were a separate application to the Privacy Protection Commission. Through the management committee all stakeholders are around the table and decisions to undertake projects can be made efficiently. In future, the permanent database can be used to develop health care quality indicators related to primary care and prescription medicines. The database could also be used to contribute new international quality indicators. While overall in Belgium it has become more difficult to undertake data linkage studies in health over the past five years, the development of the permanent sample has made it easier to analyse health insurance data than it would otherwise be. New concerns by the privacy commissioner of the risk of re-identification from small cells in aggregate tables will have to be addressed. A technical commission has been established to develop more documentation on managing disclosure risks from small cells.

Table 2.2. **Canada: Pathways of care for stroke patients**

Study title	Pathways of stroke care
Lead organisation	Canadian Institute for Health Information (CIHI).
Project description	This pilot project links data from in-patient hospitalisation data, emergency department visits and data on rehabilitative care in the Canadian province of Ontario in order to determine the additional information that could be gained on the outcomes of stroke care through data linkage (Canadian Institute for Health Information, 2012).
Project approval	The project was approved by CIHI senior management and the CIHI Committee on Privacy and Confidentiality, which includes the Chief Privacy Officer for CIHI. This committee grants approval for all projects involving the linkage of patient records across databases in the custody of CIHI. Elements of the application for approval include a description of the linkage project, the value of the project, restrictions to access to the linked data, the retention period for the linked data and protection of confidentiality of data in any published results. The approved retention period for the linked file prepared for this study was three years. The elapsed time between the submission of the application and the approval to conduct the study was three months.
Data and data linkage	Data for Ontario from several databases were linked at the level of the patient including inpatient hospitalisations, emergency department data, inpatient rehabilitation data, and complex continuing care data. Complex continuing care is a specialised programme providing continuing, medically complex and specialised services to patients over an extended period of time. It can be provided in either a free-standing facility or through designated beds within acute care hospitals for chronically ill patients. These patients require skilled, technology-based care that is not available through other long-term care facilities or home care programmes. All of the databases were in the custody of CIHI. There was no requirement for patient consent for the linkage of these administrative databases. Individual-level data is shared with CIHI from Canadian provinces to build national databases. All provinces have a health insurance number that is unique to the province and is used for health care encounters. Some jurisdictions encrypt the health insurance numbers on their databases before sending individual-level data to CIHI. These provinces apply the same encryption algorithm to each file and over time so that data linkage is possible. Other provinces provide CIHI with files with original health insurance numbers. For these provinces, one of the first data processing steps within CIHI is to encrypt the health insurance numbers. The standard encryption algorithm renders the original health insurance number unrecognizable. For this project the analytical team was provided with the data needed for the project and the team conducted a deterministic linkage using encrypted health care number, date of birth and sex. The quality of the linkage was high with over 90% of records successfully linked across the databases involved. Provincially issued HINS make it very difficult to trace patients who move province and introduce bias into record linkage studies. For the pathways of stroke care this bias was considered to be minimised because the period of follow-up was only four years.

Table 2.2. **Canada: Pathways of care for stroke patients** *(cont.)*

Study title	Pathways of stroke care
Protection of data privacy	Analytical teams never have access to data with original health insurance numbers. Access to this identifiable data is restricted to a data processing unit. Access to the analytical files necessary for the pathways of stroke care project was limited to a small number of named individuals approved to complete the work. The data is stored on a secured server. As the study has not yet been published, no external researchers have requested access to the linked database prepared for this project.
Study results and future directions	The publication of the results is expected to have implications for the care of stroke patients in Ontario and particularly for the organisation and co-ordination of their care. The analytical team would like to repeat the project in the future and to extend the linkage to other aspects of the care of stroke patients including long-term care data; home care data; pharmaceutical medicines data; and primary care data. Due to gaps in the coverage of these databases they were not included in the first pilot project. As the data holdings of CIHI expand and improve over time, greater insight into pathways of care would be possible.

Table 2.3. **Canada, Ontario: Institute for Clinical and Evaluative Sciences**

Study title	Institute for Clinical and Evaluative Sciences
Lead organisation	Institute for Clinical and Evaluative Sciences, University of Toronto.
Project description	The Institute for Clinical and Evaluative Sciences is a university-based research centre providing population-based health services research for Canada's largest province, Ontario (Institute for Clinical and Evaluative Sciences, 2011). ICES reports on many topics including health system performance; drug safety and effectiveness; population health; diagnostic services; and primary care. The research programme of ICES depends on the linkage of individual-level data from a variety of sources. ICES is sponsored by the Ontario Ministry of Health and Long-term Care and also receives academic grants for research projects.
Project approval	Under Ontario's Personal Health Information Protection Act, ICES is identified as a prescribed entity. This status enables ICES to receive and to use personal health information without patient consent for the purposes of analysis and statistics about Ontario's health care system. ICES must ensure that it can demonstrate that the collection and use of the information is in the public good and furthering medical research. Projects are approved by ICES science leaders. A form is then completed describing the project, the databases involved, the project team, and the benefit to the public of the project. This form is signed off by the lead investigator, the Chief Privacy Officer, the satellite site director if the study will take place at a satellite site and the ICES CEO. Many projects have benefited from scientific grants and have also fulfilled the requirements of the granting agencies involved.
Data and data linkage	ICES receives personal health data from the Ontario Ministry of Health and Long-term Care and also negotiates to receive data transfers from other Ontario prescribed entities, such as disease registries. Data sharing agreements are used to describe the terms of the data transfer including data privacy protection and data security. ICES data holdings of personal health data from the Health Ministry go back to 1988 and ICES is authorised to hold this data until such time as ICES is closed or the ministry ends its agreement with ICES. The legislation does not place limits on the retention of personal health data and the ministry understands that a long time series is necessary for epidemiological research. For data transfers to ICES from other Ontario data custodians, the data sharing agreement will specify a date of data destruction. The Ontario Health Insurance Number is used for all patient encounters for public health care services, including physician claims, drug benefits, and hospital encounters. The Ontario Registered Persons Database includes all HINS associated with individuals by name, address and birth date. ICES receives fully identifiable personal health data from various data custodians, including administrative health data from the Ontario Ministry of Health and Long-term Care. As a first use, ICES encrypts the health information numbers to create an ICES key number (IKN) and further de-identifies the data by removing other direct identifying variables. ICES uses the same encryption algorithm for all files and throughout time to enable data linkage. Files received from the Health Ministry do not contain the original Health Insurance Number, names or addresses. Data linkages at ICES depend on the ICES key number and tend to be undertaken using a deterministic method.

Table 2.3. **Canada, Ontario: Institute for Clinical and Evaluative Sciences** *(cont.)*

Study title	Institute for Clinical and Evaluative Sciences
Protection of data privacy	To fulfil the requirements of prescribed entity status within PHIPA, ICES policies and practices for data use and data security are reviewed by the Ontario Information and Privacy Commission on a tri-annual basis. ICES is required to have a data privacy and security framework complying with a manual prepared by the Privacy Commission. A threat and risk assessment of ICES information security takes place annually by an independent organisation hired to audit ICES data security and to try to penetrate it (ethical hacking). All data inflows and outflows to ICES use an encrypted transfer system. The retention period for linked databases is set on a project-by-project basis when a project is approved. ICES has a network of academic researchers working throughout the province of Ontario and only researchers affiliated with ICES or collaborating with an ICES researcher may be approved for access to ICES databases. Access is provided at the ICES headquarters or through a network of secure satellite centres providing access to ICES databases over a secure network. Researchers are always required to access data holdings within a secure facility. Approval is limited to the databases and the variables within the databases that the researcher needs to undertake a project. To make it easier for researchers to plan a project, ICES provides information about the databases and the variables within them on its Intranet site. There is a separation of duties at ICES where researchers analysing data can only see de-identified data and only specially designated individuals may encrypt health insurance numbers and de-identify data. There is no electronic link between computers used to process identifiable data and computers used for analysis. As a prescribed entity, ICES is able to process personal health data from administrative sources without consent. ICES informs the public about its research programme through its website. ICES also publishes an information brochure. When ICES is involved in primary data collection, such as a recent survey of homeless persons, participant consent to data linkage is sought. ICES has various documents to help staff inform staff about the protection of data confidentiality and privacy, including policies and a manual provided to new staff members. Staff are reminded of requirements and updated on new guidelines or policies through staff meetings, the Intranet site and an employee newsletter dedicated to data privacy and security. Employees are required to sign an agreement annually to protect data confidentiality.
Study results and future directions	ICES researchers have published thousands of research studies using data linkages at a population level. This research has informed about the effects of treatments in real-world settings which can differ from results of clinical trials. For example, a recent study of patients who had received an implantable cardio defibrillator determined that after six months there were important variations across care settings in the occurrence of inappropriate shocks and deaths (Krishnakumar et al., 2011). The evidence produced by ICES contributes to policy planning and evaluation within the Ontario Ministry of Health and Long-term Care. While access to databases of ICES have been restricted to staff, adjunct staff and sponsored collaborators, there has recently been an exception created by ICES and another entity, Cancer Care Ontario. Cancer Care Ontario requested a linkage of its cancer database to ICES databases for a project to be undertaken by researchers at Cancer Care Ontario. To provide data to these external researchers, ICES engaged in a more sophisticated process of data de-identification, including the removal of all direct identifying variables, the conversion of dates to indicators of elapsed time, and the inclusion of less-specific geographic identifiers. The de-identification process used was developed by Dr. Khaled El Emam and is called the Privacy Analytics Risk Assessment Tool (PARAT). The resulting de-identified data was provided to the external researcher on a CD. ICES is interested in exploring the option to use the PARAT process for other similar projects in the future. When researchers are interested in a project involving other Canadian provinces, the process has been to develop code for data processing together and then to have a researcher affiliated with each province responsible for the linkage of their own data and the analysis of that data. There is interest within ICES in further developing mechanisms to improve the ability to undertake analysis among Canadian provinces. There is also interest in further exploring linkages of health data to other areas, such as education or transportation.

