



IMPROVING AND MAINTAINING
QUALITY IN ETHNICITY DATA
COLLECTIONS IN THE HEALTH
AND DISABILITY SECTOR

Opinions expressed in this report are those of the authors only and do not necessarily reflect policy advice provided by the Ministry of Health, nor represent the views of the peer reviewers or the University of Otago.

Suggested citation: Cormack D & McLeod M. (2010).

Improving and maintaining quality in ethnicity data collections in the health and disability sector.

Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

ISBN 978-0-9864682-5-4

This document is available on the website: www.ethnicity.maori.nz

**Te Rōpū Rangahau Hauora a Eru Pōmare.
Wellington School of Medicine and Health Sciences
University of Otago (Wellington)
Mein Street, Newtown
Wellington 6242**

ACKNOWLEDGEMENTS

We would like to acknowledge those whose prior work and theorising contributed to the direction of this paper. We would also like to thank those people who commented on early drafts and who provided feedback and peer review. Any errors or omissions are the authors' own.

Thank you to Julian Apatu for design of the logo, and to Flax Digital for layout, design and production.

This discussion paper is part of a series of discussion papers developed as part of an ethnicity data project funded by Te Kete Hauora at the Ministry of Health.

Kei te mihi ki a koutou katoa.

CONTENTS

Abbreviations	vii
Introduction	1
Background	2
The health and disability system in New Zealand	4
Key health sector institutions and providers	4
Key national data collections	5
Relevant legislation and regulations	6
Relevant policy and strategy requirements	6
Key health and disability sector strategies	7
Key policies and strategies outside the health and disability sector	9
Current state of ethnicity data in the health and disability sector	10
Ethnicity data in hospitals	10
The collection of ethnicity data in hospitals	11
Studies of hospital ethnicity data quality	13
Hospital ethnicity data summary	15
Ethnicity data in primary care	17
Summary of primary care ethnicity data	21
NGOs, community services and other providers	23
Accident Compensation Corporation	23
Health and disability sector workforce	25
ethnicity data	25
Key national data collections	26
National Health Index (NHI)	28
National Minimum Dataset (Hospital Events) (NMDS)	29
National Non-admitted Patient Collection (NNPAC)	30
Mental Health Information National Collection (MHINC) and Programme for the Integration of Mental Health Data (PRIMHD)	32
Maternity and Newborn Information System (MNIS)	32
Mortality Database (MORT)	33
Primary Health Organisation (PHO) Enrolment Collection	33
Pharmaceuticals Collection (Pharms) database	34
National Immunisation Register (NIR)	34
Breast Screening Aotearoa (BSA)	34
National Cervical Screening Programme (NCSP)	35
School Dental Service	35

Vital statistics – births and deaths data	36
Improving and maintaining quality datasets: barriers and facilitators	39
Health system factors influencing ethnicity data quality	39
Strategic leadership and support	39
Funding and incentivising of collection	40
Information technology systems	41
Appropriateness and/or acceptability of ethnicity codes	43
Availability of resources	44
Health care process and provider factors influencing ethnicity data quality	44
Understanding of the concept of ethnicity	44
Concern about the reaction of those they are collecting data from	45
Training	45
Co-ordination and responsibility	47
Factors at a population or community level	48
Understanding of the purpose of the data collection or meaning of ethnicity	48
Activities and initiatives to improve ethnicity data quality	49
Ministry of Health ethnicity data improvement work programme	49
Other activities to improve and maintain quality	51
Improving ethnicity data collection in health internationally	53
Improving ethnicity data in health collections in Australia	53
Improving ethnicity data in health statistics in the United States	55
Improving ethnicity data in the United Kingdom	56
Summary of activities to improve ethnicity data in health systems internationally	56
Conclusions	57
References	60
Appendix One: Key events timeline	71

ABBREVIATIONS

ACC	Accident Compensation Corporation
DHB	District Health Board
ESR	Environmental Science and Research
NGO	Non-governmental organisation
NHI	National Health Index
NIR	National Immunisation Register
NNPAC	National Non-admitted Patients Collection
NZCR	New Zealand Cancer Register
NZHIS	New Zealand Health Information Service
PHO	Primary health organisation
PHU	Public Health Unit
TPK	Te Puni Kokiri
WAVE	Working to Add Value through E-Information

INTRODUCTION

High quality ethnicity data has been an ongoing concern for the health and disability sector in New Zealand. While ethnicity data has been collected for a number of years, there have been variable levels of data completeness and quality. In addition, approaches to data collection across the sector have not always been co-ordinated or standardised.

In recent years, a number of initiatives and activities have been undertaken to improve the quality of ethnicity data in health sector datasets, including the development of the *Ethnicity data protocols for the health and disability sector* (Ministry of Health 2004a). In spite of progress, however, challenges remain in relation to both the quality and completeness of ethnicity data.

This paper is one in a series of discussion papers examining key issues in ethnicity data in relation to measuring and monitoring Māori health and ethnic inequalities. The paper will briefly outline the rationale and framework for ethnicity data collection and what is known about the current state of ethnicity data across the health and disability sector. It will then describe recent activity to improve and maintain ethnicity data in the health and disability sector, with a focus on key datasets. The paper will also consider evidence from other jurisdictions that may be useful in ongoing efforts to achieve quality ethnicity data. It is hoped that this paper will identify both barriers and facilitators to the collection of ethnicity data and to the maintenance of high quality datasets, to support data collection policies and practices into the future.

BACKGROUND

The broad rationale for ethnicity data in the health and disability sector has been well-documented, both in the domestic context and internationally. Ethnicity data allows for the measurement and monitoring of health status and outcomes for ethnic groups and health disparities between ethnic groups (Hasnain-Wynia & Baker 2006; Iqbal et al 2009; Ministry of Health 2001a; Te Rōpū Rangahau Hauora a Eru Pōmare 2000). The Ministry of Health has recognised the need for quality ethnicity data in order to measure progress in achieving Māori health gain and reducing ethnic inequalities and to monitor the effectiveness and equity of policies and programmes. Ethnicity data collected in the health and disability sector (supplemented with data collected outside the health sector, such as births and deaths data) is used routinely to produce statistics on morbidity and mortality by ethnicity. In addition, ethnicity data is a component considered in some health funding models (Lang 2002).

Ethnicity information is important in providing Māori with relevant and timely data, but is also necessary for the Crown to meet its obligations under the Treaty of Waitangi and various international conventions. For example, as a signatory to the Convention on the Elimination of all forms of Racial Discrimination, the government is required by the United Nations to report on measures that have been undertaken to realise rights protected under the Convention and on progress the state is making in this regard. Ethnicity data is central to this reporting. In relation to Treaty obligations and health sector ethnicity data, the Waitangi Tribunal found in the Napier Hospital Claim (Wai 692) that "...failure to address adequately the known problems and limitations of ethnicity data and health outcome monitoring breached the principles of *active protection* and *equity*; ..." (2001: 368). One of the recommendations made by the Tribunal in the Napier Hospital Claim Report was that "...sufficient and accurate ethnicity data be gathered to the extent needed to measure health service results for Māori" (Waitangi Tribunal 2001: 386). Ethnicity data has been collected variably within the health sector for a number of years (Brown 1983; Kilgour & Keefe 1992; McLeod et al 2000; Pasupati et al 1980; Te Rōpū Rangahau Hauora a Eru Pōmare 1996). Official health datasets have been shown historically to undercount Māori (see, for example, Ministry of Health 2001a; Shaw, Atkinson & Blakely 2009; Te Rōpū Rangahau Hauora a Eru Pōmare 2000). In addition, variable and changing quality and completeness across datasets results in difficulties calculating rates or making comparisons over time, particularly where numerators and denominators are derived from datasets with different quality. While a number of methods have been developed to enable the production of better estimates of Māori health status and of disparities in health¹, there remains a need to improve the overall quality and completeness of ethnicity datasets in the health and disability sector and to develop activities and policies to maintain high quality datasets.

1) Methods to improve estimates of Māori health are discussed more fully in Cormack & Harris (2009). Issues in measuring and monitoring Māori health and ethnic disparities: an update. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

DEFINING 'QUALITY' ETHNICITY DATA

Data quality is generally not well defined in policy and practice (CIHI 2005), and there are a range of different definitions in play (Kerr 2009). According to the Canadian Institute for Health Information, data is “generally considered to be quality data (or to have sufficient data quality) when it would be appropriate to use for the purpose in question” (2005: 1)². This is a ‘fit-for-purpose’ definition of quality, which feature relatively prominently in the data quality literature (Kerr 2009).

In the literature on quality frameworks, some broader concepts of quality are identified. In their report on developing a data quality management framework, the Ontario Ministry of Health & Long Term Care (2006) identify four key determinants of data quality, namely: timeliness, validity, reliability, and usability.

Timeliness relates to the availability of data, which is an important consideration in measuring and monitoring the health of populations.

Validity incorporates accuracy, which is the element that is most commonly understood as representing data quality (CIHI 2005). In terms of the validity of the method used to collect ethnicity data in healthcare settings, the Institute of Medicine (2009) note that self-report ethnicity data is generally considered to be the “gold standard”, and, therefore, more valid than data on ethnicity collected in other ways. As they note, “...in this hierarchy of accuracy, self-report can be understood as being of superior validity” (IOM 2009: 146). In New Zealand, self-report is also considered the method that will produce the most accurate ethnicity data, and is relatively embedded in the health sector as the preferred method (in theory, if not always in practice). Validity also includes completeness and comprehensiveness of data (MOHLTC 2006).

Reliability is a measure of consistency or the “degree to which the standards and definitions are comparable over time” (MOHLTC 2006: 13). The reliability of the standard question used to collect ethnicity data in official statistics (the census ethnicity question) has been demonstrated. The *Ethnicity data protocols for the health and disability sector* (Ministry of Health 2004a) are designed to encourage consistency by promoting the use of standard methods of data collection, capture and reporting.

Finally, usability includes elements of relevance, interpretability and accessibility (MOHLTC 2006: 13). Usability, in the context of ethnicity data, therefore includes consideration of the methods and practices relating to the classification, coding and output of ethnicity data in health.

² Work has been undertaken in New Zealand on the possible adaptation of the CIHI Data Quality Framework. For further information, see: Kerr K. (2009). The development of a data quality framework and strategy for the New Zealand Ministry of Health. Unpublished working paper.

THE HEALTH AND DISABILITY SYSTEM IN NEW ZEALAND

The New Zealand health and disability system includes a range of policy and advisory organisations as well as public and private providers. Ethnicity data is collected at a number of different sites and levels within the health and disability sector. Key institutions and providers involved in the collection and reporting of ethnicity data are outlined briefly below.

Key health sector institutions and providers

District Health Boards (DHBs) are responsible for the provision or funding of hospital services, public health services, health needs assessment, primary care services, aged care services, treatment and rehabilitation, as well as some community services and non-governmental providers. Their purpose is to "... provide or fund a specified range of health and disability services for a specified population in each district" (Ministry of Health 2008: 25). Under the *New Zealand Public Health and Disability Act 2000*, DHBs are obliged to "... reduce health outcome disparities between various population groups". In order to achieve this, there is a requirement for DHBs to carry out health needs assessments identifying the health needs of their community to feed into district strategic plans and the development of relevant services (Ministry of Health 2008: 26). Ethnicity data to provide ethnic breakdowns and information on ethnic inequalities and to feed into funding and policy decisions is, therefore, central to the work of DHBs.

Public hospitals are funded and administered by DHBs and are required to collect ethnicity data. Primary Health Organisations (PHOs), Māori health providers and other primary care and non-governmental health providers are also contracted by DHBs in each region to provide primary health care services, including general practitioner (GP) services, to an enrolled population (Ministry of Health 2008: 29). PHOs and other primary care providers are required to collect ethnicity data on their enrolled populations, which are reported periodically to the Ministry of Health.

Ethnicity data is also important in supporting DHBs to meet their range of obligations to Māori in terms of governance and in relation to building the capacity of the Māori health and disability workforce and supporting Māori providers (Ministry of Health 2008: 26).

There are a range of other private and non-governmental organisations involved in the provision of health and disability services. These include, for example, pharmacies, laboratories, Royal New Zealand Plunket Society, the Cancer Society, private hospitals and community trusts (Ministry of Health 2008). Private and NGO providers are not always required to collect ethnicity data, although some do with varying levels of completeness, quality and accessibility.

Public Health Units (PHUs) deliver health services regionally, including communicable disease control, environmental health, tobacco control, and health promotion services (Ministry of Health 2008: 30). PHUs produce regional data on notifiable diseases. At a national level,

the Institute of Environmental Science and Research (ESR) is responsible for surveillance of communicable and non-communicable diseases. Some surveillance data is reported by ethnicity, depending on the data source.

In addition to DHBs, there are also a number of Crown Entities and Crown Agents as part of the health and disability system, namely the: Crown Health Financing Agency, Health Research Council of New Zealand, Health Sponsorship Council, New Zealand Blood Service, Pharmaceutical Management Agency (Pharmac); Alcohol Advisory Council of New Zealand (ALAC); Mental Health Commission; and, the Health and Disability Commissioner (Ministry of Health 2008).

Finally, the Accident Compensation Corporation (ACC) is responsible for providing cover for accident-related injuries in New Zealand. Ethnicity data is important in terms of understanding patterns of injury, treatment and rehabilitation in New Zealand. It is also necessary in order to monitor ACC entitlements and evaluate service provision (Te Puni Kokiri 1998).

Key national data collections

The Ministry of Health maintains a number of key national collections, holding information gathered in the various settings and institutions outlined briefly above. National datasets of particular relevance in the discussion of ethnicity data include the:

- National Health Index (NHI)
- National Minimum Dataset (NMDS)
- National Mortality Collection (MORT)
- New Zealand Cancer Registry (NZCR)
- Maternal and Newborn Information System (MNIS)
- Mental Health Information National Collection (MHINC) and Programme for the Integration of Mental Health Data (PRIMHD)
- National Immunisation Register (NIR)

The National Collections and Reporting team within the Ministry of Health has operational responsibility for national collections (formerly known as the New Zealand Health Information Service (NZHIS)).

Other key datasets include registers of publicly-funded national health and disability programmes. For example:

- National Screening Unit (NSU) administered programmes for breast and cervical cancer, antenatal HIV, newborn metabolic screening and newborn hearing screening;
- School Dental Service.

RELEVANT LEGISLATION AND REGULATIONS

The collection and use of ethnicity data in the sector is related to certain legislative and regulatory requirements. Some of these are health and disability sector specific, while others are more general but relate to data used by the health and disability sector. The Ministry of Health administers a range of Acts of Parliament and Regulations. Of these, some have specific requirements for the collection and/or reporting of ethnicity data. For example, the *Human Assisted Reproductive Technology Act 2004* requires the collection of ethnicity data by fertility service providers on donors (of embryos or cells) and donor offspring (section 47(1)(g)). The *Intellectual Disability (Compulsory Care and Rehabilitation) Act 2003* specifies that in relation to cultural assessment, the “co-ordinator must try to identify the care recipient’s culture, ethnicity, language ...” (section 23(1)). Other regulations that specifically mention ethnicity data include the *Accident Insurance (Insurer Returns) Regulations 1999* and the *Health (Infections and Notifiable Diseases) Regulations 1996*.

The *New Zealand Public Health and Disability Act 2000* is a key piece of legislation for the health and disability sector. Under this Act, DHBs are required to reduce health inequalities for Māori and other population groups. Ethnicity data is necessary to be able to monitor progress towards this.

In terms of a broader legislative context, the *Statistics Act 1975* made it a statutory requirement for ethnicity data to be collected in population censuses by Statistics New Zealand. The population census is an important source of denominator data in health and disability statistics. Population estimates are produced from the census ethnic counts and are also used as population denominators.

