Identification of Australians from culturally and linguistically diverse backgrounds in national health data collections

Ilse Blignault and Abbas Haghshenas

Abstract
In multicultural Australia, comprehensive and up-to-date information on ethnicity and health is essential to guide policy and service development in the health sector. Data collected for purposes other than research are a potentially important source of information. This study explored the extent to which indicators of cultural and linguistic diversity are currently included in national health and welfare service data collections, and the data standards employed.

We identified and reviewed 44 relevant bodies of work: 7 national data dictionaries, 15 national data sets, 10 national health data collections and 12 national surveys. Each of the large data dictionaries (health, community services and housing assistance) contained several ethnicity-related variables. Immigrant Australians were identified (usually by country of birth, sometimes by language, and occasionally by period of residence or year of arrival) in all the major national health and community data sets, health data collections and surveys. Australian Bureau of Statistics standards and classifications relating to cultural and linguistic diversity were widely used. Researchers, health policy makers and planners should fully exploit these secondary data sources, as well as undertaking or commissioning primary research.

The current situation: sources, uses and gaps in data
Australia is a culturally diverse nation, with elements of the population displaying differing demographic, socioeconomic, cultural and linguistic characteristics. Such differences are associated with significant health and social disadvantage, as well as with problems in accessing appropriate services. High quality ethnicity data are necessary to guide policy and service development in the health sector. Potential data sources include primary data, which are collected by researchers for specific purposes; and secondary data, which are collected by others for other purposes or, commonly, as a general purpose data set. In this paper, we review the data on cultural and linguistic diversity available in national health and welfare data collections, censuses and surveys, and consider their potential for contributing to the evidence base in
multicultural health. As sources of Indigenous health information and their limitations have been discussed elsewhere,\(^7\) we focus here on sources of information about the health and wellbeing of immigrant Australians.

A number of government agencies and other national bodies currently collect, collate and disseminate statistical data that potentially provide evidence relevant to ethnicity and health.\(^4\) These include the Australian Bureau of Statistics (ABS), the Australian Institute of Health and Welfare (AIHW), the Commonwealth Department of Health and Ageing, other Commonwealth Government Departments, and academic bodies such as the National Centre in HIV Epidemiology and Clinical Research. State and territory health authorities collect routine data on specific diseases, conditions and episodes of care. Inpatient statistical collections, for example, provide the main source of data on hospital inpatients and day cases.

A feature of these collections is that they include data from the whole population, thus supporting comparison between different subpopulations where they are identified, as well as facilitating analysis of time trends and geographic differences. Use of information from secondary sources has significant advantages, though there are drawbacks as well. The most obvious benefits result from savings in cost and time. Analysis of existing data provides a useful starting point for further research by identifying gaps in knowledge and suggesting research questions or hypotheses. Drawbacks generally relate to the fact that the end-user has little control over aspects such as the selection of variables, questions and output categories, or the degree of aggregation of data across different groups, time periods and regions. Such difficulties can be minimised through the use of statistical standards, which also improve the compatibility and comparability of data derived from different sources. A more difficult problem to overcome is lack of currency, particularly in the population census, conducted every 5 years.

The case for maximising the use of secondary sources for multicultural health research in Australia is a strong one, particularly since the development of statistical standards on cultural and linguistic diversity.\(^8\) At the 2001 Census, 23% of Australian residents were born overseas and 15% spoke a language other than English at home.\(^9\) In some states and regions, particularly Victoria and New South Wales, these proportions are much higher.\(^9\) A recent review for the National Falls Prevention for Older People Initiative highlighted the dearth of epidemiological data and falls prevention research among Australians from culturally and linguistically diverse backgrounds.\(^10\) Comprehensive national data are sorely needed to inform suicide prevention programs for immigrant communities,\(^11\) and to support the development of culturally competent mental health services.\(^3\)

Use of existing data requires knowledge of their existence and the means for accessing them, together with the resources to acquire them and the ability to interpret and integrate the results.\(^12\) Such information is not always easily identifiable by people working outside specialist data collection units. A further issue relates to the quality of data from various administrative collections and surveys. Not all data are equally reliable and valid. Within a given data set, data on one variable might be quite good (both complete and accurate), while data on another variable might be of more doubtful quality.\(^13,14\) One of the advantages of secondary analysis is that, since the data already exist, evaluation of their quality and appropriateness can be done in advance of actual use.