Table 2.4. **Europe: Comparing diabetes outcomes across European countries**

Study title	European Best Information through Region Outcomes in Diabetes (EUBIROD)
Lead organisation	The Department of Internal Medicine at the University of Perugia (Italy) is coordinator of the project which involves 20 partners and two collaborating institutions from at least 20 countries, including EU member states and other countries.
Project description	The EUBIROD project is amalgamating aggregate data from 18 disease registries across Europe to create a sustainable European Diabetes Register to monitor diabetes indicators including complications and health outcomes. The three-year project has been sponsored by the European Union (EUBIROD, 2011).
Project approval	It is the responsibility of regions to obtain approval to participate.
Data and data linkage	Project participants use the BIRO system to submit aggregate data from their disease registry using an on-line data transfer system. Each disease registry receives a statistical programme and technical support from EUBIROD which helps to ensure consistency in the submitted indicators. The BIRO system produces pre-defined tables with pre-defined views. It is not possible to generate new table views.

Table 2.4. **Europe: Comparing diabetes outcomes across European countries** *(cont.)*

Study title	European Best Information through Region Outcomes in Diabetes (EUBIROD)
Protection of data privacy	Some diabetes registries have a strong infrastructure for data security and protection of data privacy while others have grown up from a local advocate and have achieved governmental support with a weaker infrastructure. It was important for the EUBIROD project to understand the the management of data privacy protection among the different registries within the countries participating , as different approaches could affect the completeness of the information within the registries and the comparability of results.To build this knowledge, EUBIROD developed an on-line tool for privacy performance assessment that enables participating registers to evaluate their level of respect for the privacy principles enshrined through the EU Data Protection Directive (Di Iorio et al., 2010). Key findings of the privacy performance assessment included that privacy principles have been implemented heterogeneously across Europe, with some interpretations restrictive to the point of limiting development and use of the registries that would be consistent with research in the public interest. The EUBIROD project has sought common ground where local requirements for data protection would be respected to ensure the greatest number of countries could particpate. The project did not attempt to amalgamate individual-level data. Instead, the data submitted to the project through BIRO are aggregate indicators that would not reveal the identity of a patient. Further, as it is not possible to generate new tables, the risk of revealing the identity of individuals from repeated generation of detailed tables is mitigated. In the first published report, anticipated for fall 2011, regions contributing to EURBIROD will be identified by a study number and not by country. This is because some countries have only one region participating and identification of the country could then reveal the identity of a particular diabetes care centre.
Study results and future directions	The first EUBIROD report will cover approximately 120 000 subjects and yield 72 indicators. Regions contributing to EUBIROD will be identified by study number. In the future, as more health care centres contribute data it may be possible to publish indicators by country. The European Commission is also interested in the possibility of using the BIRO system to populate indicators required by the European Union. This is not possible in the short run, but may become possible in the future as more care centres participate and data becomes more representative of countries.

Table 2.5. **Finland: Monitoring performance, effectiveness and costs of treatment episodes**

Study title	PERFECT – PERformance, Effectiveness and Cost of Treatment Episodes (Finland)
Lead organisation	THL National Institute for Health and Welfare, Finland.
Project description	PERFECT monitors the content, quality and cost-effectiveness of treatment episodes in specialised medical care and thus contributes to monitoring health system performance. Indicators and models were created to monitor selected disease groups and procedures (stroke, premature newborns, hip fracture, breast cancer, schizophrenia, acute myocardial infarction, and orthopaedic endoprosthesis including hip and knee replacement surgery, and invasive heart surgery). These disease groups and procedures were selected because of the number of patients treated and/or the level of treatment costs. Through the linkage of individual-level data, the project is able to go beyond reporting on single health care events to examining the whole cycle of care including patient outcomes, treatments and use of health system resources for well defined, and risk-adjusted, patient groups (Häkkinen, 2011). The performance measurement data generated from the PERFECT project enables benchmarking clinical practices against best-practice guidelines and assessment of the degree to which guidelines are being followed. The data can also be used to investigate the policy-sensitive factors explaining differences between hospitals and regions The project creates a database of about 200 indicators for hospitals and regions that are identified by name.
Project approval	The project was initiated by a consortium of researchers and clinical experts who received a scientific grant for health services research. On initiation, there was no requirement for research ethics approval of this project as it involved only registry data. Two years ago, however, legislative requirements changed and a research ethics board was created within THL (National Institute for Health and Welfare) to act as the approval body for data linkage projects involving registry data. If a project requires data from Statistics Finland, such as mortality and cancer statistics, an approval process takes place within that organisation as well. As necessary, Statistics Finland will also consult with the national privacy office. Because of the on-going development of PERFECT, the team presents an application for data linkage approval at nearly every monthly board meeting. The requirements for board approval places an administrative burden on the PERFECT team and project plans must take into account the additional time required to prepare and follow-up applications.
Data and data linkage	The project depends on high-quality and linkable individual-level data from within a set of databases including hospital in-patient records, out-patient records, birth records, disease-specific registers, prescribed medicines data, social care data, death records, and data on care reimbursement (Peltola et al., 2011). The data linkages required for the project take place at THL (National Institute for Health and Welfare). The main linkage key is the Personal Identity Number which is used in all data collections for public services including health care. Overall the quality of the data linkages is high for Finland, particularly when the research team compares linkage results to those of other countries attempting similar work.

Table 2.5. **Finland: Monitoring performance, effectiveness and costs of treatment episodes** *(cont.)*

Study title	PERFECT – PERformance, Effectiveness and Cost of Treatment Episodes (Finland)
Protection of data privacy	THL has a right to collect and use personal health data under law and patient consent is not required for the use of registry data and mortality data, including data linkages. If a survey is to be included in a PERFECT study, however, survey respondents need to have been asked for consent to link their survey responses to the registries for the linkage to be approved. An agreement was negotiated between Statistics Finland, the social insurance authority and THL for the data to be shared to undertake the PERFECT study. The agreement ensures that the requirements for data security and data privacy of each data supplying organisation will be respected. The PERFECT team ensures that the data custodians receive copies of published reports so that they can monitor how their data have been used. Only one individual within THL holds the code used to encrypt the Personal Identity Number (PIN) and the data with direct identifiers is stored in a locked room. The PIN is converted to an encrypted identifier using an algorithm that is consistently applied across all of the databases required for linkage. Any names on the files are converted to an encrypted code that cannot be reversed. The original names and PIN are then removed from the databases that are provided to THL staff that will perform the record linkage with the encrypted identifiers. The linked files will contain address information and dates which have some risk of re-identification of individuals, however, all of the data are stored within the secure THL facility and are only accessed by staff who have been approved for this access and who have signed an undertaking to protect the confidentiality of the data. Computers are password protected and other computer security protections are implemented. Data can be provided by the PERFECT team to researchers outside of the THL. In this case, however, the content of the file would be reduced to the minimum needed for the researcher to undertake their project. The external researchers would be required to sign an agreement with the THL which commits them to protect data confidentiality and security and which binds them to not share the data with third parties.
Study results and future directions	Across the disease groups and procedures examined, a wide range of patient outcome indicators have been published including mortality rates at different time intervals, such as at seven days, 30 days, 90 days or one year after an event; emergency room visits and rehospitalisation rates overall and for specific reasons such as infections and surgical complications; and days spent at home and/or in long-term care facilities after events (Peltola et al., 2011). The study found that centralisation of care of new-borns in the five university hospitals would reduce infant mortality rates. This lead to a change in the law requiring new-borns to be treated in university hospitals. Further, as one of the university hospitals had a higher infant mortality rate than the other four, it was subject to a quality audit. Results published include that, after risk-adjustment for differences in the characteristics of patients across regions, as much as 20-30% of the cost of treatment of AMI patients could be contained if all regions in Finland could match the costs of the best performing region (Häkkinen et al., 2011). Further, better outcomes could be obtained for AMI and stroke patients, including a reduction in deaths. Hospitals have used the PERFECT indicators to initiate quality improvement. Currently PERFECT is focussing on monitoring the quality of hospital care but there are plans to extend this work to examine primary care, elder care and social services. Municipalities have agreed to participate in the extension and Statistics Finland has presented the proposal to the privacy office for Finland for approval. The PERFECT project is also having an international impact through the launch of EuroHOPE, which involves a set of countries developing quality monitoring indicators for hospital care using the PERFECT methodology.

Table 2.6. **Germany: Effectiveness and safety of breast cancer screening**

Study title	Evaluation of early detection guidelines for mammography screening and examination of the quality of breast-cancer mammography as a diagnostic tool
Lead organisation	Institute of Clinical Epidemiology/Institute of Cancer Epidemiology, University of Luebeck, Germany.
Project description	Two key projects evaluate breast cancer screening effectiveness and safety. A project to evaluate early detection guidelines for mammography screening requires a large population cohort to compare with health outcomes of participants in a clinical trial. Mammography screening involving spectrum mammography was delivered to women who took part in a clinical trial in Germany. This screening process exposes women to radiation and thus it is important to evaluate the benefits and harms of population-level screening. Particularly, of interest are the rates of interval cancers, which are cancers that were not detected from screening and still occurred between screening intervals. This study will examine the health outcomes of screened women and compare them with outcomes of a representative cohort of women who did not participate in the trial. A second project to examine the quality of breast-cancer mammography as a diagnostic tool focussed on a cohort of women who had experienced breast pain or a suspicious lump, who had sought medical care, and where a physician had requested a mammography screen as a diagnostic tool (Obi et al., 2011).