Vital statistics data, specifically birth and death registrations, are also a significant source of data used to produce health and disability statistics and to report key health indicators such as life expectancy and mortality. Legislation in respect of ethnicity data on birth and death registrations is, therefore, important to the accuracy of health and disability statistics. The *Births, Deaths, Marriages and Relationships Registrations Act 1995* aligned the collection of ethnicity data in vital statistics with the wider statistical system approach and had a major impact on the quality and completeness of ethnic statistics generated using this data (discussed in more detail in the section on vital statistics) (Ministry of Health 2001a).

RELEVANT POLICY AND STRATEGY REQUIREMENTS

The need for quality ethnicity data is outlined in health and disability sector strategies and policy documents. Sector ethnicity data collection is also impacted by broader policies regarding the collection of ethnicity data in official statistics, particularly Statistics New Zealand statistical standards for ethnicity data.

Key health and disability sector strategies

Improving Māori health and reducing inequalities are key goals in several health and disability sector strategies. The discussion of ethnicity data in these strategies often occurs within the context of understanding and monitoring of Māori health and ethnic inequalities. In the *New Zealand Health Strategy* (Ministry of Health 2000), three specific objectives recognise the need for the collection and availability of ethnicity data. These are to:

- Ensure accessible and appropriate services for Māori
- Enable Māori communities to identify and provide for their own health needs
- Collect high-quality health information to better inform Māori policy and research and focus on health outcomes (Ministry of Health 2000: 10).

The *New Zealand Primary Care Strategy* refers to the need for providers to “be able to identify different ethnic communities within their populations, and to provide for their different needs and priorities...” (Ministry of Health 2001b: 10). The *New Zealand Disability Strategy* (Minister of Disability Issues 2001) does not make specific statements regarding ethnicity data, but makes more general reference to the need for information in the sector.

He Korowai Oranga: Māori Health Strategy 2002 specifically discusses health information requirements in relation to improving Māori health and reducing inequalities. Objective 3.4 of the Strategy identifies the need for “Māori health information to support effective service delivery, monitoring and achievement of Māori health objectives” (Ministry of Health 2002). The Strategy acknowledges that significant improvements are needed in the health and disability sector in terms of the collection of information about Māori health, including ethnicity data. Improving the “reliability of ethnicity data” is one of 10 key recommendations in the Strategy. DHBs and health service providers are both seen to have roles and responsibilities in improving ethnicity data collection in the sector (Ministry of Health 2002).

Whakatātaka Tuarua: Māori Health Action Plan (Ministry of Health 2006) outlines a number of policy areas and objectives in relation to the collection of ethnicity data. One of the plan’s priority areas is building ‘quality data and monitoring Māori health’. The action plan states that ethnicity data is critical to monitoring Māori health, mapping trends, and monitoring how the health sector is performing, and notes problems with the accuracy and completeness of ethnicity data collection. The plan also recommends that the Ministry of Health, DHBs and providers work together on improving ethnicity data as a path to improving the development and delivery of services for Māori (Ministry of Health 2006: 6).

Specific actions around ethnicity data collection are included in the plan, namely that:

- *Māori health information will be improved to support effective service delivery, monitoring and achievement of Māori health objectives*
- *DHBs, providers and the Ministry of Health will collaborate to improve the collection and accuracy of ethnicity data in order to improve planning and service delivery for Māori*
- *Ethnicity data quality and analysis will be improved for Māori with a disability through needs assessment and service co-ordination organisations (Ministry of Health 2006: 16-17).*

In terms of information strategy and policy, the WAVE (Working to Add Value through E-Information) project in the early 2000s emphasised the need for quality, accurate and timely information for the health sector. The WAVE project report made comment on the lack of ethnicity data in the sector, issues with data collection, the need for improved data quality and the need to audit ethnicity data in health information systems annually (WAVE Advisory Board to the Director-General of Health 2001). The *Health Information Strategy for New Zealand* (HIS-NZ), released in 2005, reiterated the need for timely, appropriate information for the planning, development and evaluation of health and disability services and the measurement and monitoring of the health of populations. The Strategy also noted that increasing trends to work collaboratively in the sector require data that can be shared in a timely and safe way (Health Information Strategy Steering Committee 2005: viii). Standards were highlighted as an important way of achieving this in the sense that they would promote shared understanding of health information.

The *Ethnicity Data Protocols for the Health & Disability Sector*, released in 2004, are probably the most significant sector-specific policy development. They provide guidance for the standardisation of ethnicity data collection and output across the whole sector. Prior to this, data collection had been relatively ad hoc and of variable completeness and quality (Donaldson 2002; Kilgour & Keefe 1992). While there had been moves in 1996 to align hospital ethnicity data collection with the Statistics New Zealand approach, this was not a sector-wide requirement. As part of the protocol implementation, training resources were developed and training carried out in some District Health Boards. While the uptake and implementation of the protocols is not entirely clear, and recent work suggests variation in collection practices and quality remains, the protocols do provide guidelines for a standardised, co-ordinated approach to data collection in the sector.

Key policies and strategies outside the health and disability sector

The health and disability sector is affected more generally by official policies and standards in regard to ethnicity data, specifically official statistical standards for ethnicity. In the current context, the policy likely to be of most direct relevance is the official Statistical Standard for Ethnicity 2005 (Statistics New Zealand 2005), and the Final Report of a Review of the Official Ethnicity Statistical Standard 2009 (Statistics New Zealand 2009). The 2005 standard was produced following a major review of the measurement of ethnicity by Statistics New Zealand and outlines the current official approach to the definition, collection and classification of ethnicity in official statistics. Changes introduced by the 2005 standard that are relevant to health data collections include a change to the Level 1 classification groupings³, the introduction of 'New Zealander' as a separate ethnicity code within the official classification, and the move away from prioritisation as a standard output for ethnic statistics⁴. The official statistical standard is intended to be a whole-of-government standard. Cabinet directed in 2007 that in aligning the government sector with the standard, some elements would be considered mandatory, while others would be voluntary.

The health and disability sector ethnicity protocols were produced prior to the release of the 2005 standard and, therefore, reflect the official statistical standard in place at that time. In moving to align with the 2005 standard, consultation was undertaken in 2008 within the health sector on the implications of alignment and to seek views on proposed changes. Following consultation, changes were made as part of the 2009 National Collections Annual Maintenance Project (NCAMP) to the Level 1 and Level 2 codesets used to classify ethnicity in the sector. However, in response to the lack of support identified in sector consultation, the practice for the coding of 'New Zealander'-type responses in the health sector remained the same as that in place prior to the 2005 Statistical Standard (with 'New Zealander'-type responses coded to 'New Zealand European'). Prioritisation was retained by the sector as a method for reducing the number of multiple ethnicities to three (where the number of responses given by an individual exceeded that able to be recorded) and as an option for analysis and output of data on multiple ethnicities.

In 2009, Statistics New Zealand updated the 2005 standard in relation to practice around the collection and coding of 'New Zealander'-type responses (Statistics New Zealand 2009). The 2009 recommendations include an exemption for particular data collections from coding 'New Zealander'-type responses separately, and also propose the introduction of an alternate classification whereby 'New Zealander'-type responses are grouped with European responses. This appears to be current usual practice in the health sector.

3) Level 1 refers to the first level of the Statistics New Zealand hierarchical classification system. Level 1 is the broadest, most aggregate level of codes.

4) Implications of the 2005 Statistical Standard for Ethnicity on measuring and monitoring Māori health are discussed more fully in: Cormack & Harris (2009). Issues in measuring and monitoring Māori health and ethnic disparities: an update. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

CURRENT STATE OF ETHNICITY DATA IN THE HEALTH AND DISABILITY SECTOR

Ethnicity data has been collected within the health and disability sector in New Zealand for a number of years, with varying degrees of coordination and completeness. Issues with the quality and standardisation of ethnicity data across the health sector have been previously identified (Brown 1983; Kilgour & Keefe 1992; McLeod et al 2000; Te Rōpū Rangahau Hauora a Eru Pōmare 1996). Drawing on both published and unpublished literature, this section summarises what is currently known about the quality of ethnicity data in the New Zealand health and disability sector, with a focus on datasets of particular importance for measuring and monitoring Māori health and ethnic disparities.

ETHNICITY DATA IN HOSPITALS

Information collected in hospitals provides an important dataset for the health and disability sector. Hospital data provides "... one of the few datasets available to describe morbidity in the New Zealand context" (Swan, Lillis & Simmons 2006). Hospital data also allow for the analysis of receipt of procedures and interventions and are, therefore, vital for measuring and monitoring the quality and responsiveness of the health system. It has been mandatory for public hospitals to collect ethnicity for a number of years, but ethnicity data collection remains voluntary for private hospitals. This section refers to public hospitals and services unless otherwise specified.

Information on public hospital discharges are reported to the National Minimum Dataset (NMDS). Many hospitals maintain their own databases or data collections in addition to that which they report to the Ministry of Health. Ethnicity data collected in hospital settings is, therefore, discussed both in this section and later in relation to discussion of national data collections.

The following section on hospital ethnicity data is divided into two parts: the first part outlines literature on the process of collecting ethnicity data in hospital settings; the second part maps the literature on the quality of ethnicity data collected.

The collection of ethnicity data in hospitals

Ethnicity data can be collected at a number of different points within hospitals, including both in-patient and out-patient settings, Accident and Emergency departments, maternity, mental health and community services (Donaldson 2002; Harris, Robson, Reid & Keefe 1997; Kilgour & Keefe 1992). In the past, information about hospital admissions was drawn from hospital forms typically completed by clerks at the time of admission or following admission (Brown 1983). Forms were also sometimes sent to booked patients to complete in advance of their admission or completed over the phone, and this is still the practice in some cases. In the 1980s, the ethnicity question on hospital admission forms reportedly contained three response options: Māori, Pacific Islander, and Other. However, hospitals with computerised systems were able to have a list of 12 ethnicities (Brown 1983).

In line with shifts in the broader official ethnicity data context, there were movements in the 1980s towards a self-identified cultural affiliation approach to ethnicity data collection in hospitals (Te Rōpū Rangahau Hauora a Eru Pōmare 2000). However, there was no official standard in place at this time (Donaldson 2002).

A review of collection practices in the early 1990s found variation among the then health boards in which patients ethnicity information was collected from (e.g. admitted patients, outpatients, those treated at emergency departments) and the questions used to collect the information (e.g. 'ethnicity', 'ethnic group' or 'race') (Kilgour & Keefe 1992). While no forms asked for proportion-of-descent, which was in line with a move towards cultural affiliation approaches, forms did not indicate that people could identify with multiple ethnic groups (Kilgour & Keefe 1992). In addition, forms tended to have limited response options (i.e. Māori, Pacific Islander or Other), or to simply ask for 'Race', with no response options provided (Kilgour & Keefe 1992).

Similarly, a survey of hospital data collection practice at Hutt Valley Health in early 1995 demonstrated that ethnicity data was collected in a range of ways from individuals using hospital services (Te Rōpū Rangahau Hauora a Eru Pōmare 1996). The majority of staff reported that they always filled out the ethnicity question on the hospital admission form, with many saying that this information was usually self-identified by the individual using hospital services. However, staff also reported that in some cases information was transferred from other sources or was based on the staff member's judgement. Variable practice was also identified in relation to how multiple ethnicities were recorded, with some staff asking individuals to 'self-prioritise'. Staff also reported that while the code "other" was used "by most staff for anyone without a listed ethnicity, it was also used by some for people with mixed ethnicity, for those who did not wish to answer the question or left it blank, or for those who did not wish to be described as Pakeha or European" (Te Rōpū Rangahau Hauora a Eru Pōmare 1996: 9).

In 1996, there was formal alignment of ethnicity data collection within hospitals with the Statistics New Zealand concept of ethnicity and the population census ethnicity question. Hospital ethnicity data was officially supposed to be based on self-identification by individuals, with individuals able to identify with more than one group where applicable.

The term ethnic group has a wide meaning. It is not the same as nationality, race or place of birth. Ethnic groups are ... people who have a “sense of belonging” to the group, which may not be based on birth. It is possible to belong to more than one ethnic group. At different times of their life people may wish to identify with other groups (New Zealand Health Information Service 1996).

Further reviews of hospital ethnicity data undertaken at Health Care Hawkes Bay and Mid Central Health in 1997 and Auckland Hospital in 1998 found that variation continued to exist in approaches to ethnicity data collection in different hospital settings (Harris et al 1997; He Kamaka Oranga 1998). In some units or wards patients filled out forms, while others were asked verbally. Information was sometimes transferred from another source (Harris et al 1997: 11). A review of hospital admissions forms by Statistics New Zealand in 1999 confirmed the absence of a standardised approach (Statistics New Zealand 1999).

More recent studies have also shown that variation in ethnicity data practice remains (Donaldson 2002; Latimer 2003). A 2002 study of Otago DHB found inconsistent processes for the collection of ethnicity across the different hospital departments and public health services (Donaldson 2002). Where an admission was arranged, patients were telephoned and demographic information, including ethnicity, was collected over the phone and entered directly onto the computer system (Donaldson 2002: 17). In some cases, the forms were sent out prior to admission. The Emergency Department asked patients directly or gave them a form to complete. However, the ethnicity of babies in hospital was automatically assigned to that of their mother, based on the admission form filled out in early pregnancy with the Lead Maternity Carer (LMC) (Donaldson 2002: 20).

In spite of the change in approach in 1996, there is evidence of lower reporting of multiple ethnicities on hospital administrative collections when compared with levels of reporting of multiple ethnicities in the census. In 1997/98, according to NZHIS data on public hospital discharges for Māori, 92.6% of Māori hospital discharges were recorded as sole Māori (that is, Māori was identified as their only ethnic group), 6.7% of Māori had two ethnic groups, and 0.7% had three (Statistics New Zealand 1999). In the mid-1990s, less than 2% of hospital ethnicity records contained multiple ethnicities (Statistics New Zealand 1999). While some of this could be a setting/context-effect (i.e. administrative data collection setting compared with

Census, which is self-identified in a survey setting), this is unlikely to explain the full difference. In addition, there has been evidence of higher reporting of 'Other' or 'Not stated' than would be expected to be the case based on census data (Moala 1999; Statistics New Zealand 1999). A review of ethnicity information at Capital Coast Health⁵ for the period July 1997–June 1999 found that more than half of hospital admissions had ethnicity coded as either 'other' or 'not stated' (Moala 1999). The comparable Census ethnicity proportions for the Wellington region at the time were 1% 'Other' and 4% 'Not stated'. Historically, some computer systems defaulted to 'other' (Donaldson 2002), and surveys of collection practices indicate that staff sometimes used 'Other' as a default when the information was missing or not collected (Te Rōpū Rangahau Hauora a Eru Pōmare 1996). In addition, relatively high levels of recording of 'Other' ethnicity on hospital systems may be an artefact of data collection practices prior to 1996, when most systems only captured data for the categories 'Maori', 'Pacific' and 'Other'.

Although ethnicity data collection is not mandatory in private hospitals, there is an ethnic group field on the ADF96 Private Hospital Discharge Forms. Up to three ethnic codes from the Level 2 list can be recorded for each individual.

Studies of hospital ethnicity data quality

Several audits of hospital ethnicity data have been undertaken in New Zealand, including studies in Wellington, Auckland, Hutt Valley, Hawkes Bay and Waikato. Most of these studies have compared information collected directly from individuals by researchers with that recorded on hospital admission forms and/or hospital computer systems (Table 1).

In 1980, a study of hospital data at Wellington Public Hospital found an undercount of about 30% for Māori when comparing the ethnicity data recorded on hospital admissions with that collected by the researchers directly from patients (Pasupati et al 1980). Of the 26 patients who self-identified to researchers as Māori, six were classified as non-Māori on the hospital admission form. Ethnicity data was missing for 10% of records⁶. When asked how they felt about their ethnicity data being collected, only six patients (2.6%) had any objection (Pasupati et al 1980).