In 1992, the AIHW published the first inventory of Australian health data collections that contained information on ethnicity,\(^15\) an expanded subset of the main inventory.\(^16\) This was followed in 1996 by the information paper Non-English-speaking background and Indigenous status — identification in national health and welfare data collections.\(^4\) Since then, there has been a substantial body of work around the collection and reporting of data on cultural diversity, including the publication of Standards for statistics on cultural and language diversity and the introduction of measures to improve the collection of Indigenous statistics including the identification of Aboriginal and Torres Strait Islander peoples.\(^8,7\)

The standards were developed by the ABS in consultation with other organisations in response to growing needs and a request from government. The minimum core set consists of four variables:
country of birth of person, main language other than English spoken at home (identified, following extensive consultation, as the most general purpose language variable), proficiency in spoken English, and Indigenous status (part of the core set for collections not specifically focused on migrants). The full standard set also contains ancestry, country of birth of father, country of birth of mother, first language spoken, languages spoken at home, main language spoken at home, religious affiliation, and year of arrival in Australia.

There are four related ABS classifications covering religious groups, languages, countries, and cultural and ethnic groups. Each classification is based on a hierarchical structure that supports the capture, coding and storage of data at a base level, while allowing users the flexibility to output statistics at the level of classification that best suits their needs. Such an approach facilitates comparisons with previous data using different classifications and enables more detailed and complex analysis at a later date.

It is timely, almost a decade after the AIHW information paper, to review the collection of ethnicity data in the health and community sectors at a national level.

Method
We collated a list of relevant documents, starting with those produced by the ABS and the AIHW, reviewed them and added others. The material fell into three broad categories: national data dictionaries; national minimum data sets and data set specifications; and national surveys. We excluded data sets and surveys that were concerned solely with service or system data: establishments, business practices or labour force.

We made extensive use of the AIHW Knowledgebase, which has since been replaced by METeOR — AIHW’s Metadata Online Registry (http://meteor.aihw.gov.au/content/index.phhtml/itemId/181162). Most of the documents were accessed online, primarily from the AIHW. Some of the national reports were borrowed from the university library, while others were requested from the publisher. Clarification or further information was sought by phone if needed. For each document, we asked two questions: Which ethnicity variables are included? What data standards are used? In the case of national surveys, we also asked, How are the data reported?

Some definitions
A national data dictionary is the authoritative source of data definitions and data collection standards for a given sector or field of activity in Australia. It is designed to improve the consistency and comparability of data across jurisdictions and to make data collection activities more efficient and effective.

A national minimum data set (NMDS) is a minimum set of data elements and standards agreed by the relevant national information management group for collection and reporting at a national level.

A data set specification is a core set of data elements and standards that are not mandated for collection but are recommended as best practice.

Results
Results are presented in four sections: national data dictionaries — health and community; national minimum data sets and data set specifications — health and community; national health data collections; and national censuses and surveys — health and community.

National data dictionaries
The three main sources of data definitions in the health and welfare fields are the National health data dictionary (NHDD), the National community services data dictionary (NCSDS) and the National housing assistance data dictionary (NHADD). The NHDD is the oldest and largest of these documents, with Version 12 containing over 350 data elements. There are also several smaller data dictionaries, most of which were designed to support collection of a minimum or core data set (see following section). Four of these — the Aged care assessment program (ACAP) data dictionary, the Continence aids assistance scheme (CAAS) data dictionary, the Day therapy centre (DTC) program data dictionary and the Home and community care (HACC) data dictionary — contain definitions of data elements beyond those mandated for reporting.
Box 1 shows the ethnicity-related variables included in the seven data dictionaries and the classification systems employed. With the exception of “need for interpreter service” (NHDD) and “interpreter services required” (NCSDD and NHADD), all of the items are based on ABS classifications and standards.