Table 2.6. **Germany: Effectiveness and safety of breast cancer screening** *(cont.)*

Study title	Evaluation of early detection guidelines for mammography screening and examination of the quality of breast-cancer mammography as a diagnostic tool
Project approval	The early detection guideline is part of federal legislation for statutory health insurance. The legislation states that an evaluation of the guidelines would take place and that results would be shared with government. While the women who took part in the clinical trial provided informed consent, the comparison cohort of non-trial participants would be drawn from administrative data. The guideline authorised that the evaluation would require the use of administrative records from a large population cohort and therefore without patient consent. In Germany, legal data protection requirements are established partly on the Federal level and partly on the state (Land) level. Data protection supervision for federal public sector entities and social security administration is the remit of the (Federal) Data Protection Commissioner. Each of the 16 states in Germany has a Data Protection Commissioner with jurisdiction at the state level and their responsibilities include social security administration at the state level. Data linkages in Germany only take place at the state level and only when authorised by law. Because of the necessity to examine a population cohort for the country from administrative data without consent; the study must be incorporated into law within each of the 16 states. Thus far, a few states have implemented the changes in law required for the project. For the study of the quality of mammography as a diagnostic tool, the cohort to be studied was all women who had consented to participate. As a result, the study protocol was approved by the research ethics boards within the German states as it satisfied the requirements of existing legislation.
Data and data linkage	Both projects required data from clinical trials to be linked to records for the same patients within the cancer registries of the participating German states. There is no patient number in Germany to facilitate data linkages. Linkages take place using a set of identifiers including name, date of birth, location, and, if available, place of birth. Approved linkages take place within participating German states, and data files that have been anonymised are provided to the researcher. There is a national pseudonymisation algorithm used by all German states and this helps to limit the bias that could occur as a result of cross state mobility of patients. The algorithm is tolerant of small spelling errors in names. Name changes can occur, however, particularly for women. Where possible, the birth name of women is incorporated into the linkage framework. Linkages depend on probabilistic techniques and are resource intensive requiring skilled technicians and significant time. Other identifiers may remain on a linked file for analysis, such as date and place of birth, if they have been justified as necessary for the analytical project. Evaluations of the quality of the linkages indicate results are of reasonably high quality.
Protection of data privacy	The national data protection law and accompanying federal guidelines have strict provisions for data security. The Institute for Cancer Epidemiology in the state of Schlewig-Holstein has strong physical security including doors that cannot be opened from the outside without a key; and password-protected computers. Lap-top computers are encrypted to protect against risk in the event of a lost or stolen machine. Individuals are only authorised access to data they need for their work and only staff involved in data processing ever has access to patient identifiers. Staff involved in processing of data is under a clean desk policy where they must have all files locked away at the end of each day. Every staff member is issued a handbook on data protection. Cancer registry data may be held for 110 years after the birth of an individual, which enables long-term studies of cancer survival to take place. Data with encrypted identifiers (de-identified data) may be held indefinitely. The retention of linked data files is negotiated with the data protection authorities for each project and the scientific standard which is recommended is to retain the files for ten years. If the study has the informed consent of participants, the participants will be informed that the data will be linked for ten years.
Study results and future directions	The project to evaluate early detection guidelines for mammography screening is an example of the value to policy of research requiring data linkages, as the evaluation study was incorporated into legislation. The probabilistic linkages required for these studies are costly. It may be possible to make a case in the future to encrypt a health information number, as is the process currently in Germany for the pseudonymisation of names, but with the result being more successful and less resource intensive linkages. It remains unclear whether or not the changes that would be required to enable deterministic linkages in Germany would be supported by the population.

Table 2.7. **Korea: Quality assessment of medical services**

Study title	Quality Assessment of National Health Insurance Benefits, Korea
Lead organisation	Health Insurance Review and Assessment Service (HIRA).
Project description	This project aims to assess the clinical appropriateness and cost effectiveness of health care and to improve health care quality by reporting on quality and inducing service providers to make improvements to any services determined to be inadequate (Health Insurance Review and Assessment Service, 2011). It aims to identify underuse, overuse and misuse of therapies and to reduce variation in care practices through the regular reporting of quality indicators. Indicators from the linkage of hospital in-patient data to mortality data include 30-day case fatality for acute myocardial infarction; in-hospital fatalities within seven days after discharge, and within 30 days after surgery for coronary artery bypass grafting; 30-day in-hospital operative mortality for colorectal cancer; and 30-day operative mortality for stomach surgery, oesophageal surgery, pancreatic surgery, stem-cell transplantation, hip replacement, and percutaneous coronary intervention. Korea monitors mental-health care through a linkage of mental hospital in-patient data to prescription medicines data to produce the prescription rate of atypical anti-psychotics for schizophrenia. Through a linkage of mental hospital in-patient data to hospital in-patient data, Korea calculates the rate of readmission within 30 days of discharge from hospital for schizophrenia. Through the linkage of primary care data to prescription medicines data, outcomes of prescribing patterns including overlapping or inappropriate prescribing are monitored in Korea. Indicators include the rate of prescriptions of four-or-more component anti-hypertensive medications, parallel administration of diuretics, prescription of not-recommended parallel therapies, prescription days, and continued prescription for hypertension. For diabetes patients there is a set of monitored indicators including the rate of overlapping prescriptions; the rate of prescription of four-or-more component anti-diabetics, medication cost per administration day; and days of continued prescription. Korea also links the cancer registry data to mortality data to assess the relative survival of cancer patients and links long-term care data to survey data on the activities of daily living to estimate the percentage of patients with reduced activities of daily living.
Project approval	The Quality Assessment of National Health Insurance Benefits is conducted by HIRA under the requirements of the National Health Insurance Act. Data linkage projects are approved by a deliberation committee of HIRA on a project-by-project basis and are evaluated according to the requirements of the Privacy Protection Act.
Data and data linkage	Data linkages are conducted by HIRA and depend on the Resident Registration Number which is issued at birth and is used throughout the health care system. HIRA receives data from other government sectors, such as death data from the Ministry of Public Administration and Security.
Protection of data privacy	HIRA has internal guidelines on the protection of data security and confidentiality including specific guidelines related to data linkage. HIRA's data security and protection of data privacy are subject to audit by both the Ministry of Public Administration and Security and the National Intelligence Service. Data analysis takes place only within a designated area and by designated employees within HIRA. HIRA employees are trained in data security and privacy protection on a regular basis. External non-profit academic researchers or researchers in public-good organisations may apply to the deliberation committee for approval to access de-identified personal health data. When approved, analysis must take place only within a designated area within HIRA. Requests for access to data to prepare educational materials or to contribute to a thesis would not be approved.
Study results and future directions	Through data linkages, HIRA is able to report on the quality of services provided by physicians, clinics, hospitals and long-term care providers. These statistics are used to report on the quality of services for particular patient groups, including diabetic, heart and cancer patients. HIRA also reports on expenditures by disease categories and can therefore examine efficiency.

Table 2.8. **Sweden: Quality and efficiency assessments of clinical guidelines**

Study title	Open comparison and assessment (of clinical guidelines)
Lead organisation	National Board of Health and Welfare, Sweden.
Project description	In Sweden, there are quality and efficiency assessments of areas of care that are subject to national guidelines including cardiac and stroke care, care for four types of cancer, dental care, diabetes care, and mental health care. Generally there are clinical care guidelines developed for diseases involving large groups of patients. Clinical care guidelines are reviewed every 4-5 years. A panel of experts is convened to review and prioritise new treatments that could be incorporated into clinical care guidelines. This committee also reviews and prioritises health care quality indicators that could be developed to measure the quality and efficiency of care. These indicators require data linkages. Indicators for cardiac and stroke care were the first to be undertaken and results have been published (Socialstyrelsen, 2010). Assessments are now underway for psychiatric care, diabetes care and dementia care. The assessment takes place three-to-four years after the introduction of the clinical care guideline. For the first assessment, records for cardiac care patients were linked to health care encounters to assess how processes of care may have changed as a result of the introduction of the guidelines and to also examine results in terms of patient health. For example, the study looked at patient survival after a heart attack and the history of medications prescribed to the patients after the acute event. If, for example, a medication is found to have generated complications, its use may be given a lower priority when the care guidelines are revised.
Project approval	The government has given the National Board of Health and Welfare the mandate to assess and compare areas of care and to develop the indicators needed to assess the clinical care guidelines. As the National Board of Health and Welfare is authorised to collect and to process personal health data, no further research ethics approval was required for the project.

Table 2.8. **Sweden: Quality and efficiency assessments of clinical guidelines** *(cont.)*

Study title	Open comparison and assessment (of clinical guidelines)
Data and data linkage	The first published study involved the linkage of data for cardiac care patients in the National Quality Register to the Patient Registry. The National Quality Register is focused on the treatment of cardiac conditions. This linkage of the quality register to the patient registry enabled examination of all of the health care encounters of cardiac patients. This is particularly important when examining the effectiveness of medications, as they may have unintended health consequences. For example, a patient may be given blood-thinning medicine in response to a cardiac condition and then be admitted to hospital for gastric bleeding. The hospitalisation is not a cardiac event, but may be a complication of the medication given because of a cardiac condition. It is only through data linkage that a more complete picture of patient outcomes of care emerges. In Sweden, every person has a Personal Identity Number that is used for social security and health care and the use of the identifier is mandatory. Linkages using the PIN are of high quality. There are some data coverage issues, particularly for recent immigrants who do not yet have a permanent residence status and thus may be given a different temporary number for each health care encounter. Further, some quality registers have high patient coverage (85-90% of patients) while others have lower coverage (as low as 60%). Low coverage can bias study results. In Sweden, there are 21 county councils that govern health care. The government reaches out to these councils to implement health care guidelines and the councils work with their hospitals to participate. Participation in quality registers is voluntary and some councils work hard to gain the participation of hospitals including offering financial incentives. Others do not place the same priority on gaining high hospital participation. To participate in a quality register, a hospital would need to dedicate some staff time to data reporting.
Protection of data privacy	Patient participation in registries is mandatory. However, if patients wish to have their data removed from a registry they may appeal to the Board. Information on the registries maintained by the board is posted for patients to read in patient wards and health care centres. Quality registers in Sweden have initiated a process of patient consent and ask consent to use personal health data for research or statistical purposes. There is a specific unit within the National Board of Health that is permitted access to identifiable data and who undertake data linkages. Linked files then have identifying numbers removed before they are provided to individuals within the board for analysis. Data analysts never see identifiable data. De-identification involves removing names, personal identity numbers, addresses and full birth dates. Analysis files may contain a study number which has been assigned in place of the personal identity number. Data that has been linked for the health care quality assessment project is retained for six months and then it is destroyed. Data is stored in a locked and secure building using computers that have been protected from unauthorised access. Only employees granted access to the data can use the data and all use is tracked by a security officer. Data confidentiality rules are applied to prevent residual disclosure of patients, particularly from hospitals with very few cases. New employees receive training on data confidentiality and security and on their legal responsibilities related to the data. Data used for these assessments is not available to researchers external to the National Board of Health and Welfare. External researchers would need to approach the quality registers for access to their data.
Study results and future directions	The government has requested this project to assess the impact of the clinical guidelines and the results of the assessment are taken seriously. The conduct of the assessment itself has lead to quicker adoption of the guidelines because results are reported for each hospital by name and the media and the general public have the opportunity to examine how hospitals are performing. Hospitals who don't participate in the quality register are also named. The whole process of undertaking the assessment has been beneficial for both introducing and effectively implementing new policies. The coverage of the registers is improving and they are better now than they were five years ago. Data linkages are expected to be used more often in the future and to expand to new areas, such as to education and to social care. This would enable examination of differences in access to care and in health care quality for different groups within the population.