In the early 1990s, a study on ethnicity data on hospital discharge statistics was undertaken as part of a broader ischaemic heart disease study (the ARCOS study) in Auckland. Researchers compared the ethnicity information on hospital admission records with that collected from study participants, based on response to the question "To what ethnic group do you belong?" Participants were able to choose from 11 response options. Where individuals identified with multiple ethnicities, they were prioritised into one ethnic group. Participants were then assigned as either Māori or non-Māori. For the period 1983–1994, 92.9% of those classified in the interview as Māori were also Māori on their admission record, compared with 99.1% of non-Māori also classified as non-Māori on the hospital record (Priest & Jackson 1994). The predictive values for Māori were 87.5% for the hospital record, and 99.5% for non-Māori. This

5) Capital Coast Health admissions considered included in-patient and out-patient admissions, consultations at Accident and Emergency and pre-admission clinics at both Wellington and Keneperu hospitals (Moala 1999).

6) The context of ethnicity data collection at the time was reflected in the survey, which asked 235 patients at Wellington Public Hospital about their 'racial classification', based on:

- i) Personal preference (self-definition)
- ii) Degree of Māori blood (traced via family tree if necessary)
- i. As entered on the Admissions form.
- ii. Interviewer's classification. (Source: Pasupati et al, 1980: 2).

study showed a misclassification of Māori ethnicity on hospital admission records at a lower level than had been demonstrated in earlier work.

Several further audits were undertaken in the 1990s⁷. A 1995 study at Hutt Valley Health asked inpatients over a 10-day period to complete the 1991 Census question on ethnicity, and to answer questions on self-prioritisation, ancestry, and whether or not they objected to being asked their ethnicity (Te Rōpū Rangahau Hauora a Eru Pōmare 1996). Self-identified ethnicity was compared with that recorded on hospital admission forms and on hospital computers. The audit found that 75% of those who identified as Māori when asked the Census question were also classified as Māori on their hospital admission form (95% CI 59.4–86.3). For non-Māori, the proportion was 94% (95% CI 90.4–96.6). The misclassification was higher on computer records for Māori, with 65% of Māori correctly classified on the hospital computer system (95% CI 50.0–79.1). For those who did not identify as Māori when asked the census question, 98% were classified as non-Māori on the computer records (95% CI 94.7%–99.0%). Adjusting for self-identified ethnicity, “the number of Māori admissions was found to be 25.7% higher than the official hospital records. The adjusted number of non-Māori admissions was 3% lower than reported” (Te Rōpū Rangahau Hauora a Eru Pōmare 1996: 14). Some of the misclassification resulted from Māori hospital admission records not being coded on the computer as Māori. A further finding of importance was the low level of respondent objection to the question, with 94% of those participating in the audit reporting that they did not object to being asked their ethnicity.

An audit was also undertaken of ethnicity data in public hospitals in the Northern Regional Health Authority in 1995/1996. In comparing ethnicity data collected in interviews with patients with that recorded on hospital systems, it was found that 10% of those identifying as Māori (n=38) were classified as another ethnicity on the hospital record (Jackson 1996 cited in Latimer 2003).

Using a similar methodology to that used in the 1995 Hutt Valley study, an audit of inpatients was carried out at Hawkes Bay Health (Napier and Hastings hospitals) in 1997. Participants were asked to complete the Census ethnicity question (in this case, the 1996 Census question). Responses were categorised as Sole Māori (those who identified Māori as their only ethnic group), Mixed Māori (those who identified Māori as one of their ethnic groups), and non-Māori (those who did not report Māori ethnicity), and then compared with hospital computer records. Misclassification was found and shown to disproportionately affect Māori. Of those who self-identified as Māori when asked, only 63% were coded as Māori (sole or mixed) on the computer records, and a further 24% (n=16) of Māori were incorrectly coded as non-Māori (95% CI 14–36). In comparison, 95% of people who did not identify as Māori were coded as non-Māori on the hospital records. Only 2% objected to being asked their ethnicity (Hadi et al 1997).

In 1998, an audit was undertaken at Auckland Hospital by the Māori Health Management

7) A study was undertaken by Bashford et al at Wellington Hospital in 1993, comparing patient self-identification with hospital admission forms and computer systems. The author was unable to locate an original copy of the study, and it is not, therefore, included in the discussion. However, other reports citing the study report that among those who self-identified as Maori, 40% were misclassified as non-Maori on the hospital admission record and 57% as non-Maori on the computer system (Harris et al 1997).

Service. The audit included a survey of staff and a survey of inpatients over a three week period. The audit found discordance between patients' self-identification on the inpatient survey and ethnicity recorded on the computer system, particularly for Māori and Pacific peoples. Of the 17 people who self-identified as Māori in the survey, only 12 were classified as Māori on the computer record (71%) (He Kamaka Oranga 1998).

More recently, a study investigating the accuracy of ethnicity data in hospitals found discordance between different datasets (Swan, Lillis & Simmons 2006). This study collected self-identified ethnicity from participants in the Waikato region as part of a questionnaire on diabetes care. Participants were asked, "Which ethnic group(s) are you?" and were able to select from six options (European, Māori, Pacific Island, South Asian, Asian and 'Other') and to identify with up to three ethnicities. Ethnicity responses were then compared with hospital record data on the Regional Diabetes Database (developed and maintained by the Regional Diabetes Service). This database only captures one ethnicity, which is reported to be selected by the patient. Prioritised ethnicity was used to compare Māori and non-Māori, while total ethnic group response was used elsewhere. Almost 98% of ethnicity in the hospital record agreed with respondent ethnicity (or with at least one of the ethnicities specified where more than one was reported). There were differences in the level of agreement by ethnic group. For Māori, there was 71% concordance, compared with 99% for non-Māori. Some of the misclassification was among people who identified Māori as one of their ethnic groups, with the computer record correctly recording their other specified ethnic group (88/99) but not Māori. Although the data collected by questionnaire was considered the 'standard' in this study, it should be noted that the question used was not the standard Census ethnicity question. Some discordance, therefore, could result from the different questions used to collect ethnicity data.

Hospital ethnicity data summary

There has been ethnicity data collection in public hospitals in New Zealand for some time, demonstrating a level of support for ethnicity data collection in this setting. There have also been significant changes in the approach to collection of ethnicity data over time, notably the move in the mid-1990s to alignment with Statistics New Zealand's approach to ethnicity data and the encouragement of standardisation. However, research over a number of years demonstrates that variation in data collection practices have been commonplace.

A number of studies undertaken in the 1980s, 1990s and more recently, show misclassification of ethnicity data in public hospitals that results in undercounting of Māori hospitalisations. They also demonstrate lower than expected numbers of people recorded as having multiple ethnicities (compared with multiple ethnic responses in the population census). Historically, there have also been high levels of missing or 'not stated' ethnicity, as well as higher numbers of 'other' ethnicities coded than would be expected based on population demographics. There is some evidence that these fields ('not stated' and 'other') have sometimes been used as a default option for data entry when the information has not actually been gathered directly from the individual.

Table 1: Summary of studies reporting findings on hospital ethnicity data quality for Māori

Author	Date*	Setting	N	Comparison	Key findings
Pasupati et al 1980	1980	Wellington Public Hospital	222	Patient self-identification (to interviewer) with hospital record	23% (6/26) of Māori who self-identified as Māori to interviewers were misclassified as non-Māori on their hospital record 9.8% of records in original sample excluded due to missing ethnicity data on hospital record
Priest & Jackson 1994	1983-1994	Auckland		Patient self-identification (to interviewer) with hospital record	Among those who self-identified to the interviewer as Māori, 7.1% were misclassified as non-Māori on their hospital record Among those who did not self-identify as Māori, 0.9% were misclassified as Māori on their hospital record The positive predictive values were 87.5% for Māori and 99.5% for non-Māori
TRRHAEP 1996	1995	Hutt Hospital		Patient self-identification with hospital admission form and hospital computer system	Among those who self-identified as Māori, 25% were misclassified on the admission form, and 34% on the computer record, as non-Māori Among those who did not self-identify as Māori, 6% were incorrectly classified as Māori on the admission form, and 2% on the computer system
Hadi et al 1997	1997	Hastings and Napier Hospitals	261	Patient self-identification (to interviewer) with hospital computer record	24% of Māori were misclassified as non-Māori 1% of non-Māori were misclassified as Māori
He Kamaka Oranga 1998	1998	Auckland Hospital	179	Patient self-identification (on survey) with computer system	29% (5/17) of those who self-identified as Māori were classified as non-Māori on the computer record
Swan et al 2006	2003	Waikato		Patient self-identification with hospital record data on Regional Diabetes Database	29% of Māori were misclassified on their hospital computer records

Note: Studies where original reports could not be sourced are not included in the table (but may be discussed in the text).

*Date refers to date of study, not date of publication

ETHNICITY DATA IN PRIMARY CARE

Ethnicity data has not historically been routinely collected in primary care. This has contributed to major gaps in terms of understanding Māori health experiences, health care utilisation, patterns of treatment and intervention, and outcomes. In recent years, however, ethnicity data has been collected more comprehensively in the primary care setting. In part this increased collection is in response to changes in the funding of primary health care in New Zealand in the early 2000s, and specifically the introduction of the Population-Based Funding Formula, of which ethnicity was a variable. Primary care ethnicity data is now often collected as part of the patient registration process with a primary health care provider.

Primary Healthcare Organisations (PHOs) are required to collect 'consistent and accurate ethnicity data'. The PHO requirements document states that:

- PHO providers must ask people when they are enrolling to self-identify their ethnicity according to the Statistics New Zealand census 2001 ethnicity question.
- PHOs are required to follow the collection process outlined in the 'Ethnicity Data Protocols for the Health and Disability Sector'
- PHOs must submit ethnicity data using Statistics New Zealand classifications, as outlined in the 'Ethnicity Data Protocols for the Health and Disability Sector'
- DHBs will monitor the PHO data collection process and the quality of the ethnicity data information for consistency and accuracy
- DHBs will monitor the PHOs' enrolment communication processes to agree on effective ways to meet the needs of iwi and Māori (Ministry of Health 2004b: 2).

As noted with regard to hospitals however, there is evidence of quality issues with primary care ethnicity data (Bramley & Latimer 2007; HURA 2006).

Studies undertaken in the late 1990s and early 2000s, before the introduction of population based funding, demonstrate significant variation in levels of ethnicity data coverage between general practices, with it not uncommon for practices to collect no ethnicity data at all. In a survey of 12 general practices in the Southern Regional Health Authority in the late 1990s, it was found that five practices collected no ethnicity data, while coverage at the other seven practices ranged from 0.2% to 21.7% of patient records, with no standard coding system across practices (Tilyard, Munroe, Walker & Dovey 1998). Around the same time, data from 14 Western Bay of Plenty general practices was reviewed for completeness (as part of the development of a 'non-invasive method' to review data quality in general practice). Ethnicity was one variable considered as a data quality indicator, with assessment of data quality based on completeness (i.e. percentage of records with a recorded ethnicity). For this method, practices were assigned a pass/fail, with 100% as the threshold. All of the practices failed with respect to ethnicity data recording (Gribben, Coster, Pringle & Simon 2001), although this would not be atypical at this point in time, as there was no routine collection of ethnicity data in general practice in the late 1990s.

In 1998, a survey of current practice in general practice was undertaken in the Greater Wellington Region investigating the attitudes of staff, patients, and data users to ethnicity data (McLeod et al 2000). About half of the 62 practices surveyed (46.8%) did not collect any ethnicity data, while 25% collected data from some patients or data was collected by some of the GPs in a given practice. Ethnicity data was “routinely collected” in 29% of practices, however, practices varied in their approaches and level of standardisation. In discussing possible benefits and risks, practice staff highlighted the role of ethnicity in providing care to individual patients as a benefit and discussed the need for the data to be relevant to service provision (McLeod et al 2000). Practices also raised some concerns around perceived risks in terms of improper use of the data, causing offence to patients, and the costs of data collection. Of those patients who completed questionnaires in waiting rooms (n=55) as part of the study, 85% indicated they would “not mind being asked their ethnicity at their general practice” (McLeod et al 2000).

A study of 42 general practices from around New Zealand found that ethnicity data was not routinely collected by all practices and that there was a high degree of variation in ethnicity coverage between practices in the early 2000s (Hall, Tomlin, Martin & Tilyard 2002)⁸. In relation to the coverage of ethnicity data within individual practices, there was a wide range – from 0.7% to 99.4%. The mean percentage of patients with ethnicity recorded was 34.9%, and the median 26.9%. Eight of the 42 practices had ethnicity data recorded for more than 80% of their enrolled (registered) patient population (Hall et al 2002).

This study was undertaken as the primary care environment was changing in New Zealand, and population based funding was being introduced. The uneven coverage of ethnicity data between practices may in part have reflected the funding environment at the time, whereby Access funded practices had an incentive to prioritise the collection of ethnicity data in comparison to Interim funded practices. For the period 2004–2005, enrolees at Interim practices were more likely to have ethnicity data missing, although this decreased over time from 18% in the second quarter of 2004 to 12% in the first quarter of 2005 (Langton & Crampton 2008).

A 2001 study involving 25 Wellington general practices showed variation in ethnicity data coverage ranging from below 10% to over 90% of all patients (HURA 2006). A comparison was made between practice-collected ethnicity (prioritised) with NMDS via NHI, and NZHIS data (prioritised) to look at levels of agreement and disagreement. For the practice records with ethnicity data recorded, there was 47% agreement with that recorded on the NHI. An additional 30% of practice records had no corresponding ethnicity data on the NHI, and the remaining 23% had a different ethnicity recorded. Comparing practice ethnicity with ethnicity from NZHIS, there was also discordance, and this differed by ethnic group (47.8% Māori, 54.4% Pacific,

8) The study looked at data from 42 practice management systems (PMS) from practices which were part of the RNZCGP Computer Research Group for the Dunedin Research Unit. The practices were from around New Zealand and their combined data represented a registered patient population of 258,249 individuals (Hall et al 2002).

39.3% Asia, 44.9% Other, 46.6% European matched with NHI). The most significant difference between NZHIS ethnicity and practice-collected ethnicity was in terms of the 'Other' response category, with only 1% of registered patients classified as 'Other' on practice databases, compared with 21% on NZHIS data (0.3% of non-registered patients at practice and 7.7% on NZHIS data respectively) (HURA 2006). Similarly, a larger proportion of practice patients were recorded as 'Other' in NHI (21%) than in the practice records. Most of those coded as 'Other' on the NHI (over 75%) were coded as European on the practice-collected dataset. A similar study, undertaken at around the same time (2001) used a national sample of GP records and found that for those with an ethnicity identified on the practice records, there was 84% agreement with ethnicity recorded on NMDS (Tomlin & Hall 2004).

The proportion of people registered with a general practice who do not have ethnicity recorded appears to have reduced over time. Of those enrolled in Access PHOs in the second quarter of 2004, 5% did not have ethnicity data recorded. This decreased to 4% in the first quarter of 2005 (Langton & Crampton 2008). Another study looked at ethnicity field completeness between 2000–2006 for patients belonging to Waikato Primary Health PHO who had received a Get Checked Diabetes review, and identified 95% completion of the ethnicity field (Joshy, Lawrenson & Simmons 2008). A more recent estimate of ethnicity field completeness in five Canterbury PHOs showed variation between PHOs, but overall only 3.4% of enrolled patients had ethnicity not stated (Malcolm 2010).

There has been an increased uptake of decision support tools within primary care since the early 2000s. Differences in ethnicity classification between different databases has important implications for the care provided to individual patients (Riddell et al 2008), as well as resulting in different numbers of calculated cardiovascular-related hospital admissions by ethnicity, depending on which database is used (Marshall, Zhang, Broad & Wells 2007). Marshall et al undertook a study of the period 2002–2005, using ethnicity recorded on NHI to investigate the level of agreement with the ethnicity recorded on the PREDICT database (a web-based cardiovascular risk decision support tool used in primary care). The study found that there was disagreement between the two databases for 7.4% of records (excluding those who had missing ethnicity in either dataset). The PREDICT database had higher levels of individuals recorded as Māori, Pacific, Asian and Indian than the NHI database.

