

ABSTRACTS

Abstracts are listed in alphabetical order by presenting author and include information as provided.

CONTRIBUTED

Rates and patterns of participation in cardiac rehabilitation in Victoria

Presenter: Michael Ackland, Victorian Department of Human Services

Authors: Vijaya Sundararajan, Stephen Begg, Michael Ackland, Ric Marshall, Steve Bunker and Helen McBurney

Background: Cardiac rehabilitation improves functional status and quality of life after a myocardial infarction (AMI), coronary artery bypass graft (CABG) or percutaneous transluminal coronary angioplasty (PTCA). It is currently the recommendation of the National Heart Foundation that all such patients be referred for rehabilitation. Previous pilot data indicate that overall only 22% of patients participate in rehabilitation following a cardiac event, with rates varying according to diagnosis (AMI (14%), CABG (39%), PTCA (20%)). Because these were pilot data, based on only 20% of all participating rehabilitation centres in Victoria, a more extensive study was conducted.

Methods: 1998 cardiac rehabilitation data from all of the participating centres in Victoria (N=4474) were linked to a subset of the Victorian Admitted Episodes Dataset consisting of all cases of AMI/CABG/PTCA/angina/catheterisation based on their ICD-9 CM codes (N=30,092). Angina and catheterisation codes were included to ensure that all potential candidates for rehabilitation were included in the analysis. The linkage process matched 89% of the observations from the NHF cardiac rehabilitation dataset to the 1998 VAED subset. Linkage was also undertaken with mortality data.

Results: Of the 14,463 cases of AMI/CABG/PTCA, the median age was 68 (25-75, 59-75), with 31% being female. Rates of participation in rehab were 25% for AMI, 43% for CABG, and 26% for PTCA. Whereas 27% of men participated in rehab, only 19% of women did so. Age was also a factor with those under 70 having participation rates of 30% or more, compared to those 70 and over at 23%. Marital status, absence of comorbid disease, and in particular, the absence of congestive heart failure were also significantly related to participation in cardiac rehabilitation. Outcomes including readmission and mortality will also be discussed.

INVITED

The value of linked data for research into surveillance and adverse events

Presenter: Nick Andrews, Public Health Laboratory Service, Statistics Unit and Immunisation Division, United Kingdom

Author: Nick Andrews

When a vaccine has been licensed and introduced into an immunisation programme it is important to be able to determine how well the vaccine is working (its efficacy against disease) and also be able to detect and investigate possible adverse reactions to the vaccine. In order to do this the immunisation division at the Communicable Disease Surveillance Centre in the UK is involved in monitoring vaccine coverage, carrying out enhanced surveillance of vaccine preventable diseases and also investigating hypothesised adverse reactions to vaccines.

Enhanced disease surveillance involves merging data from different sources to ensure as many cases as possible of the disease are detected and also to enable the collection of information on vaccine status and other risk factors. This information, along with coverage data, can be used to calculate the field efficacy of the vaccine and also to monitor changes in the incidence and epidemiology of the disease post vaccination. Possible adverse events are usually flagged up through passive surveillance or during vaccine trials. They may also be known reactions to the disease itself and/or biologically plausible reactions. Formal investigation of these events involves the linkage of a data set of hospital admissions from the South East region of England with a data set of vaccine records. The relative incidence of the event in a specified risk period after vaccination compared to the background risk outside this period is calculated using the self-controlled caseseries method, which only requires information on cases. If an adverse event is found to be associated with the vaccine then the risk of the event per vaccine dose is calculated.

In this talk the risks and benefits of vaccination will be discussed and examples will be given of how we have used enhanced surveillance data and, in particular, linked data, for assessing vaccine efficacy and safety. Recent results following the introduction of the new meningitis C conjugate vaccine will be presented.

CONTRIBUTED**Statistical linkage keys: How effective are they?****Presenter:** John Bass, Western Australia**Authors:** John Bass and Carol Garfield

Existing measures of statistical linkage keys (SLKs) have usually focused on how well a key represents the source population and on the extent of duplication, i.e. multiple keys for one individual as well as multiple individuals sharing the same key. We really need to know whether the analysis of data linked by deterministic matching of SLKs leads to significantly different conclusions than would be obtained through analysis of “actual” linked data. A project using information from the Western Australian Health Data Linkage Project has been able to provide some answers to this question.

A data set has been constructed containing seven years of hospital and death records (1993–1999) of individuals older than 19 years from Western Australia (2,844,030 hospital unit records). HACC and SAAP SLKs were created for all of these records, and deterministic linkages based on these keys were performed to link records within the hospital data as well as to a copy of the WA death register to which the HACC and SAAP SLKs had been added. The data also contain a personal identifier (WA PID) created by the Data Linkage Unit, based on probabilistic linkage using full demographic data (full names, sex, date of birth, address, country of birth and indigenous status). This WA PID has been improved by linkage to other data sets such as the state electoral roll which provides historical information on name and address changes. Significant effort has also been put into validation of the links.

The primary aim of the study is to compare the results of typical analyses of linked data from the same set of hospital and death records linked by means of the HACC and SAAP SLKs as well as the WA PID. The effects of increasing the time period and of indigenous status have been examined. Two analyses are presented here – the total number of bed days per patient, and the relative risk of death. The results are significant, indicating that the linkage method should be taken into account when interpreting the results of analyses of linked health data.

CONTRIBUTED**Health service use questions addressable by the Queensland linked data set****Presenter:** Rohan Baxter, CSIRO**Authors:** Rohan Baxter, Chris Kelman, Richard Solon, Simon Hawkins, Graham Williams, Deanne Vickers, Lifang Gu and Hongxing He

The Queensland linked data set project contains patient-level linkage of Queensland hospital morbidity data, Medicare claims data and Pharmaceutical Benefits prescription data for the years 1995–1999. We describe the population covered by the avail-

able linked data, pointing out the “black holes” and limitations. We describe the original fields and derived fields available in the Queensland linked data set. Using the available meta-data for these fields, we propose cohort designs, along with suitable outcome measures and events for analysing the relationships between the hospital, Medicare and PBS services.

Comparisons of data breadth and quality are made with the long running WA linked data project (Holman et al) and the recent WA–Commonwealth hospital–MBS linked data project (Kelman et al). We show how the health services utilisation and expenditure questions to be answered by the present data set can be differentiated from the other data sets.

The current research plan, along with initial results, are given. Research questions addressed include:

- i. The relationship between use of sophisticated diagnostic technologies and major medical interventions.
- ii. Adverse drug reactions leading to hospital admission and their cause and management.
- iii. Service utilisation patterns for long-term, high-use elderly users of pharmaceuticals.
- iv. Characterisation of demographics and co-morbidities of the high-cost patients across hospital, Medicare and PBS services.

CONTRIBUTED**The effect of locational and social disadvantage on utilisation and outcomes of health care: Cardiovascular disease****Presenter:** Kate Brameld, Centre for Health Services Research, Department of Public Health, University of Western Australia**Authors:** Kate Brameld, D’Arcy Holman and Todd Owen

Introduction: This project uses the WA Linked Database to look at the effects of social and locational disadvantage and possession of private health insurance on access to and outcomes following hospital care across Western Australia at the level of the ABS Collector’s District (CD). CDs are the smallest areas for which census-derived Socio-Economic Indices For Areas (SEIFA indices) are available and comprise approximately 250 households. Misclassification error and resultant bias towards the null are much reduced using SEIFA indices of social disadvantage based on CD compared with postcode. The study will cover all major diagnostic categories of the AN-DRG system as well as specific chronic diseases and procedures based on national priorities. Some preliminary results from this study will be presented.

Method: Geocodes on the hospital morbidity records of the WA Linked Database allow indices of relative disadvantage (SEIFA) and accessibility and remoteness (ARIA) to be allo-

cated to individual records at the level of the ABS Collectors District (CD). The hospital morbidity database includes a variable indicating possession of private health insurance. Admission and procedure rates, length of stay, cumulative readmission risk and case fatality rates will be calculated for the time periods 1994–96 and 1997–99 with follow-up to 2004 according to the categories of social and locational disadvantage and private health cover. Risk adjustment will be made for age, sex, Aboriginality and comorbidity. The effects of locational and social disadvantage will be compared between 1994–96 and 1997–99.

Results: Preliminary data on cardiovascular disease indicate increasing admission rates and length of stay with increasing social disadvantage with the exception of the least disadvantaged non-Aboriginal group. There appears to have been little change in admission rates between the two time periods although a slight decrease in length of stay has been observed. However, the analysis is currently biased by the absence of geocodes on approximately 16% of records and the propensity for certain categories of records not to be geocoded, e.g. remote Aboriginal communities and institutions.

Discussion: The pattern of admission rates and outcomes following admission for cardiovascular disease in the WA population according to social and locational disadvantage and private health cover will be presented and compared with published national and international data. Methodological issues and possible solutions will be discussed.

CONTRIBUTED

The potential of linking environmental, socio-economic and health data for health policy development

Presenter: Elisabeth Bui, CSIRO Land & Water

Author: Elisabeth Bui

In order to take a true “systems” approach to epidemiology, data must be assembled to represent all components of the system that impacts on health outcomes. Socio-economic as well as environmental data are essential as they both include risk factors. Some nation-wide core environmental datasets have been generated during the recent National Land and Water Resources Audit; these cover water resources, soils, vegetation, climate, land use, sediment and nutrient budgets. The National Pollutant Inventory and State Environmental Protection Agencies are important sources of air quality data. These datasets take the form of chloropleth maps, raster data, spatially referenced point observations, and relational databases. Geographical Information Systems enable easy storage and visualization of spatial data. The linkage of health data with socio-economic and environmental data permits the study of spatial patterns of disease together with exposure patterns. It enhances the feasibility of testing of hypotheses relating chemical exposure and disease outcomes. This is all in the domain of spatial environmental epidemiology.