Table 2.9. **Switzerland: Understanding the life expectancy of a nation**

Study title	Swiss National Cohort
Lead organisation	Consortium of university researchers and the Federal Statistical Office, Switzerland.
Project description	The Swiss National Cohort (SNC) is a long-term, census-based, multipurpose cohort and research platform. It is based on the linkage of individual data from the 1990 census to the 2000 census and then the linkage of this data to mortality records from 1991 up to 2008 (Spoerri et al., 2010). It permits a better understanding of the socio-economic and socio-demographic characteristics of mortality and life expectancy and forms a baseline cohort from which additional data may be linked for specific research studies, such as to the cancer registry, the childhood cancer registry and survey data.
Project approval	The SNC began as a pilot project with the Federal Statistical Office (FSO) to evaluate if data linkage would be feasible. Given feasibility was determined; the university research team successfully obtained a grant to undertake the project from the Swiss National Science Foundation. The university consortium then entered into a contract with the FSO. The next step was to obtain research ethics board approval in each of the Swiss Cantons. The Federal Data Protection and Information Commissioner evaluated the project and provided a letter indicating that the project plan reflected good practices for data privacy and confidentiality protection. Once the baseline cohort was established, there have been requests to link other data files to the baseline cohort for specific studies. For each new linkage proposal, the FSO grants approval and, if approved, a module contract is drafted. . Whenever there is a linkage of data to the baseline cohort for an approved project, the data is deleted when the project is completed.

Table 2.9. **Switzerland: Understanding the life expectancy of a nation** *(cont.)*

Study title	Swiss National Cohort
Data and data linkage	The core of the SNC is census and mortality data, both of which are in the custody of the Swiss FSO. The FSO provides the SNC team with de-identified data. This de-identification includes the removal of names and addresses. Other identifiers remain on the file including dates of birth, nationality, marital status, sex and municipality. There is no national identifying number that could be used to conduct linkages in Switzerland and the researchers use probabilistic techniques to link the data using these less direct identifiers. Overall the quality of the linkage is quite good with more than 90% of deaths linked to the Census. For younger age groups, such as those aged 20 to 40, however, data quality is less good due to their increased likelihood to live alone, to marry, to separate and to move residence. Data quality is also higher where the date of death is closer to the date of the administration of the census. Quality problems can bias study results and the research team prefers to be responsible for the probabilistic linkages necessary to build the cohort.
Protection of data privacy	The Swiss census is mandatory and there is no requirement to obtain consent for the linkage of this data to mortality data. Additional linkages to the regional cancer registry data are conducted with patient consent. The protection of data security is part of the negotiated agreement between the SNC team and the FSO. Elements of the agreement include that the data cannot be shared with a third party, that the data cannot be linked to another file without approval and that any research results destined for publication will be first reviewed by the FSO. Researchers who will have access to the data files are identified by name. The data is stored on a secure computer. Tru-crypt software is used to encrypt data on any mobile devices, such as lap-top computers, in the event of theft or loss. SNC researchers only have access to the databases and variables they need to conduct their projects. When the SNC project was first established, the FSO inspected the SNC facilities. Researchers external to the SNC team can request access to the data and, if their proposal is accepted, they must sign a contract with the SNC team. The contract limits them to the databases and variables within the databases they need for their project, describes how the data must be securely stored and protected.
Study results and future directions	The study provides a population denominator for the cancer registry to enable calculation of the incidence and prevalence of cancer in Switzerland. The cohort also enables the study of cancer types, causes of death and life expectancy by socio-economic variables and other population characteristics available from the census. The cohort is on-going and a next step will be to link it to the new registry-based population census for Switzerland that has replaced the questionnaire-based census. The registry-based census will provide more current information but will not have the same details on the population's socio-economic characteristics as did the questionnaire-based census. The SNC cohort team will need to explore linkages to survey data. There is future potential to extend the cohort to health insurance data to enable the study of health care quality and outcomes. The research team would like to have access to higher quality patient identifiers for more successful linkages, such as encrypted names. Whether access to higher quality patient identifiers would be possible, and indeed, whether access to the same identifiers currently used by the SNC team will continue, depends on the evolution of legislation. There is a new law under development to create a national cancer registry from the existing ten regional registries. This law has the potential to influence privacy protection requirements that may have implications for the future work of the SNC cohort team.

Table 2.10. **United Kingdom, England: NHS Information Centre for Health and Social Care**

Study title	NHS Information Centre for Health and Social Care
Lead organisation	National Health Service, United Kingdom.
Project description	The NHS Information Centre for Health and Social Care collected, processed, linked, analysed and published national information for health and social care communities in England (National Health Service Information Centre for Health and Social Care, 2011). Under the Health and Social Care Act of 2012, by April 2013, it will become an executive non-departmental public body entitled the Health and Social Care Information Centre (HSCIC) and it will have broader responsibilities, including assuming data collection responsibilities previously held by other arms-length bodies (Department of Health, 2012). It will become a single national repository for data for secondary purposes, including holding and linking person-identifiable data where approved and necessary. At the time of this OECD study in October 2011, the Information Centre was both producing tabulations from hospitalisations data (HES) and conducting limited data linkages on behalf of clients. Clients included government departments, academic researchers and commercial interests.
Project approval	At the time of this OECD study, for data linkages, all requests must have had the approval of the Secretary of State for Health, who made these decisions on the advice of the UK National Information and Governance Board (NIGB) Ethics and Confidentiality Committee before the Information Centre could accept to undertake the linkage for the client. A key requirement of NIGB was that the project was in the public interest. The Information Centre itself had been approved by NIGB to conduct data linkages without patient consent when the linkage was among health care administrative databases. This NIGB approval was renewed annually and the Information Centre was required to continue to justify retaining the databases in its custody. All clients of the Information Centre entered into a written data release agreement with the Centre.

Table 2.10. **United Kingdom, England: NHS Information Centre for Health and Social Care** *(cont.)*

Study title	NHS Information Centre for Health and Social Care
Data and data linkage	At the time of this OECD study, the Information Centre was the custodian of the hospital commissioning dataset and mental health in-patient data and its data linkage service focused on linkages involving Hospital Episode Statistics (HES). The Information Centre received data from the central registry of the NHS, mortality data from the Office of National Statistics and data from the Cancer Registry. The Information Centre linked databases using the NHS number. Where unavailable, other identifiers were used to assign an NHS number to a file before linkage using both deterministic and probabilistic techniques. After data linkage, the NHS number was removed from the linked file and replaced with a study number that had no other meaning. Also removed from the file were full dates of birth, postal codes and any local patient identifying numbers. The execution of the data linkage by the Information Centre could be quite efficient. For example, for a project where all data files had NHS numbers and an established linkage algorithm was then able be used, the client may have received their results in less than one week. Generally the quality of linkages to hospitalisation data (HES) was quite good with 90% or more of clinical trial cohort participant records successfully linked. Projects where NHS numbers were not as available had been less successful.
Protection of data privacy	The collection of patient-level hospital records occurred without consent. Patients had the right, however, to refuse to have their data used for public health research. This was a rare request and, when it did occur, the Information Centre took steps to ensure the patient's records were suppressed. Patients were informed of the uses of their data thorough governmental websites and also, depending on the project, through posters in health care facilities. Researchers who had a cohort of data they had collected, such as from a clinical trial, may have requested to the Centre to link their cohort to follow the patients for subsequent hospitalisations, cancers and deaths. If the cohort was collected with the informed consent of participants, the Information Centre could approve the project and proceed. If the cohort was not collected with the consent of participants, the project must have first been approved by NIGB. Researchers requiring approval of NIGB should have planned for a six-month delay between the first submission of an application and a decision. NIGB met bi-monthly and it was typical that a researcher received questions and must revise and resubmit their application. External researchers providing a cohort of data to be linked would be able to re-identify cohort members, even though the NHS numbers are removed from the linked files. Data transfers to and from the Information Centre used a secure web transfer system. The Information Centre had a security policy which was required for its annual approval by NIGB and the security of the Centre was reviewed by security experts in the Department of Health. The computer system was protected by a firewall. Employees of the Centre only had access to data files required for their job and computers were password protected. Employee data access permissions were reviewed annually. Access logs were kept and were monitored to ensure that employees were still using the files they had approved access to use. Employees had on-line training in data protection annually that included a test that must be passed.
Study results and future directions	The data linkage service provided by the Information Centre was relatively new at the time of the OECD study and it had not yet been widely publicised. Centre staff expected that requests for data linkage services would grow in the future in response to greater awareness.

Table 2.11. **United Kingdom, England and Wales: Birth outcomes studies**

Study title	Linkage of birth data to delivery records from hospital data (1) and estimation of gestation-specific infant mortality statistics (2)
Lead organisation	1) Government of Wales. 2) Office for National Statistics, England and Wales.
Project description	Two data linkage projects are underway in the United Kingdom to improve understanding of infant health. 1) The linkage of birth data to delivery records from hospital data was undertaken by the Welsh Government to reconcile inconsistencies between multiple sources of birth data for Wales, each of which had content missing from the other. The objective is to arrive at a better, more comprehensive and more consistent picture of maternity outcomes (Welsh Government, 2010). 2) In 2002, a new system was implemented to allocate National Health Service (NHS) numbers at birth in England and Wales. The linkage of NHS birth notifications with birth and death registration records enables statistics on births and infant deaths by population characteristics, such as gestational age and ethnicity (Moser et al., 2008).
Project approval	1) The study was approved by the Caldecott Guardian who is responsible for access to and use of all data in the custody of the NHS Wales Informatics Service (NWIS). 2) The study was approved by the National Information and Governance Board which is a national research ethics board that renders decision on projects requiring the linkage of personal health data where patient consent to linkage has not been obtained. Application materials must correspond to the eight principles of data privacy protection within the UK Data Protection Act and must include a justification for each variable within each dataset that would be required for the study; evidence of the benefits to policy of the study; a literature review that demonstrates the project's contribution to knowledge; and a description of how data security will be protected. It took about 40 days to receive a decision from NIGB once the application for approval had been submitted.