There have been several studies looking at the coverage and quality of ethnicity data within primary care in recent years. One such study compared the ethnicity data on a patient's PREDICT record with patient self-identified ethnicity. The ethnicity data on PREDICT is transferred automatically from the ethnicity on a practice's patient management system (PMS), although it can also be entered or changed manually (Riddell et al 2008). A postal questionnaire was used to collect ethnicity data from patients, using the 2001 Census ethnicity question. The

study found that there was 68% agreement between the ethnicity recorded in primary care and that self-identified by respondents (at Level 2 of the Statistics New Zealand classification), however, levels of agreement varied by ethnic group. For example, there was 91% agreement between those recorded as New Zealand European in primary care and those who self-identified as New Zealand European, compared with 65% for Māori. Some of the groups with high disagreement are aggregate groups, or broad groupings including 'Not further defined' groups, and 'Other' groups. It is not entirely clear if the misclassification was individuals being recorded as another specific ethnic group or being classified in a broader aggregate grouping. Of those individuals who self-identified as Māori, 87.3% were also recorded as Māori on their primary care record.

An audit was also undertaken of PREDICT, looking at all Māori patient records and a random sample of non-Māori records from three practices, comparing the data on the MedTech Practice Management System with that on PREDICT. In their audit of 300 patients, none had a second or third ethnicity recorded (Riddell et al 2008: 15). There were seven ethnicity disagreements – five where the patient was recorded as being Māori in PREDICT versus European in MedTech, and two as Pacific in PREDICT and Asian or European in MedTech. As noted above, this misclassification has impacts for the risk assessment made with PREDICT.

In 2005, a period shortly after the introduction of PHOs, a study of 18 ProCare practices was undertaken. Overall, 1.7% and 0.2% had a second and a third ethnicity recorded respectively, much lower than the 8.2% seen in the 2001 Census for the Auckland population. The accuracy of the ethnicity recorded on MedTech was measured by comparing it with ethnicity reported by a random selection of patients in a postal survey. There was agreement for 67% of records (at Level 2 ethnicity), although this did vary by ethnic group (Māori 64.9%, Pacific 77.5%, Asian 60.0% and New Zealand European 90.4%) (Lindsay 2005). This study also demonstrated differences in the level of completeness of ethnicity data according to the enrolment status of individual patients. Patients enrolled with a PHO were more likely to have their ethnicity recorded than those who visited a practice as a casual patient (Lindsay 2005). This added to evidence from a previous study that showed higher levels of ethnicity completion for patients registered with a practice (Tomlin & Martin 2003).

Bramley & Latimer (2007) undertook an assessment of the accuracy of PHO ethnicity data in Waitemata, using a comparison of PHO register data with the data of children on the National Immunisation Register (NIR) collected in 2005. The ethnicity data on the NIR was collected as part of the National Meningococcal B Vaccination Campaign, which required parents to complete an enrolment form that included the Census ethnicity question. A comparison was made between Level 2 codes, with multiple ethnicities prioritised (in the order of Māori, Pacific, Asian, Other and NZ European). The study found that the NIR had a lower number of children

with no ethnic group recorded (1.9%), compared with the PHO register which had 12.6% of children recorded with unknown ethnicity. Misclassification of ethnicity varied by ethnic group. There was agreement for 62.9% of Māori records, but 23.3% of those who were identified as Māori on the NIR were recorded as European on the PHO register, with a further 3.7% as Pacific, 0.4% as Asian, and 9.6% as unknown. For Pacific, there was 77.2% agreement between prioritised ethnicity on the NIR and PHO register, Asian 81.4%, European 83.2% and MELAA 61.3%. Most MELAA disagreement was coded as European on practice (18.5%) or as unknown (23.8%).

Two recent studies (Malcolm & Barnett 2009a; 2009b) compared ethnicity recorded for the Partnership Health PHO registered population in Canterbury to ethnicity recorded on hospital discharge records. The study found that there were differences between the two datasets in the proportions of patients recorded as Māori, with 6.1% and 7.0% recorded as Māori on the hospital data and PHO records respectively (Malcolm & Barnett 2009a). An extension of this study (one extra practice and 6 months of additional data) compared PHO ethnicity data to NHI ethnicity data (Malcolm & Barnett 2009b). It was found that the PHO had recorded higher percentages of European (79.3% vs 72.9%), Māori (5.9% vs 4.0%), Pacific (2.1% vs 1.8%) and Asian (4.9% vs 3.6%) ethnicities than the NHI, and had a lower percentage of missing ethnicity data (7.2% vs 17.2%). Variations in data completeness across PHO practices remained.

Summary of primary care ethnicity data

Ethnicity data is important in primary care. This is in terms of contributing to our understanding of the broad health experiences and outcomes for Māori in New Zealand, and in monitoring equitable outcomes, but also because ethnicity is related to payments and funding. There has been relatively rapid movement from it being standard practice to not collect ethnicity data to comprehensive ethnicity data collection. Coverage has increased in recent years. However, it is apparent that issues remain with data quality. There appears to be misclassification occurring, with this disproportionately affecting some ethnic groups, often Māori. Some of this discordance will be the result of non-standard approaches to ethnicity data collection. It is unclear how embedded the ethnicity data protocols are in the primary care context. The *Sector Assistance Information Pack – the Capitation Enrolment & Register Handling Guidelines for Practices & PHOs/MSOs Handbook* (Ministry of Health 2007) aims to improve the quality of patient health data, but includes a non-standard template of the ethnicity question. There have also been system issues with some PMSs not able to record multiple ethnicities, or data on multiple ethnicities not being extracted from practice registers.

Table 2: Summary of studies reporting findings on primary care ethnicity data

Author	Date	Setting	N	Comparison	Key findings
Tilyard et al 1998	1995	Southern Regional Health Authority practices	81,890 patients	Assessment of completeness of ethnicity field	Five of 12 practices recorded no ethnicity data. Coverage at the other seven practices ranged from 0.2% to 21.7% of their patient records.
Gribben et al 2001	1997-1998	Western Bay of Plenty	-	Assessment of percentage of records with ethnicity entered	The 14 practices were assigned a pass/fail, with 100% ethnicity coverage as the threshold. All of the practices failed with respect to ethnicity data recording.
McLeod et al 2000	1998	Wellington	62 practices	Practice survey	Of the 62 practices surveyed 46.8% did not collect any ethnicity data, while 25% collected data from some patients or by some GPs.
Hall et al 2002	2000	NZ general practices (with data in Dunedin Research Unit)	c.258,000 patients included	Assessment of completeness of ethnicity field	There was large variation in the recording of ethnicity between the practices in the study (0.7-99.4%). Eight of the 42 practices had ethnicity recorded for 80% of their population.
HURA 2006	2001	Wellington	-	General practice records with NHI and NZHIS ethnicity	Ethnicity data coverage varied by practice from less than 10% to greater than 90%. There was 47% agreement between ethnicity recorded in the practice and that on NHI. Comparing practice data with NZHIS data, there was variation in levels of discordance by ethnicity, with 47.8% of ethnicity records matching for Māori, 54.5% or Pacific 39.3% for Asian and 46.6% for NZ European.
Marshall et al 2007	2002-2005	Patient records in PREDICT	18,239 patients - 15,902 excl missing	PREDICT ethnicity with NHI	9.9% of sample excluded from analysis due to missing ethnicity in either database. The level of disagreement in ethnicity was measured at 7.4%. The PREDICT database had a higher level of individuals recorded as Māori, Pacific, Asian and Indian than the NHI database
Riddell et al 2008	2004-2005		665 patients	Patient self-identification (on postal questionnaire) with PREDICT record	Agreement between self-identified ethnicity and PREDICT varied by ethnicity. There was 91% agreement between those recorded as New Zealand European in primary care and those who self-identified as New Zealand European, compared with 65% for Māori.
Lindsay 2005	2005	ProCare practices, Auckland	668 completed questionnaires	MedTECH ethnicity with self-identified (on questionnaire).	Overall misclassification 27.5%. Variations by ethnic groups, Māori records matched for 64.9% Pacific 77.5%, Asian 60.0% and New Zealand European 90.4%
Bramley and Latimer 2007	2005	Waitemata	57,612 patients	Parental identification on NIR (based on consent form using census question) with data on PHO register	Missing ethnicity data was higher on PHO register (12.6%) than for NIR (1.9%). Based upon ethnicity collected on NIR, misclassification of PHO ethnicity data varied by ethnic group; Māori (37%), Pacific (23%), Asian (20%) and European (17%).
Malcolm & Barnett 2009a, 2009b	2009	Partnership Health PHO	a.476,042 b.354,383	a. PHO records with hospital discharge ethnicity b. PHO records with hospital discharge ethnicity and NHI	These studies demonstrated differences between the ethnicity recorded on the PHO, and that recorded in NMDS and NHI. The size of the differences varied by ethnic group. There were also differences in data completeness by PHO.

Note: Date refers to date of study, not date of publication

NGOs, COMMUNITY SERVICES AND OTHER PROVIDERS

In addition to key settings discussed above, there are many other sites where ethnicity data may be collected within the broader health and disability sector. This includes NGO, community-based and private providers, such as hospices, physiotherapists, dentists, rest homes, rehabilitation and disability support services, and many more. Some of these providers are involved in the provision of publicly-funded health and disability services and may, therefore, collect some ethnicity data as part of reporting requirements to funders. There is limited information available on the quality and completeness of ethnicity data in the broader health and disability sector. In their study of Māori health statistics, Kilgour & Keefe (1992) identified that some community services, such as alcohol and drug services, community psychiatry, nursing care, and hearing testing, collected ethnicity data in some area health boards, with a range of different ethnicity questions in use. According to a Ministry of Health working paper, although there was a field for ethnicity data in the Client Claims Processing System for disability support services, it was not a mandatory field and had 60% coverage (although matching was undertaken with NHI to improve this) (Kerr 2003). Ministry-funded Needs Assessment and Service Co-ordination (NASC) services record ethnicity data, although there is likely to be some variation in how the data is collected from clients. The Socrates system, used to record NASC data, is able to code ethnicity at Level 4 and up to six ethnicities for each individual (personal communication, National Health Board 2010). From what is known about the health and disability sector in general, it is likely that there is considerable variation among NGOs, community-based services, and private providers in terms of the level of standardisation of data collection practices, and the quality and completeness of datasets.

ACCIDENT COMPENSATION CORPORATION

Ethnicity data collected by the Accident Compensation Corporation (ACC) is used to produce injury statistics by ethnicity and to monitor access to ACC services. However, ethnicity data collection by ACC has not always been complete or readily-available. ACC did not collect ethnicity data until 1997, at which time it was collected and recorded for some claim types only (Te Puni Kokiri 1998). In their 1997 review of ACC service delivery to Māori, Te Puni Kokiri noted that although an ethnicity question was added to entitlement claim forms (ARC1) in 1997, not all branches recorded the information on the Corporation's computer system (Te Puni Kokiri 1998). In addition, entitlement claims only represented approximately 10% of all claims at that time. Ethnicity was not collected for other claims, contracted services, or in the ACC satisfaction survey. The review report made strong recommendations about the need for ACC to collect ethnicity data on all claims and services, and noted that the lack of data meant that ACC was "severely limited in its ability to plan, deliver and evaluate effective services for Māori" (Te Puni Kokiri 1998: 39). In addition, ACC's computer system did not allow for more than one ethnic group to be recorded for each claimant (Te Puni Kokiri 1998). It is unclear how decisions were made by personnel entering the data about which ethnic group to input if multiple ethnic groups were recorded on a claim form.

In an accompanying report on the views of Māori claimants and providers of ACC service delivery to Māori (Te Puni Kokiri 1997), Māori participants were asked about ACC collection of ethnicity data. While some concern was expressed about how the data may be used by ACC, other views included that it might assist Māori workforce development within ACC (Te Puni Kokiri 1997: 19).

In a follow-up to the 1997 review, Te Puni Kokiri reported that ACC had increased efforts to collect ethnicity data for all claimants (Te Puni Kokiri 2001). The follow-up report noted that the complete collection (100%) of ethnicity data was designated a Key Performance Indicator for all branches of ACC, and reported that ACC estimates were that data were coded “accurately” for 90% of claims. For 7% of data, individuals indicated they preferred not to state their ethnicity (this is an option on the form), and 3% of data had no information provided on ethnicity (Te Puni Kokiri 2001: 23). According to the report, ACC was using the Level 2 classification categories from the 1996 Statistical Standard for Ethnicity (with an amendment to aggregate the “Middle Eastern”, “Latin American/Hispanic”, and “African (or cultural group of African origin)” categories in to one “Other” category).

The ethnicity question currently on the ACC45 claim form appears to continue to use the Level 2 classification codes as response categories for the ethnicity question. Thus, although the categories are based on the Statistics New Zealand classification, the ethnicity question itself is non-standard. It differs from the census ethnicity question in the wording, response categories and layout. Of note, the wording of the question may encourage people to identify with only one ethnic group, and a ‘prefer not to state’ response code is also included.

Figure 1: Ethnicity question from the ACC45 claim form

What is your ethnic background? *This information is collected for statistical reasons only, to help ACC develop services that are culturally appropriate.*

<input type="radio"/> NZ European/Pakeha	<input type="radio"/> Cook Island Maori	<input type="radio"/> Fijian	<input type="radio"/> Indian	<input type="radio"/> Samoan	<input type="radio"/> Other ethnic group – please specify
<input type="radio"/> Other European	<input type="radio"/> Tongan	<input type="radio"/> Other Pacific	<input type="radio"/> Other Asian	<input type="radio"/> Tokelauan	<input type="text"/>
<input type="radio"/> NZ Maori	<input type="radio"/> Niuean	<input type="radio"/> South East Asian	<input type="radio"/> Chinese	<input type="radio"/> I'd prefer not to say	

Source: <http://www.acc.co.nz/for-providers/lodge-and-manage-claims/PRV00029>

HEALTH AND DISABILITY SECTOR WORKFORCE ETHNICITY DATA

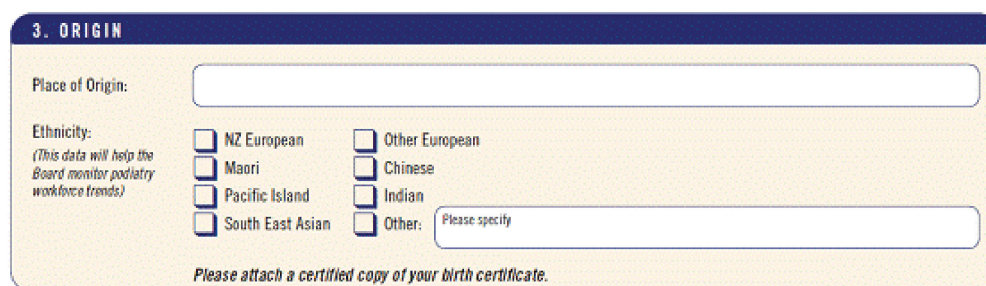
Some ethnicity data is collected on the health workforce in New Zealand. For a number of professions, this data is collected along with other demographic data at time of registration with the relevant professional organisation or through Annual Practising Certificates (APCs). This data was not collected until 1991 when an ethnicity question was included with the annual Workforce Survey accompanying the Annual Practising Certificate (Kilgour & Keefe 1992).

There appear to be a range of questions currently in use to collect ethnicity data across different professional bodies (Figures 2 & 3).

Figure 2: Example of ethnicity question from registration form for professional body

Date of birth:	Country of Birth/Ethnicity:
----------------	-----------------------------

Figure 3: Example of ethnicity question from registration form for professional body



3. ORIGIN

Place of Origin:

Ethnicity:
(This data will help the Board monitor podiatry workforce trends)

NZ European Other European
 Maori Chinese
 Pacific Island Indian
 South East Asian Other: Please specify

Please attach a certified copy of your birth certificate.

In addition, there appears to be variation across professional organisations in the ethnicity question asked in the annual workforce survey. The variability of questions between health professional organisations, as well as the absence of use of the recommended standard ethnicity question in some cases, will impact on the quality and comparability of data that is available on the Māori health workforce.