National minimum data sets and data set specifications

At the end of 2004, the AIHW Knowledgebase listed 15 NMDSs and 7 data set specifications in the health sector. All are described in the NHDD Version 12 and its Supplement.22,23 Of the 17 data sets that contained client data, 12 included country of birth and 3 also included a language variable (Box 2).

Three of the health data sets record Indigenous status: Draft National Minimum Data Set — Children’s Services; Data Set Specification — Diabetes (Clinical), and Draft Data Set Specification — Telehealth.22 However, two do not collect any data concerning the client’s cultural background: National Minimum Data Set — Injury Surveillance and Data Set Specification — Cancer (Clini-
Health Consumers

We identified another three relevant NMDSs in the community sector: the Commonwealth–State/Territory Disability Agreement (CSTDA) National Minimum Data Set,31 the Aged Care Assessment Program (ACAP) Minimum Data Set,26 and the Home and Community Care (HACC) Minimum Data Set.32 The CSTDA National Minimum Data Set was redeveloped from a snapshot collection into an ongoing collection.33 Data elements cover the characteristics of disability service users and their carers, service type and service type outlet.31 The ACAP Minimum Data Set is an ongoing, client-centred collection designed to capture data on people who receive an Aged Care Assessment Team comprehensive assessment. It covers client characteristics and circumstances (eg, socio-demographic information, health status and functional abilities), as well as details of the assessment process and the care plan/outcome.26 The HACC Minimum Data Set is an ongoing, client-centred

### 2 Ethnicity-related variables in national minimum data sets (NMDS) and data set specifications (DSS)

<table>
<thead>
<tr>
<th>Ethnicity-related variables</th>
<th>Data sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country of birth</td>
<td>NMDS — Admitted patient care V2 ✓(1)</td>
</tr>
<tr>
<td>Preferred language</td>
<td>NMDS — Admitted patient mental health care V2 ✓(1)</td>
</tr>
<tr>
<td>Main language other than English spoken at home</td>
<td>NMDS — Admitted patient palliative care V2 ✓(1)</td>
</tr>
<tr>
<td>Proficiency in spoken English</td>
<td>NMDS — Alcohol and other drug treatment services V4 ✓(1) ✓(2)</td>
</tr>
<tr>
<td>Interpreter services required</td>
<td>NMDS — Community mental health care V2 ✓(1)</td>
</tr>
<tr>
<td></td>
<td>NMDS — Non-admitted patient emergency department care V1 ✓(1)</td>
</tr>
<tr>
<td></td>
<td>NMDS — Perinatal V1 ✓(1)</td>
</tr>
<tr>
<td></td>
<td>NMDS — Residential mental health care ✓(1)</td>
</tr>
<tr>
<td></td>
<td>DSS — Acute coronary syndrome (clinical) ✓(1)</td>
</tr>
<tr>
<td></td>
<td>DSS — Cardiovascular disease (clinical) V1 ✓(1) ✓(2)</td>
</tr>
<tr>
<td></td>
<td>DSS — Health care client identification V1 ✓(1)</td>
</tr>
<tr>
<td></td>
<td>Draft DSS — Community Nursing Minimum Data Set Australia ✓(1) ✓(2)</td>
</tr>
<tr>
<td></td>
<td>Classification systems employed: 1 = Standard Australian Classification of Countries 1998, ABS Cat. no. 1269.0. 2 = Modified 2-digit level version of Australian Standard Classification of Languages 1997, ABS Cat. No. 1267.0. 3 = Standards for Statistics on Cultural and Linguistic Diversity 1998, ABS Cat. No. 1289.0. 4 = National community services data dictionary V2.</td>
</tr>
</tbody>
</table>
collection that focuses on the clients of HACC-funded agencies. Data elements relate to the characteristics and circumstances of clients (eg, socio-demographics, carer availability and residential location) and the types and level of assistance provided by the HACC Program.32,34

As shown in Box 2, all three of the data sets report birthplace and at least one language variable. Each is supported by a data dictionary,26,31,32 and the CSTDA National Minimum Data Set has an extensive guide.35 Birthplace and the language variables “main language other than English spoken at home” and “proficiency in spoken English” are based on ABS classifications and standards.