Table 2.11. **United Kingdom, England and Wales: Birth outcomes studies** *(cont.)*

Study title	Linkage of birth data to delivery records from hospital data (1) and estimation of gestation-specific infant mortality statistics (2)
Data and data linkage	1) NWIS helps to administer the health system and is the custodian of health care administrative data. NWIS is responsible for undertaking data linkages related to health care for the government of Wales. Linkage requests submitted to NWIS require at least a six month lead time before the linkage needs to take place. For this project, birth records from the Child Health System that registers births were linked to birth and delivery records within hospital data. The NWIS first pseudonymised the NHS number using a similar algorithm for both databases to be linked. The NWIS then performed the linkage using the encrypted number. In cases where a deterministic match on the encrypted number was unsuccessful, the NWIS used other identifiers, such as the mother's date of birth to clarify matches. While the quality of the linkage was high, with about 98% of infants linked to their delivery record, some issues with the underlying data were discovered as a result of the data linkage. These included that the health numbers recorded for infants from a multiple birth can become switched between the hospital record and the birth registration. 2) This study involves the linkage of birth notifications in the custody of the NHS to birth registrations and death records in the custody of the Office for National Statistics (ONS). The ONS undertook the data linkage. The linkage was deterministic and depended on the NHS number. For records that could not be linked deterministically, a probabilistic matching was undertaken using the date of birth of the mother and the infant and the postal code. The NHS data was of good quality and over 99% of live births were successfully linked. The undertaking of the linkage, however, did uncover a few data quality issues. For example, some local areas had incomplete data and some had incompletely recorded ethnicity. The ONS was able to report these findings back to the NHS which helped to further improve the quality of the NHS data.
Protection of data privacy	1) In Wales, a child health record book is given to new parents at the birth of a child or at the first physician visit in the ten days following a birth. The book contains a paper that explains to parents that the birth record will be held in a national database. There is no option to opt-in or to opt-out. The NWIS conducted the data linkage on behalf of the Wales Office for National Statistics (ONS). There is a secure web-based transfer of data between NWIS and the ONS. Within the ONS, the data are then stored on a protected network. Only a few approved persons within the ONS have access to the linked data for analysis purposes, and the analysis file has been de-identified. The ONS follows both the Welsh Government standards for information protection and the ONS code of practice. The ONS also has ISO 27001 status which confirms that it conforms to this international standard for data security. Staff members obtain training and subsequent refresher training on data security and confidentiality. Access to the data has never been requested by researchers external to the ONS. 2) As NIGB has recognised in the approval of the project that patient consent is not practical for the more than 700 000 live births occurring annually in England and Wales, it has accepted that the ONS informs patients of the use of their data. ONS produces a poster that is put up in birth and neo-natal units within health care facilities and distributes a leaflet to new parents. The poster and leaflet describe the ONS and its mandate, the data that will be collected and how it will be used and present some recent findings from the data. The ONS removed the NHS number from the linked file provided to ONS staff for analysis. The analysis file retained, however, the date of birth and the postal code as both were required for the study. The postal code was used to include in the analysis variables representing area deprivation from another data source. Access to the analysis file was restricted to a small number of approved ONS staff members and their names were provided to NIGB. The analysis file is stored on a server that is separate from the server used to perform the data linkage and analysts never have access to the identifiers used to perform the linkage. Staff members are provided with training and with regular reminders about data security and confidentiality. There are internal security audits.
Study results and future directions	1) The new maternity service in Wales has recognised that there is insufficient information on maternity outcomes. Rather than launching new data collection, data linkage was piloted as an alternative approach to reconcile databases of births to arrive at a better overall picture of maternity outcomes. Whether or not to continue this data linkage on a regular basis is still to be determined. 2) In the past, the main indicator of poor birth outcomes was infant mortality. Today, infant mortality rates are low and there is interest in better understanding outcomes for low birth-weight babies. The data linkage provides the only database that enables monitoring of low birth-weight babies in relation to their gestational age and enables reporting of birth outcomes by the age of the mother, by ethnicity, and by location to enable better targeting of programmes to support healthy infants. The study began as a pilot and now continues as an annual project.

Table 2.12. **United Kingdom, England: Mapping pathways across health and social care – Torbay Care Trust**

Study title	Mapping pathways across health and social care
Lead organisation	Torbay Care Trust, United Kingdom (England).
Project description	Torbay Care Trust is an integrated health and adult social care organisation responsible for providing and commissioning services for the population of Torbay, England which is about 140 000 people. The trust engaged the firm MedeAnalytics to develop a data management system that links patient records across health and social care databases to produce pathways of services and their associated costs on an on-going basis (Health Service Journal, 2011). This information is then used to evaluate and improve services which cross jurisdictional boundaries. Included in the information system are patient level records from acute care hospitalisations and all contacts of adult patients with community-based services (ancillary care, home care and nursing homes), All social care services that are fully or partially reimbursed by the Trust are included within the information system.
Project approval	The Trust executive approved the linkage of data for this project.
Data and data linkage	The databases included in the project are largely within the custody of the Torbay Health Trust which is either directly providing the services or is commissioning the services. The acute care hospitals in Torbay are the only other data custodians contributing records to the database. Linkage is possible because these hospitalisation records contain the NHS numbers for patients. The patient records shared with the Torbay Care Trust are the same records also shared with the NHS, as part of hospitals' mandatory data collection. Data linkage is deterministic using the NHS number. Torbay is unique in requiring an NHS number for all patients who receive social care services. When the data system was first developed, Torbay had to invest in cleaning up its databases to ensure complete and correct NHS numbers were available. Older records had patient identifying numbers that were unique to the service provider's IT system and did not have an NHS number. Torbay has now over 99% of records for social care with an NHS number. One of the means of achieving such a high proportion has been efforts to ensure that GPs include the patient's NHS number on referrals to social care services. A unique identifying number is essential to the project and the analytical value of the database was well worth the costs associated with assigning NHS numbers to older records. The database also contains dates of birth and information on the locality of the patients and their GPs.
Protection of data privacy	The analysis of the patient records is within the mandate of the Trust for the administration of its programs. Patients are made aware of the uses of their data as part of the face to face assessment that occurs when patients first enrol in social care services. The project uses a web-based system to upload and download data that is encrypted and password-protected. Torbay Care Trust uploads patient records to the system that are accessed by MedeAnalytics who prepare the analysis ready database which is then used by the Trust for analysis over a secure network. Within the Trust, only a small number of named individuals have access to the linked database. When MedeAnalytics was selected to assist the Trust to develop its data infrastructure, the data security protections of MedeAnalytics were checked. Trust staff receives training in IT governance and protection of data confidentiality. Data security protection is also explicitly part of employees' job descriptions. Staff members are only able to receive a computer account necessary to access the database if they have been recommended by the Deputy Minister. An information governance team within the Trust ensures that data security and privacy protections remain strong. The Trust has not experienced a breach of data security. If a breach occurred, the matter would be the responsibility of the Information Governance Officer within the Trust. As the database is fairly new, there has only been one request from an external researcher to analyse the data. The executive team of the trust evaluated the proposal and the project was approved. The data analysis for the project was conducted by the Torbay care staff.
Study results and future directions	The linkage of health and social care data has given a new perspective on the costs of disease management. The linked data has shown that when both health and social care expenditures are considered together, the costs of caring for patients experiencing events such as a hip fracture or stroke, are much higher and require a much longer care period than had previously been appreciated. This has raised the importance of prevention efforts to identify and support high-risk patients to avoid acute events. The data has also informed the Trust about the full picture of costs associated with caring for patients and has provided the empirical evidence needed to avoid shunting costs from one area of care to another, particularly when budgets are being reduced. Even though Torbay is a small community, analysis of its linked data has had an impact on public policy at a national level. The national government had discussed moving funds from social care services to acute care services to commit acute care providers to being responsible for their patients in the 30 days after their discharge. While this initiative was in response to a need to combat readmissions, analysis of the linked data from Torbay showed the potential negative impact on social care services that could result. The national roll-out of the policy was halted in favour of testing the policy in a few pilot locations. The Trust now has three years of linked data and, is planning to benefit from this growing resource to identify health services with better patient outcomes and that are more efficient and to introduce measures of health care quality. Future directions for the Trust include incorporating data on primary care services from General Practitioner's offices into the database which will improve understanding of how GP services form part of health care pathways. A challenge will be to engage GP's participation. Over the past five years it has become easier for the Trust to link and analyse personal health records. The availability and use of NHS numbers has been a key factor. The services of MedeAnalytics have made it easy for the Trust to harness the power of their patient records to extract information for decision making. The first year of the project involved a lot of testing, and as a result of confidence in the quality of the data, it is relied upon for policy and commissioning decisions. A concern for the future is that recent NHS emphasis on the use of pseudonymisation when data is first collected could have the consequence of making data linkages impossible as the sharing of data with common unique identifiers is a necessary prerequisite for linkage. Another challenge is the lack of standardisation of service definitions and coding across health care Trusts in the United Kingdom. Thus it is very difficult to compare social care services across Trusts. The only benchmarks for Torbay are in the acute-care sector where comparability is possible.