Public Health Units and Environmental Science and Research (ESR) ESR undertakes surveillance of various communicable and non-communicable diseases, and reports some surveillance data by ethnicity. Public Health Units (PHUs) produce regional data on notifiable diseases that is reported to the EpiSurv database maintained by ESR. Notifications to Medical Officers of Health at regional Public Health Units generally come from general practitioners who submit case report forms. The data on case report forms is then entered by a staff member at the PHU onto the EpiSurv database (Lopez, Sexton & Heffernan 2010). Ethnicity data is one of the variables collected on case report forms. Although practitioners are supposed to collect this information directly from individuals, it is unclear if this is always the case, or if it is completed by the GP or transferred from another source.

The ethnicity question included on the majority of case report forms for notification of notifiable diseases is similar to the recommended standard question outlined in the sector ethnicity protocols, with the exception of layout. Most of these forms were updated in 2007 (or later).

Figure 4: Ethnicity question from case report form for notification of tuberculosis

Ethnic group case belongs to* (tick all that apply)			
<input type="checkbox"/> NZ European	<input type="checkbox"/> Maori	<input type="checkbox"/> Samoan	<input type="checkbox"/> Cook Island Maori
<input type="checkbox"/> Niuean	<input type="checkbox"/> Chinese	<input type="checkbox"/> Indian	<input type="checkbox"/> Tongan
<input type="checkbox"/> Other (such as Dutch, Japanese, Tokelauan)		*(specify) _____	

Source: <http://www.surv.esr.cri.nz/episurv/crf.php>

In a study of giardiasis notification data for the period 1996–2000, it was found that ethnicity was unknown for 18% of all cases notified. This relatively high level of missing data made it difficult to have accurate estimates of giardiasis. It is not clear if this pattern exists for all notifiable diseases (Hoque, Hope, Scragg, Baker & Shrestha 2004). In 2009, four tuberculosis notifications (out of 300, 1.3%) had ethnicity recorded as ‘unknown’ (Lopez et al 2010).

STI clinics report data to ESR to allow for monitoring sexually transmitted diseases in New Zealand. However, only the broad categories ‘European/Pakeha’, ‘Māori’, ‘Pacific Island’, ‘Other’, and ‘Unknown’ are used to record ethnicity reported to ESR. It is not clear if there is a standard approach to collecting this information at the STI clinic level.

A recent study reporting on the coverage of routine vaccinations of 11 year olds in Counties Manukau DHB, compared ethnicity recorded on the Public Health Nurse database with that recorded at the Ministry of Education (taken from school enrolment forms). Twice as many students were categorised as ‘Other’ in the public health nurse database compared with the Ministry of Education (Loring & Curtis 2009).

KEY NATIONAL DATA COLLECTIONS

There are a number of key national health collections that hold ethnicity data, including the National Health Index (NHI) and other registers and databases maintained by the Ministry of Health. Ethnicity data is often collected during encounters with health care services and reported to the national collections. The quality and completeness of ethnicity data on these databases, therefore, reflects practice at the point of data collection and recording, as well as the interface between the collection source and the national databases, and the recording and output practices of the national databases themselves.

The Ministry of Health moved to align the collection of ethnicity for national health information systems with the Statistics New Zealand approach in the population census by providing for the collection of multiple ethnicities from 1 July 1996. Prior to this, it had only been possible to code one ethnic group in health collections. This change meant that the NHI and NMDS databases had the facility to store up to three ethnicities, with one ethnicity to be mandatory (the “Principal” ethnic group), and the other two fields optional (NZHIS 1995). Ethnicity data was recorded for national collections at Level 2 of the Statistics New Zealand classification.

Table 3: Rules for reporting ethnicity to the national collections

Field	Rules
Ethnicity_1	<ul style="list-style-type: none"> ○ Mandatory (must be populated) ○ Can be ‘Not stated’ (code 99) if ethnicity_2 and ethnicity_3 are both null
Ethnicity_2	<ul style="list-style-type: none"> ○ Optional ○ Can only be populated if Ethnicity_1 is populated ○ Cannot be the same value as Ethnicity_1 ○ Cannot be populated with ‘not stated (code 99)’
Ethnicity_3	<ul style="list-style-type: none"> ○ Optional ○ Can only be populated if Ethnicity_2 is populated ○ Cannot be the same value as Ethnicity_2 or Ethnicity_1 ○ Cannot be populated with not stated (code 99)

Source: www.nzhis.govt.nz/moh.nsf

More recently, national collections have been impacted by the move to align health sector ethnicity data with the Statistical Standard for Ethnicity 2005 (Statistics New Zealand 2005). Following consultation with the health sector in 2008, several amendments were included in the National Collections Annual Maintenance Programme (NCAMP) 2009 notification. Specifically, this included changes to ethnicity codes to align with the revised official classification. At Level 1, the code ‘Other’ (code 6) was added and the description of code 5 (‘Other Ethnic Groups’) was changed to ‘Middle Eastern/Latin American/African (MELAA)’. Four codes were added at Level 2, namely: ‘Other ethnicity’ (61), ‘Don’t know’ (94), ‘Refused to answer’ (95), and ‘Response unidentifiable’ (97). The code ‘Other’ (54) was retired (and converted to ‘97’ on the NHI).

Ethnicity data on national collections is sourced from data collected during interaction with the health system, or can be matched and loaded from ethnicity data already stored on the NHI. (Table 4).

Table 4: Source of ethnicity data on key national collections

Collection	Source of ethnicity data
National Health Index (NHI)	Hospital event, birth, GPs (registration of new NHIs), mortality (can update)
National Minimum Dataset (NMDS)	Health event (Hospital)
New Zealand Cancer Registry (NZCR)	Health event (Hospital), NHI, mortality
National Booking Reporting System (NBRS)	NHI
National Non-Admitted Patient Collection (NNPAC)	NHI
Mental Health Information National Collection (MHINC) & Program for Implementation of Mental Health Data (PRIMHD)	Health event and/or NHI
Maternity and Newborn Information System (MNIS)	Health event (submitted by LMC), hospital event
Laboratory Claims (LABS)	NHI
Pharmaceutical Claims (PHARMS)	NHI
Primary Health Organisation Enrolment Collection	Patient registration
National Immunisation Register (NIR)	Birth registration, NHI, primary care
BreastScreen Aotearoa	Programme enrolment form
National Cervical Screening Program (NCSP)	Programme enrolment form
Newborn Metabolic Screening	Enrolment card
School Dental Service (SDS)	Programme enrolment forms

Source: Modified from NZHIS Change Notification December 2008

National Health Index (NHI)

The NHI is a unique number assigned to individuals in order that they can be identified when using healthcare services in New Zealand. The NHI stores demographic information including an individual's name, address, date of birth, sex and ethnicity. Children born in New Zealand are assigned an NHI number at birth. The Ministry of Health estimate that approximately 98% of New Zealanders have an NHI number (Ministry of Health 2009a).

Ethnicity data is recorded on the NHI at Level 2 of the Statistics New Zealand official classification, and up to three ethnic groups are recorded per individual.

Historically, NHI data have primarily been collected in hospitals, but primary care and midwives can also generate new NHI numbers. Data is submitted to the NHI database electronically. In addition, data from the NHI has been used to update ethnicity data on other data sources (such as PHOs, and formerly Independent Practitioners Associations (IPAs)).

There are some documented issues with the quality of ethnicity data on the NHI. Historically, there have been higher numbers of 'Other' and 'Not stated' ethnicity recorded on the NHI collection than would be expected given population distributions of ethnicity. For example, a study of practices in the Wellington Independent Practitioner Association (WIPA) in 2001 found that out of the 110 362 patient records with an ethnicity recorded on their NHI, 55% had ethnicity recorded as either "not stated" or "other" (Love & Crampton 2001). The relatively high levels of 'Other' ethnicity on the NHI are related to an historical mapping issue resulting from migration to a new system in 1993. Prior to this, the NHI only included fields for 'Maori', 'Pacific' and 'Other'. With the system change, all records with 'Other' were migrated to the code 54 ('Other') of the Statistics New Zealand classification. If these individuals had no contact with secondary care following this, their ethnicity record would remain 'Other' (personal communication, National Collections and Reporting team, 2010). This causes issues for studies comparing NHI to ethnicity collected in other settings.

An NHI Upgrade Programme was undertaken in the mid-2000s. A number of projects were included within the NHI Upgrade Programme aimed at improving data quality, awareness and access. As part of this programme, several new data elements were added to the NHI, including a field for recording ethnicity history.

NHI is a very important data collection in terms of ethnicity data, as a number of the other databases rely on NHI ethnicity data to supplement their collections where data is missing.

National Minimum Dataset (Hospital Events) (NMDS)

The NMDS holds information on public and private hospital discharges for inpatients and day patients. Data has been submitted electronically by public hospitals since 1993. Private hospitals also submit information about publicly funded events and some surgical and non-surgical events from some hospitals.

According to the 1988 Review of Ethnic Statistics, the ethnic group codes recorded for public and private hospital discharges in the 1980s were: 'Māori', 'Cook Island Māori', 'Tokelauan', 'Samoan', 'Niuean', 'Tongan', 'Fijian (except Fijian Indian)', 'Other and mixed degree Pacific Island origin', 'Chinese', 'Indian', and 'Other'. Only one ethnic code was collected and recorded on the NMDS.

Since 1 July 1996, up to three ethnic codes have been able to be collected for each health care event and each health care user (National Health Board Business Unit 2010a). Although the NMDS Data Dictionary includes the 1996 Statistics New Zealand definition of ethnicity (i.e. the definition that is included in the 2004 protocols), the listing of Level 2 codes has been updated as part of the July 2009 NCAMP to align with the Statistical Standard for Ethnicity 2005. Where

more than three codes are reported for an individual, a prioritisation algorithm is used to reduce the number to three. The first ethnicity field (Ethnicity 1) is mandatory. There is also an extra field for recording prioritised ethnicity (based on the ethnicities recorded in the ethnicity fields).

Issues with the quality of hospital ethnicity data have been discussed in a previous section. There is evidence that there continues to be some level of undercount of Māori in the NMDS. For example, in undertaking analyses for the publication *Hauora: Māori standards of health IV* (Robson & Harris 2007), an assessment of the quality of ethnicity data for Māori on hospital discharge data was made by linking NMDS data to other datasets where the ethnicity data was considered more reliable (in this case death registrations, the Housing New Zealand Corporation tenant ethnicity database, National Children's Nutrition Survey 2002 and 2002/03 New Zealand Health Survey). For the time period 2000–2005, it was estimated that Māori hospitalisations were undercounted by between 5–15%, with the magnitude of the undercount varying by age (Harris et al 2007).

Similarly, for the publication *Tatau kahukura: Māori health chartbook 2010*, adjusters were developed for hospitalisations data by linking public hospital events to death registrations (for those people with both an admission and death in the period 2003–2006). Ethnicity data on the death registration data is assumed to be more reliable than that on the NMDS (Ministry of Health 2010a). Māori hospitalisations appear to continue to be undercounted, with the extent of the undercount again varying by age group.

National Non-admitted Patient Collection (NNPAC)

The NNPAC includes data on outpatient and emergency data events nationally. Its main purpose is for calculating Inter District Flows (IDFs), but it may also be used to “help provide information to measure health outcomes and inform decisions on funding allocations and policy” (NHBBU 2010b). NNPAC records must include a valid NHI, and ethnicity data is loaded to the NNPAC from the ethnicity data on the NHI. The NNPAC has used the 2009 Level 2 Ethnicity Codeset from 1 July 2009 that was part of the NCAMP, in order to align data coding with the official Statistics New Zealand Statistical Standard.

National Booking and Reporting System (NBRS)

The NBRS gathers data on elective services. According to the Kerr review (2003), it is reported electronically from the DHBs without ethnicity data included. However, ethnicity data could be matched from the NHI database.

New Zealand Cancer Registry

The New Zealand Cancer Registry (NZCR) records information on all new registrations of cancer in New Zealand. Notifications of a new registration are usually based on a laboratory report. Historically, the ethnicity information on cancer registrations has come from different

sources and has reflected changing approaches to ethnicity. For example, in the 1980s to early 1990s, ethnicity was recorded using the 'race' categories of 'Māori', 'Pacific Islander' or 'Other', and probably drawn from hospital notes (Shaw et al 2009). More recently, ethnicity is taken from the NMDS and/or mortality collections, and supplemented with data from NHI if no ethnicity reported on the source data. Cancer registry ethnicity data, therefore, is related to the quality and completeness of ethnicity data on NHI, NMDS and/or mortality collections.

Historically, there is evidence of undercounting of Māori registrations and missing ethnicity data on the NZCR. For example, a 2005 study estimated that cancer registrations were undercounted by approximately 17%, and ethnicity was recorded as 'not stated' for 8% of cancer registrations, for the period 1996–2001 (Cormack, Robson, Purdie, Ratima & Brown 2005). Analysis of cancer registrations on the NZCR for the period 2000–2004 were estimated to undercount Māori by about 2% to 15%, depending on age group (Cormack, Purdie & Robson 2007).

A recent study (CancerTrends) comparing self-identified ethnicity from census records for the period 1981–2004 with NZCR data showed an undercount of Māori on the NZCR that reduced over time, from about 30% in the 1981–1986 period to 15% in 2001–2004 period. There was undercounting of Pacific and Asian peoples throughout the time period, also reducing over time. In addition, it was found that the misclassification differed by age group, DHB, and cancer site. Of significance, the study also identified that a proportion of cancer registrations have missing ethnicity data – 4.9% in the 2001–2004 cohort. It has been suggested that these are individuals whose cancer is diagnosed privately, where ethnicity data collection is not mandatory (personal communication with NZHIS, cited in Shaw et al 2009).

In addition to these studies using linkage to estimate undercount and/or misclassification, there has been one audit of NZCR data, comparing data held on the NZCR with that held in patient clinical records in regional databases in Northland and Auckland (Stevens, Stevens, Kolbe & Cox 2008). This occurred as part of an audit of secondary care for lung cancer patients in these regions in 2004. Ethnicity data collected on the hospital registration form was compared with the relevant cancer registration. The audit found 3% (16 patients) of patients had discordant ethnicity between the two sources. This included six patients identified as Māori in the regional database who were identified as non-Māori (three as European, one as Pacific, and two ethnicity unstated) in the New Zealand Cancer Registry. In addition, two people classified as Pacific on the regional database were differentially classified on the NZCR, four European and four Other (Stevens et al 2008). However, the limitation of this work, as with audits undertaken in other settings, is that there is no gold standard ethnicity. The study, therefore, demonstrates discordance between data recorded on two separate databases, but it is not clear which is the more accurate. Theoretically, both data collection settings should be using the same method of data collection, i.e. self-identification with the standard question.

Of significance in discussion of data quality and completeness, the NZCR modified its approach to recording ethnicity data on the Register in 2009 in response to the issues of undercounting of particular ethnic groups. The new system uses an algorithm to assign up to three ethnicities to a cancer registration record, using ethnicity data drawn from NHI, mortality records and hospitalisation records (Ministry of Health 2010b). The algorithm has been applied to the whole dataset, from 1989 onwards. The Ministry of Health acknowledges that this method will mean that ethnicity data is likely to be more accurate for people who have more contact with the health system, because of the data sources used to assign ethnicity to a cancer registration record (Ministry of Health 2010b). In the most recent publication of cancer incidence and mortality in New Zealand, it is noted that this change in method has had an impact on rates and that data produced from the new system is not comparable with data produced under the old system. However, as the algorithm has been applied to the whole dataset from 1989 onwards, the NZCR data is comparable over time (taking into account the impact of changes to ethnicity data collection and recording practices since 1989).

Mental Health Information National Collection (MHINC) and Programme for the Integration of Mental Health Data (PRIMHD)

Historically, data for the Mental Health collection was recorded on the database using the classifications of Māori, Pacific Islander and Other. More recently, MHINC has moved to using the Level 2 Statistics New Zealand classification and collecting up to three ethnicities for each individual. However, multiple ethnicities were prioritised within the data warehouse.