Spatial environmental epidemiology seeks to unravel disease incidence, individual susceptibility and spatially varying factors. There are two traditional approaches in epidemiology to studying exposure-disease relationships. One approach is individual-level studies of case-control or cohort design, typically using logistic or Poisson log-linear regression to relate disease incidence to individual-level co-variables and suspected environmental factors. This approach is unable to discover geographical correlation and trends that are associated with unreported spatially varying environmental risk factors. Moreover, individual-level spatially referenced data are often unavailable. There are also serious confidentiality issues to consider with individual-level data.

The other approach is aggregate-level or “small area” studies. Disease counts are aggregated over geographical units and compared to aggregated co-variables with lattice-based Gaussian or log-linear Poisson Markov random field models to accommodate local spatial correlation. The aggregate-level approach offers the advantages of requiring only case counts and exposure data over geographical areas (which are often available even when individual case and exposure data are not) and of revealing and exploiting spatial correlations in the data. However this approach has some serious shortcomings: the aggregation unit of the disease and other data (socio-economic and environmental) rarely matches; the size of the aggregation unit can influence the results; the shape or size of the aggregation unit can distort the true exposure-disease relationship. The (often small) contribution of environmental risk factors to disease is difficult to estimate.

Given relationships between health and environmental risk factors can be established, health policy could address change in exposure patterns through community education, control of disease vectors, and environmental legislation to reduce emissions.

WITHDRAWN

Barriers to implementing worthwhile data linkage ideas: Lessons from Queensland

Presenter: Magnolia Cardona, Public Health Services, Queensland Health

Author: Magnolia Cardona

The benefits of data linkage for monitoring, planning, evaluation and hypothesis generation are acknowledged and have been welcomed as the basis for three real life projects in Queensland. Case 1: PBS-MBS-Inpatient data linkage. Case 2: Better Medication Management System. Case 3: IRIS project. A summary of the objectives, process and outcomes will be presented to illustrate the difficulties in translating the project ideas into the expected outcomes. Issues including confidentiality concerns, legislative requirements, cross-border agreements, infrastructure/capacity, dataset incompatibility, staff turnover, diversity of interest among parties involved, underestimates of

funding requirements, and definitions of “public good” have caused lengthy delays in the implementation of these worthwhile initiatives. Common sense actions for overcoming these barriers will be proposed.

CONTRIBUTED

Parallel computing techniques for high-performance probabilistic record linkage

Presenter: Peter Christen, Department of Computer Science, FEIT, Australian National University

Authors: Peter Christen, Markus Hegland, Stephen Roberts, Ole Nielsen and Tim Churches

Historical collections of administrative and other health data nowadays contain many tens or even hundreds of millions of records, with new data being added at the rate of millions of records per annum. Although improvements in available computing power have to some extent mitigated against the effects of this accelerating growth in the size of the data sets to be linked, large-scale probabilistic record linkage is still a slow and resource-intensive process. There have been relatively few advances over the last decade in the way in which probabilistic record linkage is undertaken, particularly with respect to the tedious “clerical review” process which is still needed to make decisions about pairs of records whose linkage status is doubtful. Unlike computers, there has been no increase in the rate at which humans can undertake these clerical tasks. The ANU Data Mining Group is currently working in collaboration with Epidemiology and Surveillance Branch of the NSW Health Department on the development of improved techniques for probabilistic record linkage. Our main focus is on the development of techniques which make good use of modern high-performance parallel computers, such as the APAC National Facility (a Compaq supercomputer with 480 processors), as well as smaller clusters of commodity PCs or workstations, which can be used as virtual parallel computers with little additional software installations.

As a first step, we will be developing prototype software that allows more efficient and faster linkage of large data sets using the “classical” probabilistic framework first described in the late 1960s by Fellegi and Sunter and extended in the 1980s by Winkler, Rubin, Jaro and others. To our knowledge, no implementation of probabilistic data linkage software is available for parallel computers. This work will involve the development of new techniques to handle the parallelisation of data distribution, blocking techniques, data preprocessing and load balancing. Secondly, this project aims to improve existing linkage techniques by using algorithms from the fields of data mining and machine learning to achieve a better linkage quality and, most importantly, to reduce the resources needed for the time consuming and tedious manual clerical review process. The suitability of techniques such as high-dimensional predictive modelling will be explored.

The developed software will be published under an Open-Source software license. The tools we are using are Open-Source software as well, namely the *Python* programming language (for rapid, object-oriented development) and *OpenMP* and *MPI*, two libraries for parallel programming. We hope to have prototype software available early in the second half of 2002. Further details of the design considerations for this software will be discussed.

CONTRIBUTED

The use of probabilistic record linkage, public key cryptography and trusted third parties to improve the protection of personal privacy and confidentiality in disease registers and tissue banks

Presenter: Tim Churches, Epidemiology and Surveillance Branch, NSW Health Department

Author: Tim Churches

Disease registers (DRs) aim to collect information about all instances of a disease or condition in a defined population. Traditionally DRs have required that notifications of cases of the target diseases be fully identified with items such as name and date of birth so that multiple notifications relating to the same case can be identified and merged. However, growing concern over the privacy and confidentiality aspects of DRs is beginning to hinder their operation, particularly in Europe. An alternative method of operation is proposed which involves splitting the personal identifiers from the medical details at the source of notification, and separately encrypting each part using asymmetrical public key cryptography. The identifying information is sent to a single population register (PR), and the medical details to the relevant DR. The shared PR does not need to capture identifying details of every person in the population, only those of people notified to a DR. The PR uses probabilistic record linkage to assign a unique personal identification (UPI) number for each person notified to it. This UPI is shared only with a single trusted third party whose only function is to translate between the UPI and separate series of personal identification numbers which are specific to each DR. The proposed scheme, which extends an algorithm described by Blobel et al. in 1995 for use in German cancer registries, would also allow linkage of records between DRs with minimal extra effort, under the supervision of the trusted third party. The scheme is directly extensible for use with tissue banks and other repositories of genetic material as well as disease registers and other health status or health event data collections. With the exception of a probabilistic record linkage engine, all of the components required by the proposed scheme are freely available in the form of reliable and well-tested free, open source software. It should be possible to retrofit existing health information systems to interoperate with the proposed system without enormous effort or expense.

CONTRIBUTED**Data linkage and the South Australian Cancer Registry**

Presenter: Wayne Clapton, SA Cancer Registry, Epidemiology Branch, Department of Human Services and Anti-Cancer Foundation of SA

Authors: Wayne Clapton, Colin Luke, Peter Chapman, Kevin Priest, Anh-Minh Thi Nguyen, Graeme Tucker, Maria Cirillo, Mary Merdo, Teresa Molik, Elaine Morton, Chris Scott, Heather Hall, Joanne Bell, Lesley Milliken

The South Australian Cancer Registry provides timely and accurate cancer epidemiological surveillance services for the State of South Australia. Cancer is a legally notifiable disease. Under specific State legislation, all hospitals, pathology laboratories and radiation oncology treatment centres are required to report all cases of invasive cancer (excluding non-melanotic skin cancers) to the Cancer Registry within one month of discharge and/or finalisation of diagnosis. The legislation also provides legal protection for those parties specified as exchanging data with the Cancer Registry. The Registry is required to protect the privacy and confidentiality of the data and not release it to any unauthorised party. Cancer Registries aim to achieve 100 percent capture of cases (100% ascertainment), to minimise case duplication and to be aware of major events, such as the death of a patient. Various strategies are employed to this end, including correlating data from different sources, obtaining data from Births Deaths and Marriages, and linking Cancer Registry data with other databases such as hospital administration systems, national cancer data accumulations at the Australian Institute of Health and Welfare, and the National Death Index. Probabilistic matching processes are used mostly; but deterministic methods also may be employed.

This paper will summarise the importance of data linkages for the SA Cancer Registry, will suggest other currently unavailable linkages which would be useful, will discuss the importance of legislation in terms of its enabling and inhibitory effects on health data linkage and will postulate on the role of a Cancer Registry in a health system where data linkage is more prevalent.

CONTRIBUTED**Consumer involvement in the Centre for Health Services Research**

Presenter: Rebecca Coghlan, Centre for Health Services Research, Department of Public Health, University of Western Australia

Author: Rebecca Coghlan

Consumer involvement in health decision making has become widely established as both a necessary part of our democratic process and part of the broader move to improve health outcomes. Many consumers are now expecting to be involved in decisions about their own health care. The partnership approach has led to the formal involvement of consumers and their groups in a large number of health initiatives. In keeping with these trends, the Centre for Health Services Research (CHSR) appointed a part time Consumer Liaison Officer (CLO) in January 2000. The position was established in conjunction with The Health Consumers' Council of WA.

Consumer participation in the CHSR has become an integral part of the research program and the CLO was involved in the strategic planning process for the CHSR covering the period 2001–2005. Dissemination of the research outcomes from the Centre's research program has been a pivotal role of the CLO. During 2000 the CLO undertook a Survey of Consumer Research Priorities which found that consumers want to participate in meaningful informed consent and an opportunity exists for health service researchers to assist them to achieve this goal. Consumers seek clear communication from health providers and they are keen to be involved with research from the ideas phase through to the dissemination of research findings in understandable language. The CLO has also given a number of talks on "How to Involve Consumers in Health Decision Making".

This paper will explore the consumer liaison role in the CHSR and examine the role it played in the successful launch in October 2001 of "Duty to Care: Physical Illness In people With Mental Illness". This study used the WA Linked Database to determine the extent to which users of mental health services, who comprise 8% of the WA population, had different rates of physical illness compared to the general population for the period 1980–98. The dissemination framework for this report will be discussed. This included the publication of a technical report and stakeholder and consumer summaries. The involvement of the CLO in this project greatly increased the profile of the study and ensured that both researchers, stakeholders from government and provider organisations, and consumers were well informed of the outcomes. The CLO has continued to communicate the research outcomes to the community whilst the project has entered a plateau phase. The presence of the CLO with the support of CHSR, together with community groups has helped to ensure that stakeholders, particularly the government, remain focused on changing the health system to better meet the needs of people with mental illness.