**National health data collections**

We identified ten national health data collections holding ethnicity data. These are listed in Box 3, together with the agency responsible for collation of the data at the national level. Some are concerned with overall mortality and morbidity, others with specific diseases (AIDS, cancer or diabetes). A few are based on a minimum data set. For example, the National Hospital Morbidity Database, which is compiled by the AIHW from data supplied by the state and territory health authorities, incorporates data from the Admitted Patient Care National Minimum Data Set. Similarly, the Perinatal Data Collection, Australia, which is collated annually by the National Perinatal Statistics Unit, is based on data from the Perinatal National Minimum Data Set.

All ten data collections contain birthplace (mostly based on the ABS classification), with maternal country of birth recorded in Congenital Malformations, Australia. Three of the collections include duration of Australian residence.

The Causes of Death collection contains data about all deaths in Australia,36 while the National Mortality Database holds non-identifiable unit

---

**3 Ethnicity-related variables in national health data collections**

<table>
<thead>
<tr>
<th>National data collection (Organisation responsible for collation)</th>
<th>Country of birth</th>
<th>Country of birth of mother</th>
<th>Period of residence in Australia*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causes of Death collection (Australian Bureau of Statistics)</td>
<td>✓(1)</td>
<td>✓(2)</td>
<td></td>
</tr>
<tr>
<td>National AIDS Registry (National Centre in HIV Epidemiology and Clinical Research)</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Cancer Statistics Clearing House (AIHW)</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Coroners Information System (Monash University National Centre for Coronial Information)</td>
<td>✓(1)</td>
<td></td>
<td>✓(3)</td>
</tr>
<tr>
<td>National Hospital Morbidity Database (AIHW)</td>
<td>✓(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Notifiable Diseases Surveillance System (Communicable Diseases Australia)</td>
<td>✓(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Mortality Database (AIHW)</td>
<td>✓(1)</td>
<td></td>
<td>✓(2)</td>
</tr>
<tr>
<td>National Diabetes Register (AIHW)</td>
<td>✓(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perinatal Data Collection, Australia (National Perinatal Statistics Unit)</td>
<td>✓(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congenital Malformations, Australia (National Perinatal Statistics Unit)</td>
<td>✓(1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Where the COD collection and the National Mortality Database uses the phrase “Period of residence in Australia”, the NCIS uses “Years in Australia”. AIHW = Australian Institute of Health and Welfare.

Classification systems employed:
1 = Standard Australian Classification of Countries (SACC) 1998, ABS Cat. No. 1269.0.
2 = National Health Data Dictionary (NHDD) V12.
3 = NHDD V8.
record files from that collection. Data are supplied to the AIHW by the state and territory Registrars of Births, Deaths and Marriages, based on death certificates and coroner’s reports. Data items include cause of death (International Classification of Diseases, 10th Revision [ICD-10]), country of birth, period of residence in Australia and other socio-demographic variables. Additional information about the deceased person and the causes and circumstances of death are recorded in the National Coroner’s Information System.

AIDS is a notifiable condition in all Australian health jurisdictions. Cases of AIDS are notified by the diagnosing doctor through state and territory health authorities to the national HIV surveillance centre. Yearly surveillance reports include age standardised average annual incidence per 100 000 population by region of birth. The National Notifiable Diseases Surveillance System, which coordinates the national surveillance of more than 50 communicable diseases or disease groups, collects birthplace as part of its enhanced tuberculosis data set.