PRIMHD is designed to develop a collection that combines national mental health data on service activity and outcomes in New Zealand, by integrating the Mental Health Information National Collection (MHINC) with the Mental Health Standard Measures of Assessment and Recovery Initiative (MH-SMART) (MOH 2010: PRIMHD page). PRIMHD includes data from some DHBs and NGOs who provide publicly funded mental health and addiction services from 1 July 2008 (as well as some legacy data from MHINC). In mid-2010, 17 DHBs were reporting to PRIMHD. NGOs are transitioning to reporting their data electronically to PRIMHD. Ethnicity data comes from the DHBs, based on information collected during encounters with healthcare services.

Maternity and Newborn Information System (MNIS)

There is currently work being undertaken at the Ministry of Health to develop a new system for storing maternal and newborn data (personal communication, National Collections and Reporting Team, Ministry of Health 2010). Previously, ethnicity data has been collected by Lead Maternity Carers and submitted to the MNIS system through the maternity claims database at HealthPAC. Following the birth of the newborn, data was drawn from the NMDS. Ethnicity data

in recent years has been collected using the 2001 Census ethnicity question, with up to three ethnicities able to be recorded for each woman. Ethnicity data has not always been collected for both the mother and the baby. In the past, it was not uncommon for babies to be assigned the ethnicity of their mother. In addition, the database has not always provided for the recording of multiple ethnicities. A 2003 review indicated that only one ethnicity field was available with multiple ethnicities prioritised at the time of recording ethnicity, although it had been a sector requirement for three fields since 1996 (Kerr 2003).

Mortality Database (MORT)

The MORT database records up to 10 fields from data collected by Funeral Directors via BDM registrations. Deaths are coded manually by the Mortality Team. The information is primarily taken from the Medical Certificate of Causes of Death. It includes underlying cause of death. Information from other sources such as the hospital discharge records on the National Minimum Dataset, coroners or certifying doctors is also sometimes used when further information is needed.

Primary Health Organisation (PHO) Enrolment Collection

The PHO Enrolment Collection was established in 2005. It includes codes for up to three ethnic groups, derived from ethnicity fields in individual practice registers. The codes accepted are the Level 2 Statistics New Zealand codes. In 2009, HDI undertook some analysis to determine differences in reporting of Māori ethnicity on the PHO register over time compared with the 2006/07 New Zealand Health Survey (Wright 2009). Records for adults (aged over 15 years) in the New Zealand Health Survey were anonymously and deterministically linked to NHI, and compared with PHO Enrolment Register data for the 2nd quarter of 2007, 2008, and 2009. Ethnicity in the NZHS was collected from individuals using the Statistics New Zealand Census ethnicity question, allowing for multiple ethnicities. PHO ethnicity data should be self-identified and should be collected using the standard question, although the literature suggests data collection processes are variable. The analysis found that there is an apparent undercounting of Māori in the PHO enrolment register data. Of those people who self-identified as Māori in the New Zealand Health Survey (the reference standard for the purposes of this analysis), 29.9% were not recorded as Māori on the PHO enrolment register in June 2007, dropping to 27.8% in June 2008, and to 27.7% in June 2009 (Wright 2009). In addition, there is low reporting of multiple ethnicities on the PHO database, with only 1.3% of enrollees recorded as having a multiple ethnicity in the 3rd quarter 2010 (personal communication, National Collections and Reporting, Ministry of Health 2010).

Pharmaceuticals Collection (Pharms) database

The Pharms database was established in 1992, and is jointly maintained by the Pharmaceutical Management Agency (PHARMAC) and the Ministry of Health. It contains information about the dispensing of government-subsidised medicines. It therefore, excludes information about medicines that are not subsidised, or where the part-charge paid by the individual meets the cost of medicine and no additional government subsidy is required (Horsburgh, Norris, Becket et al 2010). Ethnicity is not recorded on pharmacy dispensing databases. NHI has been included on the Pharms database since 2002 and allows the potential for linking of Pharms data with NHI to obtain ethnicity. However, there is some evidence that recording of NHI by pharmacists is not complete (Horsburgh et al 2010).

National Immunisation Register (NIR)

The NIR was established in August 2004 (although it contains some data from prior to this) and sources data from primary care providers and the Schools-Based Vaccination System (SBVS). It contains information on immunisations of children born since 2005 and all immunisation events from the 2005 Meningococcal B Immunisation Programme, and is an important tool for monitoring immunisation uptake in New Zealand.

The NIR includes information on ethnicity. For example, in the case of Hepatitis B immunisations for babies, ethnicity is completed by the Lead Maternity Carer (LMC) and consenting parent, and notified to the NIR (as well as the Medical Officer of Health and relevant GP) (see Figure 5).

Figure 5: Ethnicity question from Hepatitis B immunisation form for babies

Ethnic Group: (Tick all that apply) NZ European () Māori () Tongan () Chinese ()
Samoan () Niuean () Cook Islands Māori () Indian () Other (please specify) ()

Breast Screening Aotearoa (BSA)

Ethnicity is included on the Breast Screening Aotearoa register and is used for monitoring breast screening by ethnicity. Ethnicity is generally obtained at the time of admission of a screening participant on to the register. Admission forms are often sent to screening participants prior to attending a screening appointment. According to a study of ethnicity data at Dunedin Hospital in 2002, ethnicity data at that point in time was based on the NZHIS guidelines for collecting ethnicity data (using the 1996 Census question, but with the label

Pakeha removed from the label NZ European/Pakeha) (Donaldson 2002: 22). This study also found that in the Otago DHB in the early 2000s, if the patient was coming for mammography with symptoms or high risk, often through their general practitioner, their information was entered on the General Radiology Computer system (Galen Computer system), and then transferred to Oracare PMS. There was no space for ethnicity on Galen (Donaldson 2002: 21).

The 2010 Data Management Manual note that up to three ethnicities are collected for each screening participant, in line with the Ministry of Health protocol requirements. Statistics NZ prioritisation is to be used where more than three ethnicities are recorded for an individual (National Screening Unit 2010).

National Cervical Screening Programme (NCSP)

It is not entirely clear how ethnicity has been added on to the NCSP in the past. In 2003, an internal Ministry of Health paper noted that the NCSP register did not use the Statistics New Zealand codes, but rather recorded a “single two-character field for each woman” (Kerr 2003). In a study of ethnicity data in the Otago DHB in 2002, the question used to ask ethnicity for registration with the NCSP was: ‘With which ethnic group do you identify? Māori, European, Pacific Islander, Asian, Other’ (Donaldson 2002). Only one ethnicity was entered, and a priority recording system was used to determine the ethnicity to be coded (Māori -> Pacific Islander -> European) (Donaldson 2002: 28).

Although the ethnicity information to be entered on the NCSP register is supposed to be filled out by women, it has been reported that historically ethnicity data was sometimes completed by a general practitioner or practice nurse, with some practices using their own forms (Kerr 2003).

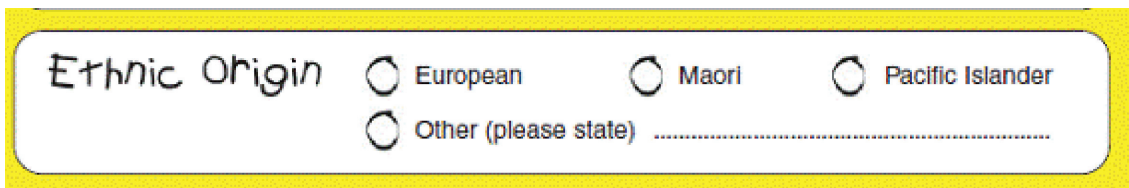
The current ethnicity question included on NCSP enrolment forms is the 2001 Census ethnicity question, and is in alignment with the ethnicity data protocols.

School Dental Service

The School Dental Service is administered through DHBs. Enrolment with the SDS occurs regionally, usually facilitated by Plunket, primary care or other well-child providers. Historically data was received from 10 DHBs and had no NHI number attached, so matching with NHI where ethnicity data was missing was not possible (Kerr 2003). It was also recorded in text form, rather than codes.

The current question on some SDS enrolment forms does not align with the Ministry of Health protocols. For example, the question and response categories in the enrolment form below are not standard, and it is not clear whether or not people can identify with more than one ethnic group, or if they do, how this is recorded on the database (Figure 6).

Figure 6: Example of ethnicity question from School Dental Service enrolment form



Ethnic Origin European Maori Pacific Islander
 Other (please state)

VITAL STATISTICS – BIRTHS AND DEATHS DATA

Ethnicity data on birth registrations has historically used the biological concept of race, asking about the degree of Māori blood of the father and mother of the newborn. This information was only collected for Māori and Pacific peoples. Following the passing of *Births, Deaths, Marriages and Relationships Registration Act* in 1995, new forms were introduced using the 1996 Census ethnicity question to collect ethnicity data on the mother, father and newborn. A question was also included asking whether or not the child was descended from a New Zealand Māori (Ministry of Health 2001a).

Similarly, death registration data prior to September 1995 only asked for degree of Māori or Pacific Island blood of the parents of the deceased person. No information on ‘race’ was recorded for deceased persons whose parents did not have Māori or Pacific Island ‘blood’. This group was usually categorised as non-Māori non-Pacific (Ministry of Health 2001a; TRRHAEP 2000). Deaths of infants less than one year of age were matched with the ethnicity on the birth registration (Kilgour and Keefe 1992).

Example: RG-28 Death Registration Form (pre September 1995)

Father: Degree of (Māori blood and tribe (if any) of father of deceased:

(Pacific island blood – state which island):

Mother: Degree of (Māori blood and tribe (if any) of mother of deceased:

(Pacific island blood – state which island):

In 1980, a survey was undertaken with funeral homes in the Wellington region to examine the collection of ‘race’ information on death registrations. The survey found that one-third of funeral directors reported not completing the ‘race’ question, citing reasons about perceived relevance or intrusiveness (Pasupati et al 1980: 6). Among funeral directors who completed the questions, a variety of approaches were reported, including making judgements based on name or appearance, or asking only some families (Pasupati et al 1980).

A study that looked at data on death registrations, comparing the data collected on the ARCOS register (from patient if nonfatal or from relative) with death registration data, found that only 50% of those who were Māori in the ARCOS study were Māori on their corresponding death registration data. This resulted in a significant 82% underestimation of Māori deaths, and a 5% overestimation of non-Māori deaths (Graham, Jackson, Beaglehole & de Boer 1989).

According to a 1996 workshop on ethnic death statistics, data collection issues related to perceived 'intrusiveness' of the question, lack of knowledge about why the data was collected, complaints about the wording/labels (notably the appropriateness of terms Pakeha or New Zealand European) (Statistics New Zealand 1996). At this time (1996), ethnicity was not specified for only 3% of death registrations (Statistics New Zealand 1996).

Example: DBM-28 Ethnic question on the Death Registration Form (1996)

Tick as many circles as you need to show which ethnic group(s) the deceased belonged to.
(1996 Census ethnicity question)

Was the deceased descended from a New Zealand Māori?

However, in 1996 there was still evidence of low reporting of multiple ethnicities. It was shown that in terms of Māori deaths registered in 1996, 91% had Māori as the only ethnicity, whereas 52% of Māori in the 1996 census were identifying only with Māori. Some of this will be an age effect, with older Māori reporting relatively fewer multiple ethnicities compared with younger Māori.

As part of a study looking at male mortality time trends, Sporle and Pearce looked at mortality data for 1994 and 1996, for Māori compared with non-Māori. They demonstrated that the percentage of male deaths recorded as Māori increased from 11.3% in 1994 (when the criteria was 50% or more Māori blood) to 21.2% in 1996 (20.9% if sole Māori only). They also found that only 1.3% of Māori deaths had multiple ethnic affiliations (Sporle & Pearce 1999).

More recently, the New Zealand Census-Mortality Study demonstrated that the alignment (general) of the death registration form in 1995 with the self-identified ethnicity question from the 1996 Census reduced the undercounting of Māori deaths in the 1996–99 period to 7%. The undercounting for Pacific ethnicity appeared to be eliminated (Blakely, Atkinson, & Fawcett 2008). In examining the undercounting in the 2001–04 period, it was found that there was 'close agreement' between the ethnicity on the mortality data and the ethnicity from the census, using both 'total' and 'prioritised' methods for classifying those people with multiple ethnic group affiliations. However, it was also found that there was less agreement when 'sole' ethnicity was

compared between the mortality and the census datasets. For Māori, the ratio was 0.86 – that is, there were higher numbers of Māori recorded as ‘sole’ Māori (those who identify Māori as their only ethnic group) on the mortality record than in the Census. According to authors of the study, this is “...because fewer mortality records were assigned two or more ethnic groups than expected compared to census data, and thus the sole Māori counts on the mortality data *exceeded* those on the census” (Blakely et al 2008: 59). The study also notes that this trend of undercounting multiple ethnicities in the mortality data relative to the census is not changing over time, and that region was associated with over-counting of ‘sole’ Māori, with the South more likely to over-count ‘sole’ Māori.

The current Notification of Death for Registration and Registration of Birth forms use the 2001 Census ethnicity question (with slightly modified wording and layout).

Ethnicity data appears to be relatively complete for birth registrations. In 2009, ethnicity was not stated for 44 children (0.07%) and 232 mothers (0.37%) on birth registration forms (Statistics New Zealand 2010).

IMPROVING AND MAINTAINING QUALITY DATASETS: BARRIERS AND FACILITATORS

Although the literature is relatively sparse, both in New Zealand and internationally, a number of barriers and facilitators to quality ethnicity data in the health sector have been identified over a period of time. Some of these are specific to particular settings and contexts, while others are more generally applicable across the sector. Recurring themes that have been identified in the literature are discussed below. For the purposes of this discussion, they are grouped into three areas, based on categories identified by Mandelblatt et al 1988 in terms of access:

- *Health system factors*: factors associated with the broader health and disability sector, as well as with the organisation and priorities of individual institutions and services
- *Health care process factors*: factors related to how services operate, how services work, and how providers and staff interact with users of healthcare
- *Patient/population-level factors*: factors associated with people using healthcare services

HEALTH SYSTEM FACTORS INFLUENCING ETHNICITY DATA QUALITY

A number of factors have been identified at the broad health system level that can act either as barriers or facilitators to the collection and maintenance of quality ethnicity data collections in the sector. In addition, many of the factors are intertwined and overlapping.

Strategic leadership and support

Strategic leadership and support has been identified as a key component in facilitating high-quality, standardised ethnicity data collections, both in New Zealand and internationally. In the domestic context, leadership was identified as an issue in consultation with the health sector over alignment with the 2005 Statistical Standard for Ethnicity (Kamira 2008). In this report, some respondents indicated that leadership and direction by the Ministry of Health was particularly relevant in relation to information systems and changes to data collection or recording practices. This was to ensure that any changes were co-ordinated and consistent across the sector. The Ministry of Health was also seen to have a leadership role in terms of relationships with IT vendors and in determining funding and implementation arrangements for alignment with the ethnicity data protocols outside of the national collections and DHBs (Kamira 2008).

In addition to the Ministry of Health, Statistics New Zealand was identified as having an overarching role in supporting health sector ethnicity data through efforts to improve general public understanding and awareness of ethnicity data collection and purposes (Kamira 2008). Statistics New Zealand recognises the role that they have in raising public awareness and understanding of official ethnicity data (Statistics New Zealand 2004).

National leadership and advocacy have also been identified in the United States as vital to improving collection of data on race, ethnicity and primary language in the health sector. A 2001 Commonwealth Fund report noted that “continued leadership by the U.S. Department of Health and Human Services, as the federal government’s principal agent in the health arena, will be pivotal to realizing the longstanding national goal of eliminating racial and ethnic disparities in health and health care” (Perot & Youdelman 2001). In the United Kingdom, the important role of senior management in supporting ethnicity data collection is noted. Guidelines on ethnic data in the UK National Health Service (NHS) highlight the need for “drive, ownership and commitment from the top of the organisation, including clinical leaders or senior service managers and senior administrators and planners” (DH/Health and Social Care Information Centre/NHS Employers 2005: 27).