CONTRIBUTED**The use of linked ambulance data to estimate the effect of comorbidity on determinants and outcomes of out-of-hospital cardiac arrest in Perth, Western Australia****Presenter:** Judith Finn, Department of Public Health, University of Western Australia**Authors:** Judith Finn, D'Arcy Holman and Ian Jacobs

In order to describe the epidemiology and survival from out-of-hospital cardiac arrests in Perth, Western Australia, three years of St John Ambulance (WA Ambulance Service Incorporated) was linked to the WA hospital morbidity and mortality data using probabilistic matching. Whilst there have been numerous publications reporting short-term outcomes of out-of-hospital cardiac arrests the challenges posed by longitudinal follow-up of cases has limited the capacity for reports of longer-term outcomes. Similarly, there have been no large studies where the effect of comorbid status on survival has been estimated. As a result of data linkage it was possible to report 28-day and 12-month survival for all arrest victims and to model the effects of comorbidity on survival outcomes.

Of the 2,303 adults who experienced out-of-hospital cardiac arrests (excluding those from non-cardiac causes) in 1996–98, 85% (1,936) of the ambulance records linked to one or more WA hospital morbidity records for hospitalisations after 1980. Three approaches to the generation of an estimate of comorbidity were examined within this study, namely:

- 1) the generation of a Charlson comorbidity index (with and without the inclusion of heart disease);
- 2) a hierarchical classification of previous hospitalisation related to cardiovascular disease; and
- 3) total previous hospitalisation episodes and sum of hospital lengths of stay.

No statistically significant association was found between any of these comorbidity “scores” and 28-day survival. Of note however, was the seemingly counterintuitive finding that the likelihood of survival was higher for those victims with some comorbid history as compared to those with no previous comorbidity (as estimated by the Charlson score). More specifically, a history of hospitalisation with mention of Ischaemic Heart Disease within 12 months of the arrest was found to be associated with a two-fold increase in the likelihood of survival. No differences in patient and/or arrest characteristics were sufficient to explain this finding. One suggestion was that perhaps therapeutic interventions performed or medications prescribed during or subsequent to the hospital admission might have affected survival from cardiac arrest. Alternatively, different pathophysiologic processes may underpin cardiac arrest in patients with existing IHD.

Data linkage also enabled an estimation of the relationship between the victim's comorbid status and the hitherto unexplained association between the location of arrest and survival. Whilst a decreased likelihood of survival in those cardiac arrests occurring at the victim's residential address has been

clearly identified within the literature, and comorbid status mooted as the likely explanation, in most previous studies there has been insufficient information about the past medical history of the victim for conclusions to be drawn. Within this study, it was shown that when the Charlson comorbidity index was added to the adjusted logistic regression model of survival, location of arrest was no longer associated significantly with survival.

This study is the first in Australia to describe the epidemiology of out-of-hospital cardiac arrest in a population-based cohort. It brought together complete data from different sources through the process of record linkage that has allowed the novel application of indices of comorbidity to be incorporated in the identification and explanation of determinants of outcome.

CONTRIBUTED**Inside the Western Australian data linkage system****Presenter:** Carol Garfield, Department of Health, Western Australia**Authors:** Carol Garfield, Diana Rosman and John Bass

The WA Data Linkage System (formerly known as the WA Research Linked Database Project) is a population-based data linkage system established in 1995. It is a collaborative venture between the Centre for Health Services Research, Department of Public Health, UWA and the Health Information Centre, Department of Health, WA. It is primarily responsible for the linkage of unit records of the core health data sets and other relevant data collections, and the provision of linked data to support health planning, purchasing, evaluation and research.

The establishment and early development of the data linkage system is described in Holman et. al. (1999). Since 1999 the linkages within and between the core data sets have been extended, and a system of monthly updates for morbidity and mortality linkages and bimonthly updates for cancer and mental health linkages now ensures that the linked information remains current. Linkages to other data sources, such as the WA electoral roll, ambulance and Medicare data have taken place, with some of these now being part of the regular schedule. This presentation will give an up-to-date picture of the status of links within the WA Data Linkage System and describe some of the inner workings of the system.

Central to the system is the storage of the links and this has been structured using a “chain of links” method developed by Dr John Bass. It consists of *chains* of *links*, where each *link* is associated with a record in one of the core data sets. All *links* in a particular *chain* have been associated with the same individual person through the process of probabilistic record linkage. This method was developed with the potential to store genealogical links.

Due to the dynamic, multi-set nature of this chain of links system a unique set of resources to manage the system have

been developed within the unit. These include tools for loading of links with features to assist in maintaining the integrity of the links, displaying records in a chain for manual verification and extraction of linked data. It also has a history mechanism that enables the state of the links at any particular date to be re-established.

Holman CDJ, Bass AJ, Rouse IL, Hobbs MST. Population-based linkage of health records in Western Australia: development of a health services research linked database. *Australian and New Zealand Journal of Public Health* 23 (5): 453–459, 1999.

CONTRIBUTED

Death subsequent to treatment for heroin and amphetamine use: A record linkage study

Presenter: Geoff Gawthorne, Next Step Specialist Drug and Alcohol Services

Authors: Geoff Gawthorne, Anne Bartu, Sarah Johnson, Jim Codde, Elizabeth Unwin and D'Arcy Holman

The risk of all cause death and drug related death for a cohort (n=4425) of heroin and amphetamine dependent clients following their first admission to a specialist drug and alcohol service between 1985–1998 was studied using record linkage. Data from the Next Step client database were linked with the inpatient mental health information system and the registrar general's death information, data files held by the Western Australian (WA) Health Services Research Linked Database. All heroin dependents were on methadone maintenance treatment. After adjustment for age and sex the following results were obtained: heroin dependents had a greater hazard of all cause death than amphetamine dependents (HR: 1.44, p=0.02) and this effect was stronger for drug related deaths (HR: 2.35, p<0.001); clients with a psychiatric admission prior to treatment at Next Step had a greater hazard compared with those without a psychiatric history (HR: 2.19, p<0.001) and similar results appeared for drug related deaths (HR: 1.62, p=0.019), male clients had a greater hazard of all cause death (HR: 2.08, p<0.001) and this was more pronounced for drug related deaths (HR: 3.055, p<0.001). A male heroin dependent client with a psychiatric history had a relative hazard of all cause death of 6.58, and a hazard of drug related death of 11.65 compared with a female amphetamine client with no psychiatric history. The Next Step research unit is involved in ongoing linkage of client records to death records to maintain surveillance of client outcomes. Record linkage provides an effective method of examining the likelihood of death among heroin dependents and amphetamine users.

INVITED

Unpacking analyses relying on area-based data: Are the assumptions supportable?

Presenter: John Glover, Public Health Information Development Unit, University of Adelaide

Authors: John Glover, Diana Rosman, Merran Smith

The majority of work in Australia describing the association between health status, health service utilisation and socioeconomic status uses an area-based measure of socioeconomic status. This proxy measure is used because there is no direct measure of socioeconomic status in the health-related datasets. Its application requires a number of assumptions, including that people who move do so within geographic areas of similar socioeconomic status; and that, despite their size, the often large areas used provide a reliable indication of the socioeconomic status and health service utilisation of the individuals in the area. Also inherent in these analyses is an uncertainty arising from the use of data as to events (e.g. hospital inpatient separations), rather than individuals.

This paper uses the WA Data Linkage System to explore the extent to which hospital inpatient separation rates vary, both overall and by socioeconomic status of area of residence, when calculated at various levels of geographic aggregation (Census Collection District, postcode and Statistical Local Area). Methods applied include the calculation of correlation coefficients and examination of hospital separation rates by quintile of socioeconomic disadvantage of area, separately for events and individuals. Results are also provided of the extent of change in socioeconomic status of area of residence between an individual's admissions over a four year period.

The paper concludes with a discussion of additional links that would add to the value of the already valuable dataset within the WA Data Linkage System.

INVITED

The value of linked data for policy development, strategic planning, clinical practice and public health: An international perspective

Presenter: Michael Goldacre, Department of Public Health, University of Oxford, UK

Author: Michael Goldacre

Health care organisations should be “learning organisations”. The systematic analysis of data from records of illness and care is one strategy that can make important contributions to local learning and to the wider advancement of medical knowledge. The added contribution of linked data is that two or more items of information about an individual, when recorded separately in data systems, may have greater value if they are considered together than when either is considered alone. Illustra-

tive examples of linkage studies will be presented, mainly using data from the English National Health Service (NHS), including the following:

- Hospital admission rates have risen, seemingly inexorably year by year, for many years. Yet waiting lists for care seem as long as ever. Is the NHS really treating more patients each year than ever before? Trends in admission rates will be outlined, comparing trends from unlinked and linked data. Their implications for policy and planning will be discussed.
- There are growing expectations by health professionals and the public that hospitals will know about, and learn from, their patients' death rates (and other outcomes). Most hospitals' routine data systems can only be used to identify deaths that occur in the initial hospital admission for the treatment of interest. Linkage to data from death certificates can be used to identify all deaths following care, regardless of when and where they occur. Examples from linkage will be given of death rates following surgery, fractured neck of femur, stroke, and myocardial infarction.
- Record linkage can provide insights into the descriptive epidemiology of disease (trends in Crohn's disease will be shown); and can provide insights into associations between clinical procedures and subsequent disease (abortion and long-term risk of breast cancer will be shown as an example).

I shall comment briefly on examples of record linkage studies from other parts of the world, including an international study in which results are being pooled from linked datasets from several countries. I shall also comment briefly on what I see as some of the key strengths and weaknesses of current data linkage systems in England. For example, most of the English experience with linkage comes from relatively small populations (e.g. 2.5 million in the former Oxford health region) which may not be representative of the general English population in some important respects. We are now piloting linkage of hospital admission data and death data nationwide in England, a population of 50 million.