The National Cancer Statistics Clearing House (NCSCH) holds the only national database of cancer incidence in Australia. It is maintained by the AIHW, with data provided from cancer registries in each state and territory. The NCSCH produces annual reports of national incidence and mortality data, with additional analyses, such as

<table>
<thead>
<tr>
<th>4 Ethnicity-related variables in national censuses and surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethnicity-related variables</strong></td>
</tr>
<tr>
<td><strong>National census or survey (Organisation responsible)</strong></td>
</tr>
<tr>
<td>Australian Needle and Syringe Program Survey (National Centre in HIV Epidemiology and Clinical Research)</td>
</tr>
<tr>
<td>Australian Secondary Students Alcohol and Drug Survey (Centre for Behavioural Research in Cancer, The Cancer Council Victoria)</td>
</tr>
<tr>
<td>Bettering the Evaluation and Care of Health (Australian Institute of Health and Welfare General Practice Statistics and Classification Unit, University of Sydney)</td>
</tr>
<tr>
<td>Clients of Treatment Service Agencies (National Drug and Alcohol Research Centre)</td>
</tr>
<tr>
<td>General Social Survey (Australian Bureau of Statistics)</td>
</tr>
<tr>
<td>Growing Up in Australia: The Longitudinal Study of Australian Children* (Australian Institute of Family Studies)</td>
</tr>
<tr>
<td>National Dental Telephone Interview Survey (Australian Institute of Health and Welfare Dental Statistics and Research Unit, University of Adelaide)</td>
</tr>
<tr>
<td>National Health Survey (Australian Bureau of Statistics)</td>
</tr>
<tr>
<td>National Physical Activity Survey (Australian Institute of Health and Welfare)</td>
</tr>
<tr>
<td>National Survey of Mental Health and Well being (Australian Bureau of Statistics)</td>
</tr>
<tr>
<td>Women’s Health Australia: the Australian Longitudinal Study on Women’s Health (Research Centre for Gender and Health, University of Newcastle)</td>
</tr>
</tbody>
</table>

* Ethnicity data collected for study child and all other members of the household.

Classification systems employed:
1 = Standard Australian Classification of Countries 1998, ABS Cat. No. 1269.0
2 = Modified from Standard Australian Classification of Countries for Social Statistics, 1990, ABS Cat. No. 1269.0
3 = Main language spoken at home by parents
4 = Australian Standard Classification of Languages 1997, ABS Cat. No. 1267.0
5 = Modified from Australian Standard Classification of Languages 1997, ABS Cat. No. 1267.0
6 = Standards for Statistics on Cultural and Linguistic Diversity 1998, ABS Cat. No. 1289.0
Health Consumers

differentials in cancer rates by birthplace, undertaken periodically.\textsuperscript{41}

The National Diabetes Register collects non-clinical information, including birthplace, about people with insulin-treated diabetes mellitus. Data are obtained from Diabetes Australia through the National Diabetes Services Scheme, the Australian Paediatric Endocrine Group and the Menzies Centre for Population Health in Tasmania. Characteristics of registrants, including region of birth, are described in an annual statistical report.\textsuperscript{42}

Statistical reports on Medicare and the Pharmaceutical Benefits Scheme, collated by the Health Insurance Commission, do not include information relating to cultural background as these data are not routinely collected.

National censuses and surveys

The 5-yearly Census of Population and Housing provides the major source of estimates of the numbers of Australians from culturally and linguistically diverse backgrounds,\textsuperscript{9,43} in addition to information on a range of topics such as education, employment, income and housing. These estimates are necessary for measuring the size of various groups and for calculating the rates of various events such as death.\textsuperscript{8} Cultural and linguistic variables in the 2001 Census included: ancestry; birthplace of individual, female parent and male parent; proficiency in spoken English; language spoken at home; religious affiliation; and year of arrival in Australia. All variables were defined in the census dictionary and were coded using the relevant Australian Standard Classification.\textsuperscript{44}

Box 4 lists 12 other national health and community censuses and surveys that have incorporated one or more ethnicity-related variables. Most collected birthplace and at least one language variable (usually language spoken at home), and half collected year of arrival in Australia. ABS standard classifications were employed in some cases but not in others. In the various reports, findings for birthplace or language group have not always been presented. While some reports refer to ethnicity in the sample description, others do not refer to it at all.