Table 2.13. **United States: Understanding health care users and health outcomes**

Study title	Linkages of population health surveys to death and medical care records
Lead organisation	National Center for Health Statistics, United States.
Project description	The National Center for Health Statistics has a data linkage group who is building a repository of surveys that are ready to support linkage projects through harmonising content and standardising variables across different surveys and survey waves (Centers for Disease Control and Prevention, 2012). Two key projects have been undertaken. 1) The linkage of survey data to the National Death Index (NDI) and 2) the linkage of survey data to data on health care for Medicare and Medicaid recipients from the Centre for Medicare and Medicaid Services (CMS).
Project approval	The NCHS is authorised by law to collect and process a broad range of statistics on illness and disability of the population of the United States. Data linkages are approved by the Research Ethics Board of the NCHS. The linkage of survey data to CMS data was first proposed by the CMS and the Social Security Administration (SSA) in response to a Congressional request in 2000. This project required the development of the first interagency agreement of its kind and involving four federal entities, the NCHS, the CMS, the Social Security Administration and the Department of Health and Human Services. The Agreement makes reference to the legal authorities under which each of the entities is able to share data to undertake the project. Each of the two projects is on-going and there is no set date where the linked data would be destroyed. Project (1) is updated every 3-4 years as part of the data programme of the NCHS. The Interagency agreement enabling Project (2) is updated and signed annually by all of the agencies involved.
Data and data linkage	The NCHS is the custodian of the survey data and the death index data involved in data linkages. The CMS is the custodian of Medicare and Medicaid enrolment and claims data. The Social Security Administration is the custodian of the SSA benefit history data. Linkages of population health survey data to the National Death Index are conducted probabilistically using seven matching criteria including a combination of Social Security Number; date of birth; first, middle or last name; and/or father's last name. The linkage of population health surveys to CMS data are conducted in three steps using a deterministic linkage method. First the NCHS prepares a file containing SSN, names and dates of birth from the population health surveys to be linked. This file is sent to the Social Security Administration (SSA) who links the file to their own databases to correct any errors in the SSN for the survey respondents. The NCHS file that has corrected SSNs is then sent by the SSA to the CMS. The CMS conducts a deterministic linkage of the NCHS file to Medicare and Medicaid records based on SSN, Medicare identification numbers and dates of birth. Once the health care records have been linked to the NCHS file, the SSN and Medicare identification numbers are removed from the file by the CMS, leaving only an NCHS assigned identity number and the file is returned to the NCHS for subsequent analysis.
Protection of data privacy	Consent requirements have changed over time. All potential survey respondents are mailed a letter about the survey that explains that survey data may be linked to other health records. During the interview, respondents are asked to provide their Social Security Number and their Medicare identification number, if they receive Medicare benefits. More recently, the NCHS has asked respondents to provide the last four digits of their identification numbers. If they do not provide the numbers, they are asked if their data may be linked to other health data without the identification numbers. The change in the consent process has resulted in an improvement in the proportion of survey records eligible for linkage, from 45% of the records for the 2006 National Health Interview Survey to 86% of the records for the 2009 administration of this survey. Death data may be linked to other data holdings without the prior consent of the decedents or their surviving family. Researchers outside of government may submit a research proposal to access linked data files within the NCHS Research Data Centers (RDC) or at the NCHS headquarters. The RDC are secure facilities for data access that are maintained by the NCHS throughout the United States. Research proposals are evaluated by NCHS staff unless the proposal involves genetic data. In that case, the proposal is reviewed by the NCHS research ethics board. Researchers are only approved to access data files and variables that are required for their project and names, SSN and Medicare identification numbers are not provided. Exact dates are provided only if their inclusion is necessary for the project. The linkage of survey data to death data is provided to external researchers through a public-use micro data file. This file has been de-identified, including perturbations of the data to avoid indirect identification. No public-use micro data file is available for the linkage of survey data to Medicare and Medicaid records. The health care administrative data is more sensitive as it pertains to living subjects. NCHS staff signs yearly affidavits attesting to their agreement to protect data confidentiality. Staff members also receive yearly training on data security.
Study results and future directions	The linkage of the NCHS survey data to health care administrative data (Medicare and Medicaid) is being repeated. The process is proceeding more quickly this second time as the Interagency agreement needs only to be updated. The survey file linked to the CMS records is a complex file to analyse as the histories are inclusive of different types of health care from pharmaceutical use to physicians and to hospitals and the observations are affected by changes in eligibility for CMS programmes over time. Despite these challenges there are a significant number of analytical projects underway with the data currently and growing interest in using the data to examine the effectiveness of health care (Looker et al., 2011; Gorina and Kramarow., 2011). The linkage of survey and death data has been used to investigate policy-relevant topics such as the characteristics of individuals who have committed suicide and socio-economic disparities in life expectancy (Denney, 2010; Dray-Spira et al., 2010). The NCHS is working on ways to make linked data easier to analyse including on-line tutorials for complex issues, such as the necessity to re-weight data to reduce the impact of non-response and linkage biases on study results.

Table 2.14. **United States: Kaiser Permanente Center for Health Research**

Study title	Center for Health Research
Lead organisation	Kaiser Permanente.
Project description	Kaiser Permanente is a closed-panel health care maintenance organisation (HMO) with eight sites and 8.5 to 9 million members in the US states of Hawaii, Oregon, California, Ohio, Washington, DC, and Maryland. As a closed panel HMO, Kaiser members tend to see Kaiser doctors and Kaiser doctors tend only to treat Kaiser patients. Within this HMO, an electronic medical record system has been implemented. Kaiser has seven research centres conducting public-domain research analysing patient-level data (Kaiser Permanente Center for Health Research, 2011). The research group is semi-autonomous from the medical group that provides direct patient care. The research arm is a non-profit entity.
Project approval	Kaiser uses an internal review board (IRB) to evaluate proposals for access to patient-level data. Each Kaiser site is identified under the US Health Information and Protection of Privacy Act (HIPPA) as a covered entity and must abide by HIPPA requirements. As a result, each site has a separate IRB and has developed its own process for application submission and approval. When a research project is proposed that would benefit from data from more than one site, a separate application must be submitted to the IRB of each site involved. It has taken effort to coach physician researchers to submit proposals to the IRB and now that processes are established, many research projects do involve multiple sites.
Data and data linkage	Kaiser has an extensive and complete array of health care data for Kaiser patients at the level of person-encounters and use of services. This includes service dates and products, including pharmaceuticals. What can be lost are health care encounters and purchases outside of the system of the health care organisation. Kaiser implements initiatives that increase adherence and minimise data loss. For example, it operates a robotic pharmacy for prescription re-fills that mails re-fill to claimants. Kaiser members are offered a financial incentive to use the robotic pharmacy. In this way, it is possible to measure adherence to prescription medicines. Kaiser patients receive a unique record number that they keep for life. This record number is used to link patients across databases. As a private company, there is no public access to data on Kaiser patients. However, Kaiser researchers may enter into collaborations with external researchers in other organisations or universities. For example, there is a current project to estimate medical care costs at the end of life for cancer patients with researchers at Henry Ford, Seattle and Group Health. The project will compare costs for all cancer cases from 2000 to 2008 in the cities of Detroit, Seattle and Denver. A virtual database was established to provide all of the researchers involved with access to patient-level data. Kaiser has been bio-banking biological samples and tissues from since 1964 and it is the oldest of the stored tissues that now have the greatest value for research into the genetic factors that have resulted in disease in patients today. For example, it may take 20 years or more of exposure to tobacco smoke, dust, and air pollution to develop disease and it is only with a comprehensive analysis that the relative influence of different potential risk factors be estimated.
Protection of data privacy	Personal health data is not available for research until the patient has provided informed consent. The Kaiser membership agreement explains to patients, or family members of patients, that when the agreement is signed they are providing consent to research with their medical records and with any biological samples or tissues. Any member who declines to consent is tracked and their data is suppressed from research studies. In Oregon, state law requires patient consent for the collection of genetic information for research. The reach of the law is broad as even gender or race may be considered genetic information. To ensure compliance with this new law, Kaiser administered a document to all members asking if they would decline to have their data used for research. About 10% of members did decline. There is a separation of duties at Kaiser Permanente. There is a separate directory space on the Kaiser computer network where employees who conduct data linkages work. The unique record number used to link databases is removed from linked files provided to researchers for analysis. There are audit trails that track who within the organisation have access to personal health information and to spot intruders and to prevent attacks on its data security. This monitoring takes place on a 24/7 basis. Facilities were data are stored are secure and can only be accessed by authorised employees. Guests must be escorted at all times. There are hierarchies of access to data at Kaiser with individuals only approved to access data required for their job. Servers and files are protected from unauthorised access. Identifiable data may not be removed from a secure facility. Researchers may have access to de-identified data for research, however, the researcher must log into a secure server and all keystrokes are encrypted. All employees sign an affidavit to protect data confidentiality annually and there is on-line training annually for employees on the protection of data security and confidentiality.
Study results and future directions	Kaiser has accumulated 50 years of expertise in conducting research with patient-level data and linkages. Kaiser, with its EMR system and bio-bank, is on the forefront of research with linked data with new research areas including genetics, genomics and personalised medicine. Bio banking is expensive and complex, requiring a system to barcode samples and storage in temperature-controlled environments; however, it represents one of the most promising avenues for new research. For example, a new study at Kaiser involving the linkage of stored samples and health care records is investigating if certain prescription medicines for mental illness are linked to producing genetic mutations in people. Other new opportunities on the horizon include the use of devices to measure risk factors, such as accelerometers to measure physical activity and devices to measure the environment around individuals, such as air quality and noise.