INVITED

The value of linked data for research into the social determinants of health

Presenter: Michael Goldacre

Author: Michael Goldacre, Department of Public Health, University of Oxford, UK

1. *Socioeconomic factors* – such as occupation, social class, income, and employment status – are strongly associated with levels of ill-health and premature death (more so in some societies than others). Much of the systematic, large-population evidence on this comes from studies of cross-sectional design, based around denominator data from population censuses. Despite the undoubted value of such studies, they also have limitations. For example, they do not generally provide any

scope for the analysis of the time-course of ill health or for the study of effects of changes over time in socio-economic or employment status. Important information which corroborates and extends their findings comes from longitudinal studies of census data linked to vital statistical records in post-censal years. Examples will be given from the English Longitudinal Study, run by the national Office of Population Censuses and Surveys. The Longitudinal Study has included samples of all people present at the 1971 and 1981 censuses, classified by their social characteristics at those times, and has followed them up by linkage to subsequent records in the National Health Service Central Register and in later censuses.

Findings on occupational social class show a widening of mortality rates between the classes over recent times. Age-standardised mortality in classes IV and V (predominantly manual occupations) was 53% higher than mortality in classes I and II (which include professional and managerial occupations) in 1976–81; and the gap had increased to become 68% higher in 1986–92. Findings on unemployment showed that men and women who were unemployed at census time had an excess mortality, compared with the employed, of about 33% over the following ten years. Men who were unemployed at both the 1971 and 1981 censuses had mortality rates which were double those of the employed. Although adjustment for social class reduced the gap a little, neither social class nor pre-existing ill health accounted for much of the raised mortality. This lends support to the hypothesis that unemployment has an independent effect on premature mortality.

2. Important *early life* determinants of ill-health in later life include low birth-weight and under-nutrition in utero and infancy. I shall comment on the contribution of historical child health records, linked decades later to mortality records, to the development of the Barker hypotheses on foetal programming and disease in later life.

3. One of the most important *socio-demographic* determinants of adult ill health is old age. The use of health services rises substantially with increasingly old age which, because of increasing longevity in industrialised countries, is a concern to health service planners. But, for many people, the chief determinants of lifetime use of health services may be proximity to death; and, because most people die at advanced ages nowadays, increased ill health and increased use of health services may be determined more by being close to death than by old age itself. Results from record linkage studies will be shown which try to distinguish between the effects of proximity.

CONTRIBUTED

Applications of record linkage in cardiovascular disease: Monitoring trends and incidence, determinants and outcome of coronary heart disease and stroke

Presenter: Michael Hobbs, Department of Public Health, University of Western Australia

Authors: Michael Hobbs, Steve Ridout, Konrad Jamrozik and Ian Jacobs

While mortality from both coronary heart disease (CHD) and stroke in Australia has been falling for over thirty years, the prevention and control of these is a National Health priority. The development of reliable methods for monitoring the incidence and prevalence of CVD and its determinants, including medical care, to guide policy and provision of services is therefore essential. Until recently there has been considerable uncertainty about the extent to which the declining mortality in CHD or stroke could be attributed to falling incidence or improved survival or the respective contributions to this of lifestyle changes or medical care. Studies such as the WHO MONICA Project suggest that both have been important but questions remain about the way that these may have changed over time or the potential for future gains. At the population level, disease registers provide the best means for answering such questions but are difficult to run and are financially impossible to sustain over the long-term. Record Linkage of administrative data sets and clinical databases, supported by intermittent disease registration offers a reasonable compromise.

Background to studies in Perth

Research into the epidemiology and medical care of selected aspects of CVD has been conducted in Perth over a period of nearly thirty years, through a combination of disease registers (myocardial infarction and stroke), community risk factor prevalence surveys and linkage of routinely collected hospital morbidity data (HMD) and death records. The research potential of the each of the above has been greatly enhanced through crosslinkage.

More recently, these studies have been extended by linkage of the above to clinical information in the form of cardiothoracic surgery and cardiology registers, ambulance and laboratory (biochemistry) data. Central to our present research has been the development of a person-based file of all persons admitted to hospital or dying from CVD in Western Australia during the period 1978–2000 (the Linked Vascular File) derived from the Centre for Health Services Research Linked Database. Examples of epidemiological studies based on the above include:

- Validation of coding of acute myocardial infarction (AMI) and stroke in HMD through linkage of HMD to disease Registers (*Perth Coronary Register, 1970–71; Perth MONICA Study, 1984–93; Perth Heart Attack Study 1998 (PHAS-98; Perth Community Stroke Registers, 1989–90, 1995–96).*
- Long-term trends in incidence, prevalence, recurrence, case fatality longer-term survival of AMI and stroke (*linkage of HMD and death data).*

- Assessment of the impact of new bio-markers of myocardial damage on epidemiological studies and on the coding of AMI in HMD (*linkage of HMD, laboratory data and PHAS-98 data.*)
- Assessment of the potential impact of medical treatment on the incidence of acute coronary events through estimation of relative and absolute risks in persons ever-admitted to hospital for CHD (*linkage of HMD and death data*)
- Development of predictive equations for CHD (*linkage of risk factor survey data to HMD and mortality data).*

CONTRIBUTED

Applications of record linkage in cardiovascular disease: Establishing illness episodes for planning and evaluation of health services for coronary heart disease and stroke

Presenter: Michael Hobbs, Department of Public Health, University of Western Australia

Authors: Michael Hobbs, Konrad Jamrozik, Steve Ridout and Robyn Broadhurst

While unlinked HMD provide an indication of total hospital utilisation for chronic diseases such as CHD and stroke, planning of health services requires information on episodes of treatment, often involving multiple hospital admissions. Similarly the assessment of disease outcomes, whether in terms of case-fatality is only meaningful in relation to completed episodes of care. Definition of disease episodes in routinely collected data is necessarily arbitrary and needs to be tailored to the particular purpose. For example, assessing the need for acute coronary care units or acute stroke services will be based on shorter episodes of care – generally 28 days – than for determining requirements for coronary artery revascularisation procedures (CARPs) after myocardial infarction or rehabilitation after stroke for which a twelve month episode might be more appropriate.

Examples of the use of record linkage to generate treatment episodes appropriate to different phases of medical care include:

- (i) The emergency retrieval of patients with acute coronary events (*linkage of ambulance and emergency department data to HMD and laboratory data for bio-markers of myocardial damage).*
- (ii) The impact of new technologies for coronary artery revascularisation on hospital utilisation (*linkage of HMD and death records to clinical registers).*
- (iii) Requirements for rehabilitation facilities and continuing care needs (domiciliary and residential care) following stroke (*linkage of HMD and death data to ACAT, domiciliary care and residential care data*)
- (iv) Monitoring the uptake of systematic secondary prevention in persons admitted to hospital for CHD or stroke (*linkage of HMD and Health Insurance Commission HIC MBS and PBS data*)

INVITED

The value of linked data for research into health outcomes

Presenters: D'Arcy Holman, Centre for Health Services Research, Department of Public Health, University of Western Australia; and John Bass, Western Australia

Authors: D'Arcy Holman and John Bass

This paper provides an overview of progress in the establishment, technical design, research applications and operational procedures of the Western Australian Record Linkage Project. The Project is research infrastructure that systematically links the available administrative health data within a single Australian state of population 1.8 million. It brings together the six core data elements of birth records, midwives' notifications, cancer registrations, inpatient hospital morbidity, inpatient and public outpatient mental health services data and death records, together with other state-based health data systems, an extensive network of research databases and limited Commonwealth health data on West Australians. The state of linkage is continuously updated as new data become available. Linkage is performed using probabilistic matching of patient names and other identifiers. Geocodes for spatial analysis are assigned using address linkage and mapping software. The project is managed jointly by the WA Department of Health and the University of WA and operational arrangements include a significant commitment to the participation of consumer and professional groups.

Research applications of the linked data are having a beneficial impact on the health system as illustrated by the case of "Duty to Care", a project that investigated the risks of physical illness in the 8% of the population who have been users of mental health services. Future directions include a further increase in community participation, cross-jurisdictional linkage to the Commonwealth Medical Benefits Scheme, Pharmaceutical Benefits Scheme and residential care data for Western Australia and the development of family links to support future population-based human genome research. The paper includes an explanation from Dr John Bass of a protocol designed to achieve cross-jurisdictional linkage, based on matching with name identifiers, which preserves the anonymity of clinical data.

By June 1997, the project had taken 2.5 years to develop the system and link 7 million core data records from 1980–95. The system is consistent with international best practice benchmarks, from four centres of excellence, in respect of the study population, core data sets, matching and geocoding, and collaborative networks. There are prospects to redress deficiencies in primary medical contact and other data resources, validation studies, tracing systems and a more supportive legal framework. The WA Linked Database will be used in combination with medical record audits to provide a comprehensive evaluation of health system performance.

INVITED

The value of linked data for research into health services use and expenditure: An Australian perspective

Presenter: Dr Chris Kelman, Information and Research Branch, Department of Health and Ageing

Author: Dr Chris Kelman, Information and Research Branch, Department of Health and Ageing

Research into health service use and expenditure in Australia has been impeded by the fact that health services are funded by separate arms of government. In order to estimate "whole of care" treatment costs on a per person or disease basis, it is necessary to have information on both inpatient and outpatient services available. This presentation discusses the work that has been done to redress this problem by the states and Commonwealth and discusses the utility of these data for policy development and resource planning.

CONTRIBUTED

Ovarian cancer in Western Australia, 1982–98: A population-based review of trends and outcomes

Presenter: Crystal Laurvick, Centre for Health Services Research, Department Public Health, University of Western Australia

Authors: Crystal Laurvick, James Semmens, Yee Leung, Anthony McCartney, Ian Hammond and D'Arcy Holman.

As part of the Western Australian Safety and Quality of Surgical Care Project (SQSCP) we reviewed the clinical epidemiology, treatment pathways and outcomes for all women diagnosed with ovarian cancer in Western Australia (WA). The WA Record Linkage System was used to obtain cancer, morbidity and death records for all patients with a diagnosis of ovarian cancer in 1982–98. Additional clinical staging information was confirmed for 169 women diagnosed at King Edward Memorial Hospital in 1995–98. Patients were identified using the International Classification for Diseases diagnosis and procedure codes pertaining to ovarian cancer.

During the study period, 1,336 women were diagnosed with ovarian cancer. For women with staging information, 61% (102/169) were diagnosed with International Federation of Gynaecology and Obstetrics (FIGO) stage III or IV. The mean age at diagnosis was 62 years (SD15). The incidence rate decreased on average 1% per year ($p=0.05$) and the mortality rate remained stable. Rates of surgical intervention increased.