The 2001 National Health Survey was the first in a new series of regular, triennial health surveys designed to collect information about the health of Australians including long-term medical conditions and recent injuries; the use of health services, health-related aspects of lifestyles, and demographic and socioeconomic characteristics. Four questions on ethnicity were included: country of birth, main language spoken at home, proficiency in spoken English and year of arrival (ABS classifications and standards used). About 26 900 people from all states and territories and across all ages groups were surveyed.\textsuperscript{45} Information on ethnicity and health obtained in the 2001 and earlier surveys has been reported in several ABS and AIHW publications, usually at a broad or general level.\textsuperscript{41,46,47}

The 2002 General Social Survey (GSS) also represented a new approach. It included questions on health and disability, housing, education, income, labour force, transport, crime, recreation activities and a range of other topics. The same four questions on ethnicity were asked as in the 2001 National Health Survey. Over 14 500 people aged 18 years or older were surveyed nationally, and broad analyses by ethnicity were reported for a range of topics.\textsuperscript{48} The GSS is conducted 4-yearly.

The AIHW conducts a regular (2- to 3-yearly) survey on use of tobacco, alcohol and other drugs — the National Drug Strategy Household Survey. In the 2001 survey, almost 27 000 people aged 14 years and over provided information on their drug-use patterns, attitudes and behaviours. Questions were asked about birthplace, year of arrival in Australia and language (main language and other languages spoken at home). Summary tables of drug use for special populations included a breakdown by main language spoken at home (English or other).\textsuperscript{49} Work has commenced on the 2004 survey, which will be the eighth in the series.\textsuperscript{46}

The Australian Secondary Students Alcohol and Drug Survey monitors the use of tobacco, alcohol and drugs by secondary school students. It was developed from a triennial national survey of tobacco and alcohol use, first conducted in 1984.\textsuperscript{50} In the 2002 survey, data were obtained from 23 417 male and female students aged 12–17 years in 363 schools. Limited personal and demographic information was collected, with main language spoken at home being the only ethnicity-related variable. Results were not reported by language group.\textsuperscript{50,52}
The Bettering the Evaluation and Care of Health (BEACH) Program is a continuous survey of about 1000 general practitioners (GPs) each year. Data collected include characteristics of the GP–patient encounter, the GP and the patient, patient reasons for encounter, problems addressed at the consultation and their management, and tests and investigations ordered. Patient cultural background is recorded simply as NESB (non-English speaking background), Aboriginal or Torres Strait Islander.53

National oral health surveys include the National Dental Telephone Interview Survey (NDTIS) and the Child Dental Health Survey. The NDTIS is a periodic survey of the general population. Topics include access to dental care, self-assessed dental health status, present dental-health needs, use of dental services and preventive behaviours, satisfaction with dental services, and experience of and attitudes to dentistry. It was last conducted in 2002. Results from earlier surveys have been reported by country of birth (Australia or other) and language spoken at home (English only or other).54 The Child Dental Health Survey is an annual clinic-based survey. Although birthplace of both child and mother used to be collected,55 this information was not included in recent reports due to increasingly limited recording by state and territory school dental services.56,57

The National Physical Activity Survey, a telephone survey conducted in 1997, 1999 and 2000, collected details of physical activities in the week before the interview.46 Results were reported by sex, age group and education but not by language group.58

Clients of Treatment Service Agencies is a 1-day census of clients (both users and friends/relatives of users) of all drug and alcohol treatment agencies across Australia. It is conducted on an irregular basis. The last time was in 2001 when 5304 clients were surveyed. Data were obtained on country of birth and language spoken at home. Published results included a broad breakdown of drug use and services received by birthplace (overseas or Australian-born) and language (non-English-speaking or English-speaking background).59

The annual national Needle and Syringe Program (NSP) Survey collects data (including country of birth and main language spoken at home by the client's parents) from clients attending selected NSPs during the designated survey week. It forms the basis of HIV and Hepatitis C surveillance among injecting drug users in Australia. The 2003 survey recruited about 2500 participants. The Australian NSP Survey National Data Report 1999–2003 gives number of respondents and HIV and HCV antibody prevalence by region of birth (Australia, other Oceania, Asia, UK and Ireland or other) and home/parental language (English or other).60