Summaries of further examples of policy-relevant data linkage projects

Country	Study title	Project description
Australia	Care pathways for older patients with chronic diseases	The Australian Institute of Health and Welfare explored care transitions for older people with dementia, cardiovascular disease, arthritis and musculoskeletal conditions. Records for all people aged 65 and older who were assessed under the Aged Care Assessment Program were linked to data for six major aged care programs and to mortality data over a four-year period. The linked data enabled examination of care pathways and the factors influencing different care paths. Of particular interest were the entry into and the time to entry into residential care and how this may have been influenced by the use of community care (AIHW, 2011).
Australia	Study of the health effects of exposure to low-dose radiation from CT scans in childhood	This study, led by researchers at the University of Melbourne, with data linkage undertaken by the Australian Institute of Health and Welfare, explored whether exposure to low-dose computerised tomography scans in childhood increases the risk of cancer. Records of all children aged 0 to 19 who received medical services between 1985 and 2005 in the Medical Benefits database were linked to the Australian cancer registry and death databases. The cancer incidence of those exposed to CT scans was compared to that of the non-exposed of similar age and sex. Results will be compared with those of parallel studies in the United Kingdom and the United States to inform international guidelines for CT scan use in childhood.
Belgium	Studies of health care quality and outcomes for cancer patients	The Belgian Cancer Registry has several projects underway where data linkages are helping to generate new information about the quality of care and health outcomes of cancer patients. These include a linkage of breast cancer screening results and the cancer registry to measure the quality of screening through the identification of the occurrence of interval cancers. A second project involves the continuous linkage of cancer registry data with nomenclature (medical procedures and pharmaceutical data) and vital statistics. This analysis will produce indicators of care quality including cancer survival and variability in treatment practices. The registry has also elaborated indicators of the quality of oncology care for breast and testicular cancer that could be implemented within cancer centres (Stordeur et al., 2011; Vlayen et al., 2011).
Belgium	Disease registries for cystic fibrosis and neuromuscular diseases	Two projects are underway within the Scientific Institute of Public Health to develop disease registries for cystic fibrosis and neuromuscular diseases. To create the registries, patients with these two conditions who were treated in expert centres would be followed up for subsequent health outcomes.
Denmark	Monitoring of cancer pathways	This project of the National Board of Health reports on wait times in cancer treatment pathways.
Finland	Drug and pregnancy project	Through the linkage of a number of registers this project evaluates patterns of medication use during pregnancy to estimate the effect of drug use on pregnancy outcomes. Results enable monitoring of the safety of medications used during pregnancy (Artama et al., 2011).
France	Monitoring health care use and expenditures	France has developed a national insurance information system including a permanent sample of beneficiaries (SNIR-AM) to create a national picture of health care consumption and expenditures. Longitudinal data for patients are available for the current year and the previous two years. The database is used to study certain chronic diseases where there is a 100% reimbursement rate and certain prescribed medicines (Tuppin et al., 2010). This linked data has been used to study implantation of pacemakers and cardioverter-defibrillators; hospitalisation rates for low-income subjects with full health insurance coverage; and the use of the medication Benfluorex among Diabetic patients and the occurrence of valvular heart disease.
Germany	Quality of health care of patients after hospital discharge for myocardial infarct	A project is being planned to assess the quality of care of patients discharged from hospital for a myocardial infarct through the linkage of claims data from AOK Berlin-Brandenburg and clinical data from the Berliner Herzingfarktregister.
Israel	Quality of health care projects	The Ministry of Health in Israel is undertaking data linkage projects to monitor quality of health care in order to determine health policy in specific areas. These projects include examination of post-operative clinical complications, re-hospitalisations and mortality. The projects are on-going in several areas: colon surgeries, craniotomies, fractures of neck of femur and appendix (Simchen et al., 2011). Another study is exploring mortality among psychiatric patients in order to improve community mental health care (Haklai et al., 2011).
Korea	Annual cancer statistics	The National Cancer Centre established a National Cancer Incidence Database incorporating: the Korea Central Cancer Registry, a medical review survey, eleven population-based regional cancer registries, and site specific cancer registries. The centre reports nationwide cancer statistics, including incidence, mortality and survival rates, and their trends (Jung et al., 2010).
Norway	Social inequalities in health	The Norwegian Institute of Public Health is working to describe trends in social inequality in Norway from 1960 to the present through the linkage of mortality and population records. The project has described socio-economic inequalities in mortality for children and adults by cause of death and also socio-economic inequality in life expectancy (Næss et al., 2007)
Singapore	National Chronic Disease Management Programme	The national programme evaluates the quality of primary care providers by examining health care providers' adherence to recommended care processes, as well as their success in preventing hospitalisations related to those diseases. The programme encourages patients with chronic conditions to work closely with their doctors to avoid acute exacerbations or complications that could lead to hospitalisations while encouraging doctors to follow evidence-based disease management protocols (Ministry of Health, 2011).
Sweden	Quality and Efficiency in Swedish Health Care: Regional Comparisons 2010	This project uses data from approximately 30 health data and health care quality registers to generate statistics to openly compare processes and outcomes of health care (Socialstyrelsen, 2012).

Summaries of further examples of policy-relevant data linkage projects *(cont.)*

Country	Study title	Project description
United Kingdom	National Cancer Intelligence Network	This network is a UK-wide initiative to drive improvements in standards of cancer care and clinical outcomes by improving and using information collected about cancer patients for analysis, publication and research (Morris et al., 2010 and 2011; Lambert et al., 2011).
United Kingdom	Hospital Episode Statistics linkage to mortality	This project links mortality data from the UK Office of National Statistics to hospital episode statistics from the NHS Information Centre for Health and Social Care in order to add a unique anonymised patient identifier to the mortality database. This variable will help to perform analysis of hospital patients who have subsequently died (NHS Information Centre for Health and Social Care, 2011).
United Kingdom	Cancer survival	This project produces statistics on cancer survival for England and for the United Kingdom. The UK results are provided to the OECD for the Health at a Glance publication (Office for National Statistics, 2011).

Bibliography

AIHW – Australian Institute of Health and Welfare (2011), "Pathways in Aged Care: Do People Follow Recommendations?", *AIHW Bulletin,* No. 88, Cat. AUS 137, Canberra.

Altman, B.M. and S.P. Gulley (2009), "Convergence and Divergence: Differences in Disability Prevalence Estimates in the United States and Canada Based on Four Health Survey Instruments", *Social Science and Medicine*, Vol. 69, No. 4, pp. 543-552.

Artama, M., M. Gissler, H. Malm and A. Ritvanen (2011), "Nationwide Register-based Surveillance System on Drugs and Pregnancy in Finland 1996-2006", *Pharmacoepidemiology and Drug Safety*, Vol. 20, No. 7, pp. 729-738.

Canadian Institute for Health Information (2012), "Pathways of Care for People with Stroke in Ontario", Ottawa, *https://secure.cihi.ca/estore/productFamily.htm?locale=en&pf=PFC1695*, accessed 9 October 2012.

Centers for Disease Control and Prevention (2012), National Center for Health Statistics Data Linkage Activities, *www.cdc.gov/nchs/data_access/data_linkage_activities.htm*, accessed 10 October 2012.

Commission Technique d'Échantillon Permanent au Conseil Général de l'INAMI (2011), *Échantillon Permanente Steekproef 1/19, Annexe – CGSS 2011/097*, Belgium, *www.riziv.fgov.be/information/fr/sampling/pdf/report.pdf*, accessed 6 April 2012.

Denney, J.T. (2010), "Family and Household Formations and Suicide in the United States", *Journal of Marriage and Family*, Vol. 72, No. 1, pp. 202-213.

Department of Health and NHS Informatics (2012), *Informatics: The Future – An Organisational Summary*, *www.dh.gov.uk/health/2012/07/informatics-future/*, accessed 18 February 2013.

Di Iorio, C.T., R. Andány et al. (2010), "Privacy Impact Assessment Report", EUBIROD Consortium.

Di Iorio, C.T., F. Carinci, M. Brillante, J. Azzopardi, P. Beck et al. (2012), "Cross-border Flow of Health Information: Is 'Privacy by Design' Enough? Privacy Performance Assessment in EUBIROD", *European Journal of Public Health,* Epub May 4, IF: 2.313.

Di Iorio, C.T., F. Carinci, J. Azzopardi, V. Baglioni, P Beck et al. (2009), "Privacy Impact Assessment in the Design of Transnational Public Health Information System, the BIRO project", *Journal of Medical Ethics*, Vol. 35, No. 12, pp. 753-761.

Dray-Spira, R., T.L. Gary-Webb and F.L. Brancati (2010), "Educational Disparities in Mortality Among Adults with Diabetes in the U.S.", *Diabetes Care*, Vol. 33, No. 6, pp. 1200-1205.

DUQuE (2011), "Deepening Our Understanding of Quality Improvement in Europe", *www.duque.eu*, accessed 25 April 2012.

EUBIROD (2010), *WP5 Data Collection: Privacy Impact Assessment Report, www.eubirod.eu/documents/downloads/D5_2_Privacy_Impact_Assessment.pdf*.

EUBIROD (2011), *www.eubirod.eu/index.html*, accessed 25 October 2011.

EUROCARE (2011), *www.eurocare.it*, accessed 25 April 2012.

EuroHOPE (2012), *www.eurohope.info/*, accessed 10 October 2012.

EURO-PERISTAT (2011), *www.europeristat.com*, accessed 25 April 2012.

Gorina, Y. and E.A. Kramarow (2011), "Identifying Chronic Conditions in Medicare Claims Data: Evaluating the Chronic Condition Data Warehouse Algorithm", *Health Services Research*, Vol. 46, No. 5, pp. 1610-1627.

Gulley, S.P. and B.M. Altman (2008), "Disability in Two Health Care Systems: Access, Quality, Satisfaction and Physician Contacts Among Working-age Canadians and Americans with Disabilities", *Disability and Health Journal*, Vol. 1, No. 4, pp. 196-208.

Häkkinen, U. (2011), "The PERFECT Project: Measuring Performance of Health Care Episodes", *Annals of Medicine*, Vol. 43, Suppl. 1, pp. S1-S3.

Häkkinen, U., A. Malmivaara and R. Sund (2011), "PERFECT – Conclusions and Future Developments", *Annals of Medicine*, Vol. 43, Suppl. 1, pp. S54-S57.

Haklai, Z., N. Goldberger, N. Stein, I. Pugachova and I. Levav (2011), "The Mortality Risk Among Persons with Psychiatric Hospitalizations", *Israel Journal of Psychiatry and Related Sciences,* Vol. 48, No. 4, pp. 230-239.

Health Insurance Review and Assessment Service (2011), "Quality Assessment Service", Seoul, *www.hira.or.kr/eng/activity/01/04/activity01_04.html*, accessed 30 April 2012.

Health Service Journal (2011), "HSJ Best Practice Awards 2011", *www.hsj.co.uk/journals/2011/11/21/o/c/i/HSJBP2011.pdf*, accessed 6 April 2012.

IARC – International Agency for Research on Cancer (2011), *Cancer Incidence in Five Continents*, World Health Organization, monograph series, *www.ci5.iarc.fr*, accessed 25 April 2012.

Institute for Clinical and Evaluative Sciences (2011), Toronto, *www.ices.on.ca/index.html*, accessed 30 April 2012.

Jung, K.-W., S. Park, H.-J. Kong et al. (2010), "Cancer Statistics in Korea: Incidence, Mortality and Survival in 2006-2007", *Journal of Korean Medical Science*, Vol. 25, No. 8, pp. 1113-1121.

Kaiser Permanente Center for Health Research (2011), *www.kpchr.org/research/public/default.aspx* accessed 30 April 2012.