The majority of women (77%) received surgery to manage the disease, with a median inpatient stay of 14 days (IQR:10-19) and an in-hospital mortality of 4.4%. The overall relative sur-

vival at 5 years was 38% (95% CI:34-41). Survival among women with advanced disease was 39% (95% CI: 25.1-53.1) at 3 years compared to 92% (95% CI: 80.1-100) survival for women with stage I or II disease.

The present study demonstrated that there has been little change in the incidence and mortality from ovarian cancer in the past two decades. Surgery remains a cornerstone in the primary management of ovarian cancer and the increasing trends in surgical intervention suggests the demand for surgery will continue. The high proportion of women diagnosed with advanced disease limits treatment pathways for cure. New strategies are required for the diagnosis of early and often asymptomatic disease if real improvements in morbidity and mortality are to be expected.

CONTRIBUTED

Trends of cataract surgery and post-operative endophthalmitis in Western Australia (1980–1998): A population-based study

Presenter: Jianghong Li, Department of Public Health, University of Western Australia

Authors: James Semmens, Jianghong Li and Nigel Morlet, on behalf of teamEPSWA

Background: Endophthalmitis is a serious intra-ocular infection that can lead to severe visual loss. Cataract extraction is the most common cause of post-operative endophthalmitis. Although the reported incidence of post-operative endophthalmitis is low, the prevalence of endophthalmitis has increased significantly with the rising number of cataract procedures performed in most developed nations. During the last 30 years there have been significant changes in surgical practice for cataract extraction, from intracapsular extraction with intraocular lens implantation to extracapsular extraction in the late 1970s and to small incision phacoemulsification from early 1990s onwards. It is unknown if the transition to phaco-emulsification has reduced the incidence of endophthalmitis. As phacoemulsification surgery is almost universally performed today in Australia, the lack of information regarding the risk of post-surgical endophthalmitis is a significant gap in our knowledge of the safety of this new surgical method. Most existing analyses of post-operative endophthalmitis have been handicapped by small and incomplete data, making comparisons and statistical validity of data difficult. Our study provides a population-based estimate of the incidence rate of cataract procedure and postoperative endophthalmitis and examines their socio-demographic trends, using the hospital morbidity data from the Western Australian Record Linkage Project.

Methods: Cataract procedures and endophthalmitis diagnoses for 1980–1998 were identified using the international classification for diagnosis and procedure codes (ICD-9) and its modification (ICD-9-CM). The aggregated data for this

study was complemented by other data sources, including the microbiology and anaesthetic databases from Royal Perth Hospital and surgeon logbooks. The accuracy of coding for the patient with post-operative endophthalmitis was validated by chart review. Annual age-specific and age-standardised rates for cataract procedures per 100,000 PY were estimated using the World Standard Population as the standard set of weights. Poisson regression was used to estimate the trends in the incidence of cataract procedures. The incidence rate of endophthalmitis was calculated as a case per 1,000 cataract procedures and differentiated by gender, age, surgical type, locality and hospital. Differences in the incident rate of endophthalmitis were assessed with the 95% confidence interval and the Z-test.

Results: During the 19-year period, the number of cataract procedures performed annually increased 7 times, from 1,335 in 1980 to 9,652 in 1998. The increasing trend is much stronger for the metro area and private hospitals than rural areas and public hospitals. Despite of the dramatic increase in cataract operations and the changes in surgical techniques over the last 30 years, there was no significant change in the incidence of endophthalmitis, which averaged about 2 per 1,000 procedures. The incidence rate fluctuated with time (1 to 3 per 1,000) and varied with the location of surgery (0.65 to 16.4 per 1,000).

Conclusions: There is a significant reduction in the risk of endophthalmitis following the transition from intracapsular to extracapsular extraction. There is however no further significant reduction following the shift from extracapsular to phaco extraction. The incidence rate is higher in the metro area and among elderly female patients.

CONTRIBUTED

From evidence to practice: Population-based monitoring of the use of breast-conserving surgery for the treatment of breast cancer using record linkage of routinely collected data

Presenter: Kim Lim, Epidemiology and Surveillance Branch, NSW Health Department

Authors: Kim Lim and Tim Churches

By the early 1990s there was substantial evidence that breast conserving surgery (BCS) plus radiotherapy was an effective but less disfiguring alternative to mastectomy in the treatment of early breast cancer. This study examines trends in and predictors of the use of BCS as opposed to mastectomy in the NSW population.

Methods: Cancer registry records of all cases of breast cancer in women resident in NSW diagnosed in the period 1991 to 1998 were linked using probabilistic techniques with records of breast surgery extracted from hospital separations data for 1991 to 1999. The trend in the propor-

tion of women undergoing BCS was examined and logistic regression was used to examine the influence of geographical remoteness, age, degree of spread, insurance status and hospital throughput.

Results: Between 1991 and 1998, 26,718 NSW women were diagnosed with breast cancer. 25,138 (94%) of these cases were linked with a total of 98,678 hospital admission records. The proportion of metropolitan women who received BCS increased monotonically from 37% to 54%, and from 29% to 43% for rural women. All covariates were significant in predicting the probability of mastectomy versus BCS.

Conclusions: Changes in clinical practice occur slowly across the entire population. The urban/rural difference may be due to relative access to radiotherapy services. On a population basis, it is not clear what proportion of women with breast cancer are candidates for BCS.

CONTRIBUTED

Data linkage to estimate resource and service utilisation for palliative care clients

Presenter: Andrew McAlindon, Data Linkage Consultant, International Institute of Hospice Studies, Flinders University of South Australia

Authors: Andrew McAlindon and Amy Abernethy

Background: Demand for palliative care services is growing within the Australian Health System. Palliative care is increasingly being offered to clients with terminal cancer and other end-stage illnesses. Up-to-date and accurate costs associated with the provision of these services are required for the successful negotiation of resources from the state and Commonwealth governments and for local and system-wide service planning.

Setting: Southern Adelaide Palliative Services (SAPS) is a comprehensive palliative care program serving the southern region of Adelaide, South Australia, including specialist nursing and medical support, social work, inpatient acute hospital and hospital care, outpatient visits, home care, nursing home consultations, a bereavement program, volunteers, complementary care, and day care. SAPS interfaces with public and private hospitals, district nursing and domiciliary care.

Objective: To determine the optimal record linkage methods for linking together a range of databases to estimate the resource and service utilisation for palliative care clients in a regional palliative care service.

Methods: Access to a range of databases from local service providers, and state and Commonwealth governments was negotiated for the 931 SAPS clients who died in 1999. These databases contained service utilisation and claims data for public and private hospitals (inpatient and non-patient), Medicare, Pharmaceutical Benefits Scheme, Royal District Nursing Service (RDNS), Southern Domiciliary Care (SDC) and Aged

Care Services. A Master Client Index (SMCI) for the SAPS clients was constructed from these databases and from the South Australian Enterprise-wide Patient Master Index (EMPI) using a combination of deterministic and probabilistic record linkage methods. Using an iterative process, the SMCI was progressively refined and used to extract the final utilisation data.

Results: In 1999, 931 SAPS clients died with an average length of stay of 111 days. Eighty-nine percent of clients were hospitalised, with an average cost of AU\$7,888 per client. Half of all hospitalisation costs from the last year of life occurred within the final three months. On average each client also received 18 prescriptions (AU\$489), 25 district nursing visits (AU\$1,031) and 6 domiciliary care services (AU\$281). Nineteen clients used respite care with an average length of stay of 25 days (AU\$1600) and 74 clients used permanent care with an average length of stay of 93 days (AU\$7600)

Conclusion: A combination of deterministic and probabilistic record linkage methods are required to accurately estimate the resource and service utilisation for palliative care clients. The construction of a Client Master Index allows patient identifiers from a range of databases to be spliced together to further increase the accuracy of the linkage process.

CONTRIBUTED

The sensitivity of probabilistic record linkage: Estimating the number of "false negatives" in a linkage involving client records of a large domiciliary care organization

Presenter: Luke Marinovich, Silver Chain Nursing Association

Authors: Luke Marinovich, Stuart Fuller, Michael Hobbs and Gill Lewin

The continuing Care Linkage Study has been established to examine patterns of health service utilisation, including hospitalisation, domiciliary care, geriatric services and residential care, in persons with continuing disability. As part of this study, the Data Linkage Unit of the Department of Health (WA) performed a linkage between client records from Silver Chain Nursing Association (Western Australia's largest home care provider) and records from the Western Australian Hospital Morbidity Data System (HMDS). Personal identifiers used in the matching process included family name, given names and date of birth.

Where a client is referred to SCNA from a hospital, this is recorded within the SCNA record. It is therefore possible to examine the characteristics of clients who are recorded as having at least one hospital admission according to Silver Chain records (from referral), but did not link to an actual hospital admission record in the HMDS. The present paper will describe an analysis of these "false negatives" comparing their characteristics with those for which a matching hospital admission record could be found.

A previous linkage of SCNA records with hospital admission records was performed in 1998 without the use of client names. Comparisons will be made between the groups of matched and unmatched records in this initial linkage and the more recent linkage using full names and dates of birth.

CONTRIBUTED**Issues in the use of unique patient identifiers to link statistical data**

Presenter: Ric Marshall, Victorian Department of Human Services

Authors: Ric Marshall, Neil Powers and Carmel Heaney

Using practical examples, this paper describes the current approach adopted by the Victorian Department of Human Services to make linked administrative data on health service utilisation available for research on health outcomes.

Like all other state and territory health authorities, the department has made de-identified unit record data from its hospital morbidity collection available for ethically approved research projects for many years. However, the potential for data linkage has been limited by the anonymisation process, which ensures that names and addresses are not provided to the central data collection. Thus linkage is only possible through the probabilistic matching of data items such as the hospital-specific unit record number, the Medicare card number and dates of birth, admission or discharge. Recent developments in legislative provisions have enabled development and trialling of registers of patients who use services across Victorian public hospitals. While the ultimate objective of these registers is to enable patients to build a whole-of-life record that is accurate and portable, with timely access to this information for clinicians, these registers also have the potential to enable improved linkage of statistical data sets for planning, policy and approved research developments at state and regional levels.