The adult component of the 1997 Survey of Mental Health and Wellbeing covered mental disorders and physical conditions, disability, health service utilisation, perceived health needs and socio-demographic characteristics (including country of birth, language usually spoken at home and year of arrival in Australia).61 In all, there were 10,641 fully participating respondents. General results were reported on prevalence of mental disorders by birthplace (Australia, other English-speaking country or other).62

The last National Nutrition Survey was conducted in 1995, and there are no plans to conduct another one.46 The 1995 Survey collected data on birthplace, language spoken at home and duration of Australian residence.63 The 2003 Survey of Disability, Ageing and Carers, another ABS survey, did not collect any ethnicity data.64

The Australian Longitudinal Study on Women's Health (also known as Women's Health Australia [WHA]) is a longitudinal population-based survey. Established in 1995, it was designed to examine the health of 40,000 Australian women over a 20-year period.55 Ethnicity data collected in the main study include birthplace, year of arrival in Australia, language other than English spoken at home and proficiency in spoken English. WHA has explored the influence of ethnicity and acculturation on a number of women's health issues, such as menopausal transitions and symptoms, contraceptive behaviour, and health service utilisation, with results reported by region of birth or broad birthplace/language grouping (non-English speaking countries or other).66,67

Growing Up in Australia is the Longitudinal Study of Australian Children. Data collection for the first wave of the study was completed in November 2004, with over 10,000 families recruited. Results
Health Consumers were released in May 2005, and the second wave of data collection commenced in September 2005. In addition to socio-demographic data (including county of birth, language spoken at home and year of arrival in Australia — recorded for the study child and all other members of the household), information is being collected on family functioning, child care and early education, and child and parental health.

**Discussion**

There are many important reasons for collecting information on ethnicity and health at a national level. High-quality data are necessary to support identification and monitoring of health issues of concern for specific groups as well as risk and protective factors, and of barriers and inequities in access to and use of services and in health outcomes. Data are also needed to assist in planning and provision of both mainstream services and ethnic-specific services to meet the needs of all Australians. Over time, the effect of duration of residence and exposure to the Australian environment and lifestyle for different cohorts of immigrants can be examined.

**The concept of “ethnicity”**

Ethnicity is an abstract and complex concept, and there is no single widely-accepted definition or measure. In Australian statistical collections, the usual practice has been to make use of one or more surrogate variables, rather than to rely on a direct question as is done in other countries such as the United Kingdom. In the collections examined, the most commonly used surrogates were country of birth and language used, typically language spoken at home or main language other than English spoken at home. Proficiency in spoken English is also critical, as it directly affects access to services and information. These are the minimum data that should be collected.

**Including religion**

The NHDD Version 12 contains six ethnicity-related variables, four of which are based on ABS standards and classifications. The NCSDD Version 3, which contains seven, also includes religious affiliation.

With growing numbers of migrants coming to Australia from the Middle East, Africa and Asia, religion is receiving increasing attention. A new question on religion has been added to the Household, Income and Labour Dynamics in Australia Survey (known as HILDA). Consideration should be given to the inclusion of this item in future health surveys.

**Collecting data on language**

One-third of the data sets record information on language used. Proficiency in spoken English is reported as part of the Aged Care Assessment Program Minimum Data Set and need for interpreter as part of the Commonwealth–State/Territory Disability Agreement National Minimum Data Set, but neither variable is included in any of the health data sets or collections. This information is critical in both supplying and planning health services and programs.

**County of birth**

Country of birth is included in the majority of health and community national minimum data sets and data set specifications that record client-centred data. A recent evaluation of the Perinatal National Minimum Data Set recommended that country of birth of mother be reported. Arguably, birthplace of father should also be reported, given the high rate of intermarriage among some birthplace groups.

All the health data collections include country of birth (or maternal country of birth in the case of Congenital Malformations, Australia). The three major death collections also collect duration of Australian residence. These collections cover the whole population, making it possible to estimate mortality, morbidity and hospitalisation rates for different countries or regions of birth, at least for the larger population groups. For rare events and diagnoses, such as suicide or birth defects, and where population numbers are small, data can be aggregated over several years.