Funding and incentivising of collection

The costs associated with collection of ethnicity data are raised as a challenge to having a consistent, standardised approach across the sector. Concerns about costs are often discussed within the context of changes to systems to be compliant with ‘best practice’, including changes to practice management systems, software and data collection forms (Smedley, Stith & Nelson 2002). For example, in the primary care sector where practices have traditionally operated as fee-for-service businesses, the potential cost of changes to systems to enhance ethnicity data collection has been cited as a barrier (Kelaheer et al 2010). As mentioned briefly in relationship to leadership and co-ordination, the responsibility for funding of implementation of the protocols across the whole health and disability sector, particularly in primary care and other private providers such as NGOs and community-based organisations, is seen as an issue. There have been some calls for recognition of potential increased costs and the need for central co-ordination of funding (Kamira 2008).

Funding has also been discussed in relation to the incentivising of ethnicity data collection. In situations where funding is linked directly to the availability and/or quality of ethnicity data, this can act as a facilitator for improving data collection. In New Zealand, changes to the primary care funding environment in the early 2000s by which ethnicity became a variable in the Population Based Funding Formula, provided a specific funding related incentive to collect ethnicity data. However, it is possible that this may incentivise increased coverage but may not necessarily provide improved data quality, unless incentive payments are also linked to measures of data quality.

The Institute of Medicine (2009) identified the role of incentive programs in implementing improved collection of data on race, ethnicity and language. The report recommends that where incentive programs are in existence, receipt of incentives should include consideration of the recommended collection of data on race, ethnicity and language. Linking improved data collection to incentives or accreditation provides an opportunity to encourage movement towards implementing and maintaining high quality data.

Information technology systems

Information technology (IT) systems have consistently been identified in New Zealand as important to the quality of ethnicity data collection, recording and reporting. IT systems can either facilitate or hinder high-quality, standardised ethnicity data in the sector. An example of systems acting as a barrier to quality data collection is the inability of many systems in the past to record multiple ethnicities for healthcare users, even where that data was being collected (He Kamaka Oranga 1998). This resulted in decisions about which ethnicity to record at the point of data entry that would have had downstream effects on data quality, and are likely to explain, at least in part, lower than would be expected levels of multiple ethnicities in health sector data. There have been improvements in the ability of systems to record up to three ethnicities for each individual, as is required by the protocols. However, the software does not always facilitate easy recording or extraction of this data. In addition, software does not always display all recorded ethnicities (Donaldson 2002).

Systems issues are compounded by the fact that there are a number of different IT systems being used within the broader health and disability sector, with different approaches to collecting, storing and extracting ethnicity data and variability in alignment with the protocols. A study of ethnicity data practices at Otago DHB in 2002, for example, showed that there were different systems operating in different parts of the same institution, resulting in a lack of standardisation and data comparability issues. The recently released National Health IT Plan notes the relatively large number of different information systems currently in use across the DHBs in New Zealand (National Health IT Board 2010). At a national level, the involvement of a number of IT systems and vendors can act as a barrier to sector standardisation in the absence of a co-ordinated and systematic approach to IT issues. In addition, there are likely to be data flow issues related to the existence of different systems within and across organisations (IOM 2009).

IT and software issues have also been identified as an issue internationally (Iqbal et al 2008; Kelaher et al 2010). In their report on improving the identification of Aboriginal and Torres Strait Islanders in mainstream general practice in Australia, Kelaher et al (2010) identify similar systems issues to those identified in New Zealand. Specifically, systems barriers to identification included: the ease (or not) of being able to record someone as Aboriginal; the

ability to see on software if an individual had been asked their Aboriginal status; the ability for the GP to see the status when recorded; and, the practice of software defaulting to non-Aboriginal status if no information was entered (Kelaheer et al 2010).

A range of suggested solutions for addressing IT system challenges for ethnicity data have also been identified, both in New Zealand and overseas. Having software that assists frontline staff to collect the data easily is seen as an enabler, including increasing automation where possible (Kamira 2008) or aligning the format of the data entry interface with that of the registration form (He Kamaka Oranga 1998).

Generic specifications for IT vendors have also been recommended in the report on the consultation undertaken in relation to alignment of the health sector with the Statistical Standard for Ethnicity 2005 (Kamira 2008). In their report, Kamira notes that: “Generic specifications for IT vendors should be developed and IT vendors encouraged to include them in future upgrades of their products. This will ensure a standard approach to the collection and transfer of data between systems is achieved” (Kamira 2008: 2).

Some further suggestions that may be facilitators of systems issues were identified in the 2008 consultation (Kamira 2008) with the health sector over alignment with the ethnicity data protocols:

- *incorporate a flag and prompt for staff to collect incomplete data*
- *add a timestamp so that staff are not asking people to reconfirm their data too frequently*
- *all software should use the same standard interface technique where the user is required to select from the Level 4 codes*
- *the impact of changing the data capture/storage to Level 4 is primarily in the data capture area. However, there will also be changes to the storage of the data, interface files, and minor report changes. There will be some data conversion required for compatibility*
- *information systems' user interfaces should include automated input (such as auto-complete for partial text-based input, or pick-lists)*
- *there will be more errors in input even with auto-complete as when options appears users can easily incorrectly click one of a set of similar selections.*

Source: Kamira 2008: 24-25

Appropriateness and/or acceptability of ethnicity codes

Concerns about the appropriateness and/or usefulness of codesets can act as a barrier to standardised data collection. In the United States, for example, it has been identified that data collectors have felt that “patients often did not “fit” the categories that were given” (Hasnian-Wynia & Baker 2006), and that this was a barrier to the collection of ethnicity data. Similar concerns have been raised in talking with data collectors in New Zealand in the past (Donaldson 2002).

Consultation undertaken in 2008 as part of the move to align health and disability sector datasets with the 2005 Statistical Standard for Ethnicity (Statistics New Zealand 2005) identified some concern relating to the codes in use. Essentially, these concerns were specifically related to the coding of ‘New Zealander’ responses. It was generally felt that it was not helpful in the health sector to have a visible ‘New Zealander’ code at Level 2 of the codeset. The ethnicity protocols contain the 1996 Statistics New Zealand classification, which does not include a separate code for ‘New Zealander’-type responses. However, the 2005 official standard codes ‘New Zealander’ responses separately at Level 4, within the broader ‘Other’ classification. More recently, Statistics New Zealand has recommended that some administrative datasets be exempt from the separate coding of ‘New Zealander’ responses.

Related to this are concerns about the appropriate level of recording of ethnicity data. In New Zealand, it is mandatory for health collections to record data at Level 2 of the Statistics New Zealand classification as a minimum. However, the sector is encouraged to collect at more detailed levels (Levels 3 and 4) where possible. There has been some indication from the sector that it is important that a greater level of detail is able to be collected and recorded than Level 2 (He Kamaka Oranga 1998; Latimer 2003). In some cases, more detailed data is collected, but may not be able to be recorded because of the IT systems being used.

In a review of ethnicity data at Waitemata DHB, Latimer noted there was a need for disaggregated data for the “monitoring of population health and service utilisation” (Latimer 2003). The report noted that this was particularly the case for information about Asian ethnic groups. A further issue identified was that if an individual’s ethnicity information was aggregated to Level 2 at the input stage, staff were then not able to print out details of what is on their record for the individual to check, as it will not match the information they gave. The report identifies this as an issue that should be addressed by the Ministry of Health (Latimer 2003).

This issue has also been raised in the United States, where data collectors and institutions have been concerned about whether or not the categories drawn from the census question are granular enough for the purposes for which the data is used in the health sector (Hasnian-Wynia & Baker 2006; IOM 2009).

Availability of resources

The availability of resources to support ethnicity data collection, recording and output is another factor identified as a potential facilitator or barrier in health care settings. This includes resources that outline the sector requirements and those that support staff to standardise collection and respond to questions or concerns from those they are collecting the data from. The health sector, for example, has identified ethnicity brochures for staff as a useful resource to support ethnicity data collection. In addition, there have been some requests for the availability of the standard ethnicity questions in various electronic formats to facilitate its incorporation into data collection forms (Paua Interface Ltd 2007).

HEALTH CARE PROCESS AND PROVIDER FACTORS INFLUENCING ETHNICITY DATA QUALITY

Understanding of the concept of ethnicity

A consistent issue identified at the level of data collection is understanding of the concept of ethnicity. There is some evidence from surveys and interviews with data collectors that they are not always clear of what the concept of ethnicity means in the health sector context (Donaldson 2002). This is seen to be related to the ability of data collectors to respond appropriately to questions about ethnicity data collection (He Kamaka Oranga 1998). Health sector consultation in 2008 did identify some confusion about the meaning of ethnicity among frontline staff, and the potential for this confusion to lead to resistance in asking for ethnicity (Kamira 2008).

Understanding of the purpose of the data collection and perceived importance

Research in New Zealand has demonstrated that personnel involved in ethnicity data collection in health settings are not always aware of why the question is being asked (Donaldson 2002; He Kamaka Oranga 1998). This appears to particularly have been the case among frontline staff (Moala 1999). The need for staff to understand why they are collecting ethnicity data also came through as a theme in the 2008 health sector consultation:

...some say it is important that frontline staff clearly understand the reasons for collecting ethnicity data. This enables them to explain it to people, to remove suspicion about why it is collected, and to minimise sensitivity to the way ethnicity data is coded (Kamira 2008: 18).

Internationally, staff understanding about the purpose of ethnicity data has been cited as an important barrier to data collection in the United Kingdom and the United States (Iqbal et al 2008; Iqbal et al 2009; Regenstein & Sickler 2006). In Australia, both GPs and other practice staff identified that being unsure about why the question is asked is a barrier to identification of Aboriginal and Torres Strait Islander status (Kelaher et al 2010).

Closely related to understanding the rationale behind collecting the data is how important the data is perceived to be within healthcare settings. Ethnicity data collection is not always

seen as a priority and this can impact on the quality of data collected (Donaldson 2002). For example, in acute settings such as Accident & Emergency, some have cited that ethnicity data is not always a priority, and that staff in these situations do not have time to follow up on incomplete forms (He Kamaka Oranga 1998). Others have identified that ethnicity data is a secondary concern to health issues (Hadi et al 1997).

In international literature, it has been identified that ethnicity data is not collected if it is not seen to be relevant, and is not prioritised if its importance and usefulness is not recognised (Iqbal et al 2008). In interviews about collecting data in mainstream general practice in Australia, some GPs and practice staff identified that collecting this information on Aboriginal and Torres Strait Islander status was not seen to have any benefit and was therefore not a priority (Kelaheer et al 2010). In a survey in the United Kingdom, the main reasons given for not collecting ethnicity data by health professionals were that it was not perceived as important or 'relevant to patient care' (Iqbal et al 2008: 51).

Concern about the reaction of those they are collecting data from

In addition to issues of knowledge and perceived priority, it has been indicated that the perception that people do not like being asked the question can act as a barrier (Donaldson 2002; Hadi et al 1997; He Kamaka Oranga 1998; Regenstein & Sickler 2006). This is a recurring theme in New Zealand and internationally. Data collectors have identified that they sometimes worry that patients may object to them collecting this information. This may be because they have experienced a negative reaction in the past when collecting this information (He Kamaka Oranga 1998). Among those staff who indicated that they were not comfortable collecting ethnicity in a hospital setting, some cited "their concerns over verbal abuse received from patients who thought that it was offensive to ask their ethnicity" (Hadi et al 1997).

Fear of offence has also been suggested in a recent survey of NHS data collectors in the United Kingdom (Iqbal et al 2008) and among those working in mainstream general practice in Australia (Kelaheer et al 2010).

Research in New Zealand, in both healthcare settings and in broader testing of ethnicity questions has found that generally people are not upset about answering the question (see, for example, Hadi et al 1997; Pasupati et al 1983). However, that does not necessarily mean that it does not act as a barrier for data collectors who have had a negative experience collecting this data in the past, or who perceive that they may elicit a negative response because of their own concerns about the data.

Training

Training has been identified as a facilitator of good ethnicity data collection practices. There is evidence that many people have not received training in ethnicity data collection (Hadi et al 1997; Harris et al 1997; He Kamaka Oranga 1998; Moala 1999). For example, in the mid-1990s,

staff at hospitals in the Hawkes Bay and Midcentral regions reported that while most staff knew why the data was collected, the majority were unaware of guidelines for collection and most had not had any training (TRRHAEP 1996). Similarly, in a survey with staff at Auckland Hospital in 1998, although staff had had some informal training about why ethnicity data is collected, none of the participants had received any training specifically about how to collect the data (He Kamaka Oranga 1998). Although training was generally thought to be useful, some data collectors in one site identified that they felt training would not make much difference, because of the established processes of data question or because they perceive the issues are related to the question itself, rather than the process (Donaldson 2002).

In general, however, training was perceived to be necessary in regard to overall requirements, such as information and awareness about the ethnicity data protocols, but also when changes were being made to codesets. Training priorities were identified in the mid-1990s in reviews of ethnicity data collection in hospitals in the central Regional Health Authority (Harris et al 1997). This included training on the meaning of ethnicity, the purposes of data collection, the collection of multiple ethnicities, and the importance of self-identification of ethnicity. In addition, staff were seen as requiring specific training relating to how to ask for ethnicity data, how to collect from next-of-kin, and how to address objections should they occur (Harris et al 1997: 4-5).

Many of these points are addressed in the sector ethnicity data protocols and in accompanying support and training materials. However, there has been uneven implementation of the protocols across the sector and more recent research demonstrates that training remains an issue (Kamira 2008). A range of suggestions around training were raised during the 2008 consultation with the health sector:

- *That work be undertaken to identify further training needs (including issues and barriers for staff collecting ethnicity data) and determine what methods and material will provide accessible, clear and appropriate support to staff and healthcare recipients.*
- *That training material and methods be developed to support the codeset changes and the new residual codes, and that its development includes a process of user testing.*
- *That training material and methods address (a) the technical or mechanical practices to collect and record, and (b) the knowledge of staff about what ethnicity is and why it is collected, and (c) the confidence of staff in collecting and explaining ethnicity to each other and to healthcare recipients.*
- *That the material, training and support methods reduce the burden of collection and recording, and reduce errors.*
- *That training, user interfaces and other support should follow a staff/ethnicity respondee-centric approach that aims to increase the support and compliance for collecting and recording ethnicity data.*

- *The staff/ethnicity responsee-centric process should include identifying the target audience specifically by asking: who is not collecting, who is not recording, who is resisting giving the data, why, what works?*
- *The process should include focus groups (frontline staff, non-migrant healthcare recipients, and migrant healthcare recipients).*
- *The new material and approaches should be tested by cohorts of frontline staff, non-migrant healthcare recipients, and migrant healthcare recipients.*
- *The new material and approaches should be piloted and reviewed before being finalised and deployed nation-wide.*

Source: Kamira 2008:27-28

Barriers to training in the health sector have been identified as including staff availability, high staff turnover, and lack of continuity of training (Kamira 2008). A number of improvements for training have been suggested, including the provision of an upgraded training package, integration of training into staff induction, and shared mobile trainers (Paua Interface Ltd 2007). At a PHO level, it has been suggested that training could be integrated into the data collection software and occur at the time of data entry, accompanied by suitable online support (Kamira 2008: 25).

Training has also been identified internationally as key to quality ethnicity data collection. In the guidelines for data collection in the NHS in the United Kingdom, the importance of providing training for new staff, as well as keeping existing staff up to date, is noted (DH/Health and Social Care Information Centre/NHS Employers 2005). In the United States, the Institute of Medicine identify that training of staff at all levels within healthcare organisations is central to the collection of standardised, accurate data (IOM 2009). They also note the need for this training to be tailored to address the different contexts within which data on race/ethnicity is being collected and used.