However, the development of this type of data linkage capacity places an obligation on service managers to codify guidelines for the use of unique patient identifiers or statistical linkage keys in de-identified administrative data collections. These guidelines need to address such issues as:

- minimisation of potentially identifying information,
- supervision of the use of data,
- data editing, and
- subsequent use and destruction of data sets.

Examples of studies incorporating these guidelines and using de-identified unit record datasets will be provided and the linkage results discussed.

Victoria is currently leading the development of broad standard national guidelines as a first step towards a model code of practice for custodians of de-identified health data collections. The distinction between identified and de-identified or anonymised health information is not always clear-cut; consequently this type of code of practice is increasingly becoming essential to adequately address the rapidly developing privacy requirements that formally limit the collection, use and disclosure of identified personal health information.

CONTRIBUTED**Linked health data in the Northern Territory**

Presenter: Stephen Moo, Business Information Management, Northern Territory Department of Health and Community Services

Authors: Cherie Shepherd, Mary-Anne Measey and Stephen Moo

The Northern Territory Department of Health and Community Services (DHCS) has had a functional data warehouse for almost six years. It contains data from over twenty departmental systems as well as from non-government organisations. One of the most important benefits of the data warehouse is that it allows for the linkage of data from various feeder systems. Data are linked using a unique identifier. This identifier is used by all operational systems within the department as well as by some external agencies. Data from the warehouse can be used to provide information on the range and cost of services accessed by a particular population group at one point in time as well as track changes in health status and service access over time. The data contained in the data warehouse are routinely used for a variety of purposes including:

- Management reporting and decision making.
- Health economics analyses.
- Epidemiological analyses.
- Routine reporting for various agencies.

Some specific examples are:

- Analyses of data to help measure the costs of providing services across the care continuum for various population groups within the NT.
- Finance and Management Reporting for Tiwi and Katherine Health Boards.
- Practitioners in remote communities have used the data to monitor their clients care plans and medications over time.
- The Northern Territory has a rich and valuable data source that will assist the department to make decisions on the basis of sound information. It will also enable researchers to undertake various epidemiological studies on the health of the population of the NT.

CONTRIBUTED**Electronically linked health records: Current proposals in New South Wales****Presenter:** Chris Puplick, Privacy NSW**Author:** Chris Puplick

The draft Health Records and Information Privacy Bill was released as a "public exposure draft" by the NSW Minister for Health in December 2001. It is (at the time of writing) expected that the Bill will be tabled in Parliament in March 2002. This paper will review the role of consent in the linking of health records, comparing the recommendations of the NSW Ministerial Advisory Committee on Privacy and Health Information with the provisions contained in the draft Bill. Mr Puplick will briefly outline how the Bill, if enacted, will affect public and private health service providers and administrators.

CONTRIBUTED**The use of record linkage in epidemiological research into maternal and child health****Presenters:** Gail Reading, Peter Cosgrove and Hoan Nguyen, Telethon Institute for Child Health Research Perth, Western Australia**Authors:** Maxine Croft, Hoan Nguyen, Gail Reading, Peter Cosgrove, Melinda Berinson and Gillian Smith

The Maternal and Child Health Research Database (MCHRDB) was established in the mid 1980s. The database is a population-based set of data of all births registered in WA from 1980–1999. The midwives' notifications of births are linked to Registrar General birth registrations creating a composite birth record. These records are then linked to hospital morbidity data, mortality data and census data. Birth records are also linked to the cerebral palsy register, the birth defects register, the disability services commission database and other study cohorts as required. In addition, siblings have been linked producing a valuable data resource for research into family health and future intergenerational studies.

Over the last 15 years the database has proved to be an effective resource for maternal and child health research. In the area of policy and planning, the Department of Health of WA have recently funded a project at the Telethon Institute for Child Health Research called the Collaboration for Applied Research and Evaluation (CARE). The aim of this project is to facilitate partnerships between researchers and policy makers undertaking research and evaluation projects and to assist agencies to translate research of various forms into system and program level policy. The MCHRDB is one of the main data sources for this new project.

The presentation will start with an overview of the Maternal and Child Health database and the CARE initiative mentioned above. A "snapshot" of a number of different research projects

that have used data from the database will be presented including examples of mapping to illustrate the geographic distribution of study cohorts and/or populations.

CONTRIBUTED**Surgical site infections following orthopaedic surgery: Statewide surveillance using linked administrative databases****Presenter:** Thomas V. Riley, Department of Microbiology, University of Western Australia**Authors:** Claudia Thomas, Helen Cadwallader and Thomas V. Riley

In many institutions prospective surveillance programmes to monitor the incidence of post-operative infection can be difficult to implement due to limited human and technical resources. In addition, prolonged patient follow-up, up to one year, may be required for implant surgery. Traditional methods of surveillance can be enhanced by using administrative databases to assist in case finding and facilitate overall surveillance activities.

The aim of this study was to identify the incidence of surgical site infection (SSI) in patients who had undergone total hip replacement (THR) or total knee replacement (TKR) surgery in all Western Australian hospitals in 1999 using the Western Australian Health Services Research Linked Database. The database was used to identify patients who underwent THR or TKR surgery in 1999 using ICD-9-CM and ICD-10-CM codes. Of these, patients who had been given an infection diagnosis code plus the external cause "surgical operation with implant of artificial internal device" codes (E878.1 in ICD-9-CM, Y83.1 in ICD-10-CM) were identified. This allowed all patients to be followed for at least one year after surgery. Patients who died from other causes during the follow-up period were identified from linked mortality data. A total of 1476 THR and 1875 TKR procedures were identified from 21 WA hospitals (11 public, 10 private) during 1999. There were 169 infections identified, giving an overall cumulative incidence of 5%, (95% CI 4.3-5.7) [THR (4.86%, 95% CI 3.77-5.95) and TKR (5.15%, 95% CI 4.15-6.15)]. Only 23 infections were coded specifically as "infection and inflammatory reaction due to hip or knee replacement". The remainder had been given alternative infection codes, but had also been given the external cause code "surgical operation with implant of artificial internal device". There was a total follow-up time of 5012 person-years, giving an incidence rate of 33.72 infections per 1000 person-years. Most infections (96%) occurred within one year of surgery. Patients aged over 80 years experienced a significantly higher rate of infection after THR, when compared to patients aged 80 or less (z -test, $z=2.56$, $p=0.0150$), but not for TKR ($z=0.35$, $p=0.726$). Eighty-five patients (50.3%) developed a SSI during the same hospital admission as the sur-

gical procedure, the remainder were re-admitted to hospital at a later date. For THR, 58% developed a SSI during the same admission, compared to 44% for TKR. The incidence of SSIs for procedures performed in public hospitals was higher, 6.4% (95% CI 5.2-7.66) compared to 3.8% (95%CI 2.97-4.75) for private hospitals ($z=3.39$, $p=0.007$). The WA Health Services Research Linked Database provided a unique opportunity to review the incidence of SSIs in patients undergoing THR or TKR surgery in WA hospitals.

INVITED

Record linkage: Data quality, tool development and substantive research

Presenter: Leslie Roos, Manitoba Centre for Health Policy, University of Manitoba, Canada

Author: Leslie Roos

Science is shaped by “entry points” which facilitate the study of important problems. Such entry points may be concepts or theories; they may be tools which permit asking new questions or approaching old questions much more efficiently. I believe two such entry points for the study of health and health care are longitudinal administrative data and record linkage techniques. Research with administrative data has been conceptualised as a triangle, with the base of the triangle being data acquisition, maintenance and enhancement. In the middle is tool development, tools to facilitate population-based, provider-based, spatial and temporal analyses. At the top of the triangle is substantive, multi-disciplinary research in health care and health. Record linkage is an essential tool for dealing with both the base of the triangle – data and its quality – and the top of the triangle – substantive research. I will discuss the role of record linkage in both these areas. Other tools (such as our Concept Dictionary and our hospital costing algorithms) valuable for the task of creating an information-rich environment will be described. Almost twenty years ago, we recognised the need for a record linkage system which would be relatively simple and integrate well with our existing operations. Such a system has been used to build the Manitoba research registry from periodic snapshots of the population registries used to administer Canada’s provincial health insurance systems. The critical design feature of the research registry is its ability to provide key variables for the entire population at any time since 1970.

Linkage can also be used for assessing data base quality. By substituting for record abstraction when both hospital and physician files are available, record linkage can cost-effectively highlight the quality of both variables and data sets.

Our substantive examples will highlight the active use of record linkage to enhance existing data sets and support new kinds of research. The 1986 Census/Manitoba Linkage Project has been designed to create a multi-purpose, population-based data set from the linkage of 5% of the households on the 1986

census with comprehensive health information (hospital, physician, and registry) for 1983–1989 ($N=44,629$). One research program based on these data is oriented toward the new field of social epidemiology in its concern with neighbourhood, as well as individual, characteristics. Another has emphasised the relationships between unemployment and health.

A second example uses record linkage to support a multi-method research program building on clinically defined cohorts. The designs typically look backward and forward from an index event (treatment at a university-based Sleep Disorders Centre). Individuals in these cohorts have been linked to the research registry, and then to substantive files (hospital, physician). Control groups have been selected from both the research registry and a provincial Heart Health Survey. Utilisation and expenditure data are being used in a series of studies.

The presentation will conclude with a discussion of challenges for the future. Such challenges include issues of privacy/confidentiality and the appropriate investments in tools and data bases.

INVITED

Inequalities in child health: Bringing together new data sources to assess the roles of family, community, education and health care

Presenter: Noralou Roos, Manitoba Centre for Health Policy, University of Manitoba

Authors: Noralou Roos, Marni Brownell and Diane Watson

Objectives: The goal of this Research Program is to identify determinants of variability in health and functioning of children, and factors that support resilience. Our analyses will focus on informing policy on ways to reduce inequities and improve the health and functioning of children in Canada. Questions to be addressed include: To what extent do children’s health and functioning differ across Manitoba’s educational and health jurisdictions? To what extent do individual, family, school and community characteristics, as well as the delivery of medical care across jurisdictions contribute to, or buffer against, inequities in children’s health and functioning? Which policy options have the greatest potential for reducing inequalities in child health and functioning?