**Limitations of surveys: size and barriers against inclusion of immigrant Australians**

With the exception of the population census, which is unique in its coverage of the entire population, surveys do not encompass everybody. The relatively small size of the sample, even in larger surveys,
necessarily limits the degree to which data can be disaggregated by cultural or ethnic group. Furthermore, people with limited English proficiency are often excluded due to the additional costs associated with translation of documents and the employment of bilingual interviewers in a range of languages, limiting the generalisability of the findings. This is of added concern given the association between lack of English proficiency and socioeconomic disadvantage. In the longitudinal children’s survey, special efforts were made to involve children and families from culturally and linguistically diverse backgrounds. Bilingual interviewers were employed and publicity materials were translated into Arabic, Chinese, Turkish and Vietnamese.73

The changing health status of immigrant Australians
Advances in information technology, coupled with the demands of performance management, have resulted in the production of more complete, accurate and timely health statistics, including an enhanced evidence base for multicultural health. We know that people born overseas generally enjoy better health than Australian-born people.47 However, this health advantage is reduced with increasing length of residence in Australia,74 an effect that may become even more evident as many of the young migrants of the 1950s and 1960s reach ages at which they are at greater risk of a range of chronic conditions.46 Moreover, certain diseases and risk factors are more prevalent among some birthplace groups,47 and some migrants continue to face special problems in dealing with the Australian health care system.54

Using information on ethnicity and health
Information about ethnicity and health is essential to addressing health inequalities. It should be employed to inform clinical and public health interventions, to monitor progress and to evaluate outcomes at a number of levels. For example, in NSW, the Chief Health Officer’s Report includes chapters on country of birth and refugee health,75 and information from Area Health Services feeds into the Ethnic Affairs Priority Statement (EAPS) Program, under which all government agencies are required to prepare detailed annual plans that document the key ethnic affairs initiatives and priorities of the agency.76

Health service managers have a responsibility to ensure that ethnicity data are of acceptable quality and that they are used appropriately. Health professionals and clerical staff are responsible for ensuring that these data are recorded fully and accurately. This requires questioning patients or their relatives directly rather than making assumptions, and the use of professional interpreters when necessary.

In the future, use of secondary sources for multicultural health research should be maximised, along with opportunities for data linkage (bringing together data from different sources in order to obtain a greater understanding of situations or individuals).77

Primary research is also needed
Research involving the collection of primary data is also needed in order to better understand the relationships identified using secondary sources, as well as the needs of specific migrant groups (whether defined by birthplace, language, religion, gender, socioeconomic status or a combination of these). For example, the growing health and other needs of elderly people belonging to Australia’s more established ethnic communities will be of increasing concern.43 While English language competence generally improves with longer residence in Australia, dementia is usually associated with reversion to the first language learnt.78 It is important that attention is paid not only to factors such as birthplace, language and religion, but also to cultural beliefs about health and illness and expectations of health and community services.79

Qualitative, quantitative and mixed methods approaches all have a place. Local or regional studies conducted in areas of high migrant concentration and state/territory-wide surveys should make use of Australian standards and classification whenever possible. This will assist with later comparison and aggregation of the data.

Conclusion
Although migrants to Australia generally enjoy good health, there are specific health conditions and risk factors that are elevated among certain ethnic
groups. People who are not proficient in English face barriers in accessing health care and are less likely than others to benefit from broad health communication strategies. Over the last decade, there have been significant developments in the collection and reporting of cultural and linguistic health data and information. Australians from culturally and linguistically diverse backgrounds are identified to a limited extent in all major health and community data sets, and health data collections and surveys. Health researchers and policy makers and planners should fully exploit these secondary data sources, as well as conducting or commissioning primary research in multicultural health. In so doing, they will be making an important contribution to the evidence base for ethnicity and health in this country and, ultimately, to promoting and maintaining the health and quality of life of all Australians.

Acknowledgements

This work was commenced while the authors were employed at the Centre for Culture and Health, University of New South Wales. Thanks are due to staff at the Australian Bureau of Statistics, the Australian Institute of Health and Welfare and other organisations who provided clarification and additional details on data collection and reporting.

Competing interests

The authors declare that they have no competing interests.

References