Krishnakumar, N., J.V. Tu and D.S. Lee (2011), "Why Are Implantable Cardioverter Defibrillator Outcomes in Practice Different from Clinical Trials?", *Cardiac Electrophysiology Clinics*, Vol. 3, No. 4, pp. 511-520.

Lambert, P.C., H. Moller et al. (2011), "Quantifying Differences in Breast Cancer Survival Between England and Norway", *Cancer Epidemiology*, Vol. 35, No. 6, pp. 526-533.

Looker, A.C. et al. (2011), "Hip Fracture and Risk in Older U.S. Adults by Treatment Eligibility Status Based on New National Osteoporosis Foundation Guidelines", *Osteoporosis International*, Vol. 22, No. 2, pp. 541-549.

Ministry of Health (2011), "Medisave for Chronic Disease Management Programme (CDMP) – The Third Year", Singapore, *www.moh.gov.sg/content/moh_web/home/Publications/information_papers/2010/medisave_for_cdmp-thethirdyear.html*, accessed 30 April 2012.

Morris, E.J. et al. (2010), "Comparison of Treatment and Outcome Information Between a Clinical Trial and the National Cancer Data Repository", *British Journal of Surgery*, doi: *10.1002/bjs.7295*.

Morris, E.J. et al. (2011), "Thirty-day Postoperative Mortality After Colorectal Cancer Surgery in England", *Gut*, doi: *10.1136/gut.2010.232181*.

Moser, K., K. Stanfield and D. Leon (2008), "Birthweight and Gestational Age by Ethnic Group, England and Wales 2005: Introducing New Data on Births", *Health Statistics Quarterly*, Vol. 39, pp. 22-31.

Næss, Ø. and A.H. Leyland (2010), "Analysing the Effect of Life Course Area of Residence in Multilevel Epidemiology", *Scandinavian Journal of Public Health*, Vol. 38, Suppl. 5, pp. 119-126.

Næss, Ø., B.H. Strand and G. Davey Smith (2007), "Childhood and Adulthood Socioeconomic Position Across 20 Causes of Death. A Prospective Cohort Study of 800 000 Norwegian Men and Women", *Journal of Epidemiology and Community Health*, Vol. 61, pp. 1004-1009.

National Health Service Information Centre for Health and Social Care (2011a), *The Health and Social Care Information Centre Annual Report and Accounts 2011/12*, *www.official-documents.gov.uk/document/hc1213/hc03/0332/0332.pdf* accessed 18 February 2013.

National Health Service Information Centre for Health and Social Care (2011b), "Linked ONS-HES Mortality Data", *www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1299*, accessed 30 April 2012.

Obi, N., A.Waldmann, V. Babaev and A. Katalinic (2011), "Record Linkage of a Large Clinical Practice Patient Cohort with the Cancer Registry Schleswig-Holstein", *Gesundheitswesen*, Vol. 73, No. 7, pp. 452-458.

Office for National Statistics (2011), "Index of Cancer Survival for Primary Care Trusts in England – Patients Diagnosed 1996-2009 and Followed Up to 2010", *Statistical Bulletin*, 13 December.

Peltola, M., M. Juntunen, U. Häkkinen et al. (2011), "A Methodological Approach for Register-based Evaluation of Cost and Outcomes in Health Care", *Annals of Medicine*, Vol. 43, Suppl. 1.

Simchen, E., I. Weiss Salz, A. Ekka-Zohar, L. Freedman et al. (2011), *The National Project for Quality Measures in Israeli Hospitals, 1st Report: General Surgery,* Ministry of Health, Jerusalem.

Socialstyrelsen (2010), *Open Comparison and Assessment 2009 – Cardiac Care*, Stockholm, *www.socialstyrelsen.se/publikationer2010/2010-9-2*, accessed 30 April 2012.

Socialstyrelsen (2012), *Quality and Efficiency in Swedish Health Care 2010*, Stockholm, *www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/18336/2011-5-18.pdf*, accessed 30 April 2012.

Spoerri, A. et al. (2010), "The Swiss National Cohort: A Unique Database for National and International Researchers", *International Journal of Public Health* , Vol. 55, pp. 239-242.

Spoerri, A., M. Egger and E. Von Elm (2011), "Mortality from Road Traffic Accidents in Switzerland: Longitudinal and Spatial Analyses", *Accident Analysis and Prevention,* Vol. 43, No. 1, pp. 40-48.

Stordeur, S., F. Vrijens et al. (2011), *Quality Indicators in Oncology: Breast Cancer*, KCE Reports 150C, Brussels.

Tuppin, P., J. Drouin, M. Mazza et al. (2011), "Hospitalisation Admission Rates for Low-income Subjects with Full Health Insurance Coverage in France", *European Journal of Public Health*, Vol. 21, No. 5, pp. 560-566.

Tuppin, P., A. Neumann, E. Marijon et al. (2011), "Implantation and Patient Profiles for Pacemakers and Cardioverter Defibrillators in France (2008-2009)", *Archives of Cardiovascular Diseases*, Vol. 104, No. 5,pp. 332-342.

Tuppin, P., L. de Roquefeuil, A. Weill et al. (2010), "French National Health Insurance Information System and the Permanent Beneficiaries Sample", *Revue d'Épidémiologie et de Santé Publique,* Vol. 58, No. 4, pp. 286-290.

Van Eycken, R., K. Haustermans et al. (2000), "Evaluation of the Encryption Procedure and Record Linkage in the Belgian National Cancer Registry", *Archives of Public Health*, Vol. 58, pp. 281-294.

Van Klaveren, R.J., M. Oudkerk et al. (2009), "Management of Lung Nodules Detected by Volume CT Scanning", *New England Journal of Medicine*, Vol. 361, pp. 2221-2229.

Vlayen, J., F. Vrijens, K. Beirens et al. (2011), *Quality Indicators in Oncology: Testis Cancer*, KCE Reports, Vol. 149, Brussels.

Wahlbeck, K., J. Westman, M. Nordentoft et al. (2011), "Outcomes of Nordic mental health systems: Life Expectancy of Patients with Mental Disorders in Denmark, Finland and Sweden 1987-2006", *British Journal of Psychiatry*, Vol. 199, No. 6, pp. 453-458.

Welsh Government (2010), "Adding Value to the Analysis of Welsh Birth Data, 2008", *http://wales.gov.uk/docs/statistics/2010/101215addingvalueen.pdf*, accessed 30 April 2011.

Further reading

Canadian Institutes of Health Research (2002), *Secondary Use of Personal Information in Health Research: Case Studies*, Ottawa.

Chappel, A. (2011), "Multi-payer Claims Database (MPCD) for Comparative Effectiveness Research", Presented to the NCVHS Full Committee Meeting, 16 June 2011.

Coleman, M.P., D. Forman, H. Bryant et al. (2011), "Cancer Survival in Australia, Canada, Denmark, Norway, Sweden, and the United Kingdom, 1995-2007 (the International Cancer Benchmarking Partnership): An Analysis of Population-based Cancer Registry Data", *The Lancet*, Vol. 377, No. 9760, pp. 127-138.

Faeh, D., J. Braun, S. Tarnutzer and M. Bopp (2011), "Public Health Significance of Four Cardiovascular Risk Factors Assessed 25 Years Ago in a Low Prevalence Country", *European Journal of Cardiovascular Prevention & Rehabilitation*, doi: *10.1177/1741826711430282*.

Faeh, D., S. Tarnutzer, J. Braun, and M. Bopp (in press), "Individual and Combined Impact of Traditional Risk Factors on Cardiovascular Mortality in Switzerland", *European Journal of Cardiovascular Prevention & Rehabilitation*.

Gill, L.E. (1999), "Ox-Link: The Oxford Medical Record Linkage System", Record Linkage Techniques – 1997 Proceedings of an International Workshop and Exposition, National Academy Press, Washington.

Gissler, M., M. Artama, A. Ritvanen and K. Wahlbeck (2010), "Use of Psychotropic Drugs Before Pregnancy and the Risk for Induced Abortion: Population-based Register-data from Finland 1996-2006", *BMC Public Health*, Vol. 30, No. 10, p. 383.

Goldacre, M. (2002), "The Value of Linked Data for Policy Development, Strategic Planning, Clinical Practice and Public Health", Symposium on Health Data Linkage Proceedings, Australian Government Department of Health and Ageing, Sydney, pp. 95-98.

Karmel, R., P. Anderson and A. Peut (2009), "Pathways Through Aged Care: A First Look", *Bulletin* No. 73, AIHW Cat. No. AUS 116, Canberra.

Malm, H., M. Artama, M. Gissler and A. Ritvanen (2011), "Selective Serotonin Reuptake Inhibitors and Risk for Major Congenital Anomalies", *Obstetrics and Gynecology*, Vol. 118, No. 1, pp. 111-120.

Paulinski, D. (2011), "ASPE/CMS Multi-payer Claims Database", Presented to Academy Health ARM, 14 June 2011.

Piro, F.N., Ø. Næss and B. Claussen (2007), "Area Deprivation and its Association with Health in a Cross-sectional Study: Are the Results Biased by Recent Migration?", *International Journal for Equity in Health*, Vol. 6, No. 10.

Spycher, B.D., M. Feller, M. Zwahlen et al. (2011), "Childhood Cancer and Nuclear Power Plants in Switzerland: A Census-based Cohort Study", *International Journal of Epidemiology,* Vol. 40, No. 5, pp. 1247-1260.

Sund, R. (2008), "Methodological Perspectives for Register-based Health System Performance Assessment – Developing a Hip Fracture Monitoring System in Finland", Stakes, National Research and Development Centre for Welfare and Health, *Research Report* No. 174, Helsinki.

Weill, A., M. Paita, P. Tuppin et al. (2010), "Benfluorex and Valvular Heart Diseases: A Cohort Study of a Million People with Diabetes Mellitus", *Pharmacoepidemiology and Drug Safety,* Vol. 19, No. 12, pp. 1256-1262.

Westman, J., M. Gissler and K. Wahlbeck (2011), "Successful Deinstitutionalization of Mental Health Care: Increased Life Expectancy Among People with Mental Disorders in Finland", *European Journal of Public Health,* Vol. 22, No. 4, pp. 604-606, doi: *10.1093/eurpub/ckr068*, Epub 2011 Jun. 9.