Methodology: A cross-sectional methodology, using population-based data for Manitoba, will be adopted. Small area variation techniques will be used to describe individual characteristics, family circumstances, school environments, community contexts, and medical care utilisation, as well as health and educational achievement of Manitoba’s children. The entire Manitoba child population will be used to calculate expected health and functioning values for each region, given the region’s socioeconomic characteristics. These values will be

compared to the observed values to highlight which region's outcomes are higher or lower than expected. We will examine factors that attenuate or protect against inequities in health and educational achievement across socioeconomic levels during two periods of childhood: early in elementary school and late in high school. A longitudinal, population-based analysis will be used to evaluate the additive and cumulative effects of individual characteristics, family circumstances, school environments, and community context on child health and functioning, educational achievement, and the inequalities in health and education across socioeconomic groups. A matched cohort design using a neighbourhood-level measure of household income will be used to compare familial and community characteristics of families that move out of the lowest income quintile with those that remain in the lowest socioeconomic group. Differences in children's health and functioning will be examined for the two groups. Differences in the cost of health services for these two groups will be used to estimate the potential size of savings in health care spending due to upward movement in a family's socioeconomic status from the lowest quintile.

This Research Program builds on the existing population-based, anonymised research data repository at the Manitoba Centre for Health Policy and Evaluation. The expanded infrastructure will include education, family services and health data. Anonymised individual-level data from a local school division, the Department of Genetics, and a preschool cohort study will be added as part of this Research Program. We are exploring the potential use of student data from a major college. Having these various data sources together and linkable in one place presents a unique and exciting opportunity to identify policy options for reducing health inequities and supporting resilience.

CONTRIBUTED

Estimating the cost of injury using the Western Australia Data Linkage System and Injury Cost Database

Presenter: Diana Rosman, Data Linkage Unit,
Department of Health, Western Australia

Authors: Diana Rosman and Delia Hendrie

Injury is a major cause of death and disability in Australia. The overall burden from injury ranks third when measured as years of life lost due to premature mortality and fifth when measured as disability adjusted life years. It has also been estimated that in Western Australia, injuries account for nearly 10% of hospital bed-day costs or about \$50 per head of population per year.

The WA Data Linkage System, comprising linked hospital admission and death records for all of WA since 1970, is unique to Western Australia. This resource allows injury index admissions and all subsequent related admissions to be accumulated across this time period. In addition, an Injury Cost Database has been established in WA from injury compensa-

tion claims and other related sources. A severity scoring program developed at Johns Hopkins University (ICDMAP) provides the link between the diagnosis codes recorded in hospital admission records and injury costs components in the Cost Database.

This paper will describe a process for using the tools described above to attach cost estimates when available information is limited to unlinked hospital admission records. The resultant total cost estimates will be presented for injuries resulting from road crashes and from falls over the three year period July 1994 to June 1997 (with follow-up to June 1999).

CONTRIBUTED

Measuring data and link quality in a dynamic multi-set linkage system

Presenter: Diana Rosman, Data Linkage Unit,
Department of Health, Western Australia

Authors: Diana Rosman, Carol Garfield, Stuart Fuller, Alexia Stoney, Todd Owen and Geoff Gawthorne

The international quality management standard ISO 9001:2000 consists of a generic set of guidelines to be applied in any industry. In Health Information Management both product (information) and service (information delivery) aspects need to be considered.

The WA Data Linkage System comprises 3.7 million chains of linked records drawn from six population-based datasets and is updated monthly. In addition, links to six additional datasets are updated six monthly and ad hoc linkages to special cohorts also occur. Although matching algorithms and the levels of clerical review have been tuned to minimise mis-matches and missed matches, the system also needs to be efficient to achieve a quality product and a quality service. In probabilistic record linkage, variations in the values of key variables are tolerated in order to achieve optimum linkage. Consequently, among a group of linked records, variations in the values of key variables may be due to inherent data errors or may signal that incorrect links have been made. Elsewhere, link quality has been assessed by theoretical or empirical methods. Results have been measured against an expected linkage rate or compared with a "gold" (or even a "platinum") standard. However, these methods are problematic in a dynamic multi-set linkage system as there is no benchmark for comparison.

In January 2001, a random sample of 5,000 records for admissions to hospital during 1990 was selected. Complete chains of links for each of these were extracted for manual review. It was determined that 28 (0.6%) of the 4854 distinct chains contained one or more records that had been linked incorrectly. One year later the task was repeated and errors were discovered in just 15 (0.3%) of the 4868 selected chains.

This paper examines the factors contributing to variations in

key variables within linked chains and proposes a system for on-going monitoring of the quality of the links with rules based on those used by the six expert linkage operators during the two pilot manual linkage audits.

CONTRIBUTED

Evolving treatments for primary urolithiasis: Impact on services and renal preservation in 16,679 patients in Western Australia

Presenter: James Semmens, Centre for Health Services Research, Department Public Health, University of Western Australia

Authors: D'Arcy Holman, Stan Wisniewski, James Semmens and John Bass

Background: The quality and safety of health care is now a major health issue, which is being addressed in all western countries and is a high priority in Australia. Increasing public awareness, including a more active participation in individual health care and active consumer organisations, along with concomitant political forces, are changing the face of health care in developed countries. Improving the safety and quality of health care is now a central concern of all those in the health care system. Clinicians and hospitals are now being held accountable for the level of health care provided while at the same time governments are pushing for improved economic efficiency in the health care system. In keeping with this focus, the Western Australian Safety and Quality of Surgical Care Project (SQSCP) was established in 1996, as a collaborative quality assurance venture of the Royal Australasian College of Surgeons, the Department of Public Health (University of Western Australia) and the State Health Department. The overall aim of the SQSCP is to evaluate surgical outcomes at a state population level to improve the safety and quality of surgical care within the state.

Objectives: The present study uses data from the SQSCP to examine changing treatments for primary presentation of urinary lithiasis and their effects on readmissions, repeat procedures, cumulative hospital utilisation and renal preservation.

Patients and methods: Linked hospital morbidity records were used to identify first-time admissions for renal and ureteral calculi from 1980 to 1997 in the population of Western Australia. The cases were followed to mid-1999 and actuarial methods were used to estimate risks of further hospital admissions and procedures including the loss of a renal moiety.

Results: Between 1980–1997 the total rate of inpatient procedures for urinary stones more than doubled at a time when the rate of first-time hospital admissions increased by only 13% and the conservative management of stones remained constant at around 59%. The predominant procedure for stone management was initially open lithotomy, replaced in the early 1980s by percutaneous nephrostolithotomy (PCN) and soon supple-

mented by extracorporeal shock wave lithotripsy (ESWL). The changes in technology led to a four-fold increase in procedural re-admissions within 30 days of primary separation. This was due to repeated, staged or postponed interventions, often involving the use of stents or a second treatment with ESWL. The risk of surgical intervention fell from 48% to 32%, whilst the cumulative length of stay over the first 12 months fell from 7.8 to 3.9 days. The risk of kidney loss fell significantly from 2% to less than 0.1% during the period.

Conclusions: The main reason for more interventions were short-term procedural readmissions. ESWL reduced the need for invasive procedures and cut cumulative hospital stay despite more readmissions. Renal preservation improved by a factor of ten.

CONTRIBUTED

Linking the prairies to the outback: What can New South Wales learn from the Manitoba data linkage project?

Presenters: Marian Shanahan and Kees van Gool, Centre for Health Economics Research and Evaluation, University of Technology, Sydney.

Authors: Marian Shanahan and Kees van Gool

Establishing a population-wide linked data project is both complex and expensive. It is therefore necessary to ensure that if New South Wales is to undertake such an exercise the costs are minimised (avoid making costly mistakes) and that research outcomes warrant the substantial investment. Using the experiences gained in the design and establishment of the Manitoba Population Health Information System (POPULIS), this paper will examine possible methods in which data can be linked in NSW. Starting with a description of POPULIS, we analyse the lessons learned from the Manitoba experience and assess whether its successes can be replicated in NSW.

The aim of this paper is to draw out the steps that will be necessary to ensure the establishment of an effective research tool. Linked data can be used for a wide variety of research efforts, the focus of this paper however will be its use in health economics and health services research. In particular, we examine how a future NSW data linkage project could assist in answering questions regarding resource allocation, improving access and equity and improve our understanding of the health care system's impact on health outcomes and utilisation. Whilst there is little doubt that linked data can potentially help our understanding of how to improve the efficiency and equity of the health care system, the extent to which this occurs depends on the validity and accuracy of the data and the data links. The question that remains is whether, in the NSW context, the expected return warrants the investment.

INVITED**The value of linked data for policy development, strategic planning, clinical practice and public health: An Australian perspective****Presenter:** Richard Smallwood, Chief Medical Officer, Department of Health and Ageing**Author:** Richard Smallwood

Australia's health data systems were designed originally to provide a mechanism and infrastructure to reimburse people and providers for their health care costs. At this time, the data was collected with no knowledge of, or interest in, why the person needed care, or whether the care was successful. Separate datasets were collected by various levels of government and often "protected" from unscheduled use by legislation. Since then, Information Technology has come of age. Its power has made large-scale data collection, analysis and monitoring not just a possibility but a responsibility for any industry – including health care. It seems that we are only just starting to take up this challenge – we certainly have huge resources of collected data about health services, however most of it is still oriented toward health service payments and seriously lacking in the kind of clinical detail that allows the evaluation of health outcomes. Much of it is still not readily available for research purposes. We need an evidence-based, systems approach to the planning of services and interventions. We need to monitor interventions to see what works as intended, what doesn't, and what works better than expected, and then start the whole process again. Good information is central to understanding and monitoring the effect of interventions properly.

The challenge is thus to change the focus of data collection away from health service payments to one of ongoing monitoring and surveillance of clinical outcomes to allow the quality cycle to take place. We need to ensure that the information we gather can test key assumptions about current service provision and can force greater emphasis to be placed on the systems' effectiveness in meeting the needs of individual clients. To achieve these ends, the use of linked data is fundamental. To be able to provide a holistic view of health services and outcomes across primary care, inpatient and residential care is a basic requirement. The collection of appropriate data and timely access to data provides the foundation for an infrastructure which will allow health system appraisal and performance measuring and support a rational basis for future policy development.

To aid in the appropriate management of these resources, an extensive infrastructure of ethics committees is being established, guided by the Australian Health Ethics Committee of the National Health and Medical Research Council. In addition, as part of a number of activities to develop national networks and privacy principles for electronic health records, business rules for linking statistical collections using unique patient identifiers are with Commonwealth and State/Territory Health Ministers for endorsement.

CONTRIBUTED**From data linkage to policy and back again****Presenter:** Christine Stone, Department of Human Services, Victoria**Authors:** Christine Stone, Jenny Howard, Andrew Ramsden, Maree Roberts, Jane Halliday

Background: Comprehensive and valid neonatal data are required for effective monitoring and planning of neonatal services in Victoria. The most recent DHS review of neonatal services in Victoria, undertaken in 1998, recommended that a Neonatal Services Advisory Committee should be established to "advise and assist the Department in the provision of the best possible standard of neonatal care to Victorian babies" and that issues of data collection were an important consideration in providing such advice and assistance. The Data and Information Sub-Committee was set up to explore these data issues and a data linkage project was established to explore the feasibility of linking maternal and neonatal data from routinely collected datasets and usefulness of the linked dataset.

Aim: This presentation describes the technological and political issues that arose during the process of this data linkage project.

Method: The routinely collected data sources were the Perinatal Data Collection Unit (PDCU), the Victorian Admitted Episode Dataset (VAED) and the Australian and New Zealand Neonatal Network (ANZNN). The records of the neonates, their mothers and their subsequent transfers were linked using probabilistic record linkage software, AutoMatch and a complex multistage approach.

Result: This data linkage was more complex and resource intensive than the usual linkage projects for a number of reasons including the complexity of birth data and differentiating multiple births, the multistage process that was necessary to gradually build up the linkage variables and the political process of working across organisations with different datasets requirements.

Conclusions: Record linkage of routinely collected data provides potential to inform policy however there is still work to be done to improve policy around record linkage. In a field which is expanding so rapidly, we believe that this is an iterative process.

CONTRIBUTED

Linkage of the Victorian admitted episodes dataset

Presenter: Vijaya Sundararajan, Victorian Department of Human Services

Authors: Vijaya Sundararajan, Toni Henderson, Michael Ackland and Ric Marshall

Background: The Victorian Admitted Episodes Dataset (VAED), the state's hospital morbidity dataset, is an episode-of-care level dataset. Turning the VAED into a case-level dataset has potential benefits in epidemiologic, health services research and quality of care research. However, at this time, there is no unique variable which can be used to separate the dataset into cases.

Methods: Initially, for the fiscal years 1994–2000, we evaluated the quality of data by comparing the agreement of identifiers using well-matched pairs of observations. Next, four linkage variables were created for use:

- 1) year/day/month of birth/postal code/continent of birth/gender (link1);
- 2) hospital code/hospital record number (link2);
- 3) 3-digit medicare suffix/8-digit medicare number (link3);
- 4) year/day/month of birth/postal code/gen*der (link4).

Link1 was the variable used for the first pass at linkage and was the basis for the newly created variable "caseno". After this first pass, two derivative sets were created: one with observations which did not group with any others based on agreement of their link1 (the "orphans") and one with observations which did group (the "many"). For the second pass, link2 was compared between observations in the orphan set to those in the many set – if there was agreement on link2, the orphan observation was given the "caseno" from the many dataset's observation whose link2 agreed with it. Passes 3 and 4 were conducted in the same fashion after creating the second and third orphan and many sets. The pilot project was conducted using SAS 8.0.

Results: The public VAED had Medicare numbers for 86.6 to 100% of its observations; the private VAED was more limited with only 30–34% of its observations with a Medicare number until fiscal year 2000 when the rate was 100%. The coding error rate was low: most of the potential linkage variables had 98–99.9% agreement in pairs of well-matched observations. These data have been used in a number of studies which will be briefly described.

CONTRIBUTED

The use of linked data in health surveillance in New South Wales

Presenter: Lee Taylor, Epidemiology and Surveillance Branch, NSW Health Department

Authors: Lee Taylor, Kim Lim and Tim Churches

The Epidemiology and Surveillance Branch of the NSW Health Department began experimenting with probabilistic record linkage in 1993, and has carried out a data linkage program since 1994. The program has two components: annual linkage of selected data sets that are of particular importance for population health surveillance; and ad hoc linkages to meet specific needs.

Linkage of the following datasets occur annually: NSW Central Cancer Registry (CCR) data to the NSW Inpatient Statistics Collection (ISC), to provide information on patterns of treatment for cancers. For example, we have compared patterns of treatment for breast cancer in women living in urban and rural NSW. Internal linkage of the ISC, to enable estimation of the numbers (and rates) of persons with certain chronic diseases. For example, we have estimated rates of renal failure among Aboriginal and Torres Strait Islander people. Internal linkage of the Notifiable Diseases Database, to yield improved estimates of the number of people with chronic infectious diseases such as hepatitis B and C. NSW Midwives Data Collection (MDC) to the Registry of Births, Deaths and Marriages birth registration data, to allow improved estimation (using capture-recapture methods) of the number of babies born to Aboriginal and Torres Strait Islander parents. We also return information on unregistered births to the Registry. MDC to the ISC, to enable reporting of post-natal length of stay and obstetric interventions by health insurance status.

Examples of recent ad hoc linkages include:

- Linkage of the Royal North Shore Spinal Injury database with the ISC to examine morbidity following spinal injury.
- Linkage of pregnancy ultrasound data from the Bankstown-Lidcombe Hospital with the MDC and NSW Birth Defects Register to examine false positive and false negative results of prenatal diagnosis with ultrasound.
- Linkage of the St George Hospital Hypertension in Pregnancy Database with the MDC to support a study of the risk of recurrent hypertension in pregnancy.
- Internally linked ISC data are being used to examine pregnancy rates following tubal sterilisation, and revascularisation procedure rates following angioplasty.
- Linkage of the Sydney Obstetric and Gynaecological Ultrasound database with the MDC to examine miscarriage risk associated with amniocentesis.

Within the NSW public health system, data linkage occurs in a complex legal, ethical and policy environment. The collection and use of personal health data are governed by the NSW Public Health Act 1991, the NSW Health Administra-

tion Act 1982 and the NSW Privacy and Personal Information Protection Act 1998. Policy for linkage of personal health data is described in the Health Department's Information Privacy Code of Practice 1998. The policy provides for referral of certain proposed linkage projects to the Department of Health Ethics Committee (DoHEC), which operates in accordance with NHMRC guidelines.

CONTRIBUTED

Data linkage in the public health assessment of ecosystem change

Presenter: Shilu Tong, School of Public Health, Queensland University of Technology

Authors: Shilu Tong, Rod Gerber, Rodney Wolff and Ken Verrall

Earth's ecosystems are being disrupted by the combined weight of rapid industrialisation, population growth and intensive consumption. The resultant climate change, stratospheric ozone depletion, loss of biodiversity and other environmental changes pose significant risks to human health, and perhaps survival. The assessment of the public health consequences of ecosystem change will provide important data for the formulation of public policies and international cooperation to mitigate and adapt these changes. It is abundantly clear research of this kind is both of massive importance in its field and of significant benefit to Australian health and medical research. This paper will report some preliminary work undertaken in Queensland and will illustrate how data linkage analysis is conducted in this study.

CONTRIBUTED

Putting data into context: Findings from linking Medicare health service use and expenditure data with longitudinal health survey data.

Presenter: Anne Young, Research Centre for Gender and Health, University of Newcastle

Author: Anne Young

Introduction: The Australian Longitudinal Study on Women's Health (ALSWH), funded by the Commonwealth Department of Health and Ageing, is a study of the health and well being of three large cohorts of Australian women. The ALSWH has made extensive use of linked survey and Medicare/Department of Veterans' Affairs data. Results are presented to illustrate the value of the linked data for informing policy makers about provision of health services and for monitoring compliance with best practice guidelines.

Methods: The project recruited three large, nationally representative cohorts of women, aged 18–23 years (n=14,228),

45–50 years (n=13,338) and 70–75 years (n=12,317) in 1996. Self-administered postal surveys are completed every three years and include a wide range of measures of demographic, social and health-related factors. Almost 23,000 of the women have given written consent for the release of their individual records from the Health Insurance Commission. Data relating to more than 1.5 million Medicare/DVA services provided to these women during 1995–1999 have been linked to the first two phases of their survey data. Changes in health, health service use and the costs of services were examined according to age, urban/rural residence and socioeconomic status. Analysis of the linked data for subgroups of women, such as frequent attenders to general practice, and the use of best practice guidelines for diabetes care were also examined.

Results: For all age groups, women with lower socioeconomic status tended to have lower out-of-pocket costs for general practice visits. However, women in rural and remote areas reported poorer access to doctors who bulk bill and Medicare data showed these women had higher out-of-pocket costs than women living in urban areas. Many of the very frequent attenders to general practice had suffered a major personal illness, and the survey data showed that many also had very difficult personal and social circumstances. Women with diabetes, and those who developed diabetes, reported poorer health and greater use of health services and medications than women without diabetes. Medicare data helped to quantify the increased health service use and expenditure over time (for services outside hospital) for these women. However their Medicare data also showed that compliance with best practice guidelines for diabetes care, such as monitoring HbA1c, was sub-optimal.

Conclusions: The linked data provide information on medical conditions and social circumstances which are valuable for understanding health service use. Inequalities in the provision and costs of health care services were identified. The linked data can be used to monitor compliance with best practice guidelines for care and to determine the impact of strategies designed to improve the health and well being of women.