Co-ordination and responsibility

In an international report on collecting ethnicity data, it was noted that some types of health sector data are checked but that generally “Hospitals may not have the same checks in place for the collection of race and ethnicity data. In many instances, no one person is responsible for the collection of these data so it does not get done” (Hasnian-Wynia et al 2004: 16). This lack of clear responsibility or accountability for ethnicity data within institutions can be a barrier to the standardised collection and the maintenance of high quality. In New Zealand, responsibility for ethnicity data has often been associated with Māori health, although the issue is broader than this. Reviews of ethnicity data collection in New Zealand have identified uncertainty about who is responsible for ethnicity data collection (Moala 1999). The need for a designated individual with clear responsibility for ethnicity data within each organisation has been identified as important in improving ethnicity data quality (Donaldson 2002; Hadi et al 1997; Harris et al 1997).

FACTORS AT A POPULATION OR COMMUNITY LEVEL

Understanding of the purpose of the data collection or meaning of ethnicity

In discussing the collection of ethnicity data in healthcare settings, community understanding of the purpose of collecting the data has been cited as a potential barrier by data collectors (Baker et al 2005; Donaldson 2002; IOM 2009; Iqbal et al 2008).

However, there is some evidence internationally from talking to patients that they are aware of a range of different purposes behind collecting the ethnicity data (Hasnian-Wynia & Baker 2006; Iqbal et al 2008). The United States has undertaken some research on how patients feel about the collection of data on race and ethnicity. For example, patients at a General Internal Medicine clinic in Chicago were asked to self-define their race and/or ethnicity, and were asked their opinion on the importance of collecting information and how comfortable they felt within particular contexts. It was found that most participants agreed (somewhat or strongly) that information on race and ethnicity should be collected in health care settings (80%), and that the information be used to support work on health disparities (97%) and train staff to better meet needs of diverse patient populations (66%). Those who disagreed were more likely to be African American or Hispanic than white (Baker et al 2005).

Most participants also indicated that they would be comfortable reporting their information to a clerk (63%), although the levels were lower for Black participants than whites. Those with lower levels of comfort would be more comfortable identifying their race/ethnicity to a doctor or nurse. The study did find, however, that in spite of most people indicating that they were comfortable with providing the information, there was concern by some people that the information would be used in discriminatory ways, particularly among Black people (Baker 2005).

Baker et al (2005) conclude that there are concerns but they can be addressed, and that the most important thing is that providers “who collect this information must be true to their word: these data must be linked to measures of quality of care to examine disparities, initiatives to eliminate any disparities must be funded and executed, and the results must be openly shared with patients and communities” (Baker et al 2005: 899).

A small proportion of people, however, may be concerned about data purpose and possible misuse (Hasnian-Wynia & Baker 2006; Iqbal et al 2008). In the UK, it was identified that where this was a barrier, it was important to explain the purposes behind collecting the data (Iqbal et al 2008). Communication and awareness raising were, therefore, potential solutions to this barrier. In addition, it has been suggested that having resources available for healthcare users and communities, such as brochures or posters, may help them understand why the data is being collected (IOM 2009).

ACTIVITIES AND INITIATIVES TO IMPROVE ETHNICITY DATA QUALITY

Issues with the quality and completeness of ethnicity data in the health and disability sector have been known for some time. There have been a number of initiatives and attempts to improve ethnicity data collection in health settings, both in New Zealand and internationally. This includes activities aimed at increasing the coverage of ethnicity data, encouraging standardisation, and evaluating data quality. However, there is limited documentation of many of these activities and it is, therefore, difficult to undertake any systematic review. In addition, most activities have not been formally evaluated. The examples of activities to improve and maintain the quality of health sector ethnicity data in New Zealand outlined below, along with discussion of related work undertaken internationally, thus aim to provide a feel for the types and range of activities that have happened rather than to quantify health sector activity (see Appendix One for a timeline of key events related to ethnicity data in the health sector).

MINISTRY OF HEALTH ETHNICITY DATA IMPROVEMENT WORK PROGRAMME

The Ministry of Health, and formerly the Department of Health, has been progressing an ethnicity data work programme for a number of years. The need for more reliable ethnicity data has been advocated for strongly by those working in Māori health policy.

KIA PIKI TE ORA – A CASE STUDY OF AN ACTIVITY TO IMPROVE ETHNICITY DATA

The Department of Health, at the request of Te Wahanga Hauora Māori (Māori Health Policy Unit) undertook a project investigating the collection of Māori health statistics in 1991. The project incorporated a review of the literature, key informant discussions, interviews with three area health boards, and a survey questionnaire with eleven area health boards. The objectives of the project were to understand how and where information on Māori is collected and recorded, and to identify appropriate practices and strategies to improve data collection and standardise Māori health statistics (Kilgour & Keefe 1992).

The study identified the range of locations where data was collected and variation in the completeness of Māori data captured and the questions and forms used to capture the data. It also made a series of 24 recommendations about how ethnicity data could

be improved. The recommendations related to the purposes for which the data was collected, the need for consistency in definitions and approach to collecting the data, the need for alignment with Census definitions, gaps in data collection to be addressed, and mechanisms for monitoring.

A number of these recommendations have subsequently been addressed (in part or in full) and have contributed to the increased standardisation of ethnicity data collection practices in the health sector and alignment with official statistical standards.

In 2002, the Ministry undertook a project focused on improving ethnicity data in the health and disability sector. This project (known as the Ethnicity Data Improvement Project) was in part response to the recommendations of the WAVE report, which identified ethnicity data improvement as a top priority. In the early stages of the project, an internal review was undertaken of the processes for collecting and recording ethnicity data in Ministry of Health national collections (Kerr 2003). The review paper identified inconsistencies in processes associated with ethnicity data in national collections (discussed more fully in the earlier section on ethnicity data quality). The paper made several recommendations, including the need for the development of data quality indicators for national collections.

As part of the Ethnicity Data Improvement Project, stakeholders were consulted and a work group established to work with the Ministry of Health to develop protocols to improve ethnicity data across the different agencies and health care settings. The protocols aimed to encourage standardised approaches to collecting, inputting and outputting ethnicity data by providing guidance on recommended best practice. The *Ethnicity data protocols for the health and disability sector* were released in 2004.

A train-the-trainer package and supporting resources were also developed to support the implementation of the ethnicity protocols. Training was undertaken with some DHBs and resources made available via the New Zealand Health Information Service website. The focus of implementation of the protocols has been primarily on national collections and DHBs to date, although they are intended for the entire health and disability sector.

The protocols were developed and released prior to the finalisation of the 2005 Statistical Standard for Ethnicity, and therefore, reflect official policy and practice in place before the new standard. They have been updated to incorporate changes made to official statistics policy. This involved consultation with the sector and recommendations about alignment with the revised Statistical Standard. Changes to the codesets were notified as part of the 2009 NCAMP. The roll-out of these changes has been limited at this time to national collections and the changes have not been implemented across the broader health and disability sector.

In spite of the uneven alignment with the protocols across the broader sector, they represent a significant development towards standardising ethnicity data in the health and disability sector. They provide a blueprint for improving the quality, consistency and comparability of ethnicity data.

A work programme focussed on improving the quality of ethnicity data in the health sector is ongoing at the Ministry of Health. More recently, this has involved Investment Logic Mapping, the commissioning of discussion papers, and the funding of the development and piloting of an ethnicity audit tool for the primary care sector.

OTHER ACTIVITIES TO IMPROVE AND MAINTAIN QUALITY

Attempts have been made for a number of years to improve approaches to hospital ethnicity data. In 1996, the New Zealand Health Information Service released a pamphlet to provide information about the collection of ethnicity data in hospitals.

In the mid-1990s, Te Rōpū Rangahau Hauora a Eru Pōmare developed an audit tool to be used to evaluate the quality of ethnicity data in hospital settings. It was specifically focused on investigating ethnicity data issues for Māori and levels of misclassification and undercounting in hospital data. The audit tool built on work at the Centre and by other researchers over a number of years identifying issues with ethnicity data quality, accuracy and completeness, and projects undertaken by medical students to evaluate ethnicity data in public hospitals.

The audit tool involved mapping of current ethnicity data policy and practice in the health service, a survey of staff, and a survey of patients. The patient survey collected self-identified ethnicity from inpatients in a census type activity, and then compared this with data held on hospital records (paper and/or computer records) (Robson et al 1996) to identify levels of misclassification of Māori. The method has been used to audit ethnicity data in a number of hospitals, including Hutt, Napier, Hastings and Auckland hospitals in the 1990s.

MĀORI HEALTH ETHNICITY DATA IMPROVEMENT PROJECT – WAITEMATA DHB

In 2003, Waitemata DHB undertook an '*Ethnicity Data Collection Baseline Review Project*' as part of a broader *Māori Health Ethnicity Data Improvement Project*.

The review was aimed at assessing ethnicity data within the DHB and formulating recommendations for improvement.

As well as examining previous studies and audits of ethnicity data quality in New Zealand, the review project undertook analysis of data collection practices in some of its provider arm services, in order to answer a number of questions (Latimer 2003: 18):

Does the service have access to the DHB patient management system and the National Health Index?

How does the service collect ethnicity data?

How many staff are involved in collection of ethnicity data?

Are there written instructions for staff on how to collect ethnicity data?

Is there training available to staff on collection of ethnicity data?

The review looked at the forms that were used to collect ethnicity and indicators of data quality. As a result of the review, a number of recommendations were developed for the DHB. The recommendations covered a number of areas of identified need, including an implementation plan outlining operational procedures and processes, training and education among staff, the inclusion of indicators in performance reporting, and the development of a routine audit programme (Latimer 2003: 14).

Some quality improvement and audit processes are ongoing within the health sector in relation to ethnicity data. For example, in the West Coast DHB, records with 'not stated' ethnicity are routinely audited, and information on ethnicity collection is included as part of IT training and general staff training (personal communication West Coast DHB 2010). Other examples of activities include a project undertaken by the Southern Cancer Network in 2009 (Southern Cancer Network 2009) and Princess Margaret Hospital in Canterbury in 2003-2004 to assess and improve ethnicity data quality.

There has been a significant increase in recording of ethnicity data in primary care since the 2000s, when it was not uncommon for general practices to collect no ethnicity data at all. There is limited literature documenting how this increase in coverage has happened and it is difficult,

therefore, to identify the strategies that have been successful in improving coverage of ethnicity data. More generally, there has been a recent focus in quality standards in primary health care to evaluate ethnicity data and increase standardisation, including through the CORNERSTONE accreditation process for primary care.

IMPROVING ETHNICITY DATA COLLECTION IN HEALTH INTERNATIONALLY

In spite of the frequency with which ethnicity data is reported in the health literature, there is limited published information on ethnicity data collection for health statistics (Iqbal et al 2009). A recent international review noted that while articles in the health literature frequently included ethnicity as a variable, many did not provide any information on how the data was collected (Iqbal et al 2009), similar to a pattern shown in New Zealand by Thomas in the early 2000s (Thomas 2001). The review article identified that most published literature on ethnicity data collection methods was from the United States, although the majority of guidelines identified were from the United Kingdom (Iqbal et al 2009).

This section briefly maps some of the evidence from international jurisdictions on the quality of ethnicity data and activities to improve ethnicity data collection in relation to health statistics, focusing on the countries where there is the most literature, namely the United States, Australia, and the United Kingdom.

Improving ethnicity data in health collections in Australia

In recent years, there has been work undertaken to improve the identification of indigenous peoples in health care systems in Australia. The context of ethnicity data collection in Australia is different from New Zealand in that the focus is on collecting data on Aboriginal and Torres Strait Islander peoples, rather than on collecting ethnicity data for the general population. The population census in Australia does not contain a general question on ethnicity, but has a specific question on whether or not people identify as Aboriginal, Torres Strait Islander, both or neither (questions are also asked in relation to ancestry and country of birth). There is also a specific context to the collection of data on Indigenous populations in Australia, as the definitions have tended to be exclusionary in the past.

The need for standardised, complete data on Aboriginal and Torres Strait Islander status is recognised (ABS & AIHW 2005). The National Health Data Dictionary in Australia includes the standard question used in the Census to collect information on Aboriginal and/or Torres Strait Islander identification. However, there are problems in relation to administrative data that relate to "... a lack of understanding about why the Indigenous status question is being asked and why some people may be reluctant to identify under these circumstances" (ABS & AIHW 2005). This has resulted in undercounting of Aboriginal and Torres Strait Islander peoples.

Recently, a project was undertaken to improve the identification of Aboriginal and Torres Strait Islanders in mainstream general practice in Australia. The project outlined 'best practice' principles for collection of this data at point-of-care, and identified strategies that had been used to improve the collection of this data in mainstream general practice, and how mainstream general practice could be supported to improve this data (Kelaheer et al 2010). The report concluded that the overarching principle of collection of information on indigenous status should be to ensure better and more appropriate health care. Administrative strategies that could improve data at the practice level were identified from consultation with key informants and case study interviews. Strategies seen as important in relation to improving point-of-care identification of Aboriginal and Torres Strait Islander people included:

- Having improved, clear registration forms to improve completion rates of the question
- Including ethnicity data within updating of patient information
- Triaging new patients and collecting the information in a more private environment with the nurse
- Follow-up by the GP if the question is not completed on the practice form
- Having clear policy on how data will be collected in the practice and guidance on staff roles (Kelaheer et al 2010).

Kelaheer et al also identified regional strategies undertaken by General Practice Networks, such as use of promotional material. At an overall national level, the COAG 'Closing the Gap' initiative was underpinned by the need to have data on indigenous status of healthcare users (Kelaheer et al 2010). However, the report also concluded that strategies to improve identification have had "...no formal evaluation and all have been implemented on an ad hoc basis with minimal resources and guidelines as to what constitutes 'best practice'" (Kelaheer et al 2010: 53).

In 2010, the Australian Institute of Health and Welfare introduced guidelines for collecting indigenous status data in health. The *National best practice guidelines for collecting indigenous status in health data sets* outline how the question on indigenous status should be asked and how responses should be recorded (Australian Institute of Health and Welfare 2010). They also provide information on how the data can be collected in specific situations (such as from next of kin or where an individual objects) and include recommendations on ways to support best practice in the sector. Among the recommendations for best practice in the guidelines are the importance of staff training, and the need for data quality assurance and validation at the local provider, state- and territory-level, and the national-level. As part of the dissemination and implementation of the guidelines, a National Indigenous Data Improvement Support Centre (IDISC) was established to provide ongoing support, including a helpdesk.

Improving ethnicity data in health statistics in the United States

The context of ethnicity data collection in health and for health statistics is somewhat different in the United States compared with New Zealand. In the US context, the focus is on collecting race, ethnicity, and language (preferred language) in order to identify populations, monitor care and develop targeted programmes (Baker et al 2005). There is evidence of the ad-hoc nature of the collection of race, ethnicity and language within health settings in the United States, as well as some active resistance to the collection of information on race and ethnicity. In California, Proposition 54 was defeated but had proposed to ban collection of data on race and ethnicity.

A body of work has been undertaken for The Commonwealth Fund on the collection of race/ethnicity data in the United States. Interest in improving data collection practices have been driven by the need to better understand, monitor, and address health disparities (Perot & Youdelman 2001; Hasnian-Wynia, Pierce & Pittman 2004). A 2001 Commonwealth Fund report found that in spite of the policy imperatives to collect data on race, ethnicity and primary language and lack of legislative barriers, there was large variation in practices across different agencies (Perot & Youdelman 2001).

The Institute of Medicine recently released a major report on the collection of data on race, ethnicity, and primary language in health care, aiming to provide guidance on the standardised categories to be used in healthcare data collection. Recommendations on how health care institutions should collect data on race, ethnicity and primary language from individual health care users, include the promotion of the use of 'granular' data as opposed to broad, aggregate categories. The report also recommends the development and piloting of different question formats to identify the most appropriate way of collecting complete and accurate data (IOM 2009). At a national level, the development of national standards, the incorporation of standards into accreditation processes, the provision of guidance on data comparability, and incentivisation of collection where appropriate, are advocated. In discussing how to improve data collection, the report identify several 'opportunities and challenges' seen to be shared across the healthcare sector, namely:

- How to ask patients and enrollees questions about race, ethnicity, and language and communication needs;
- How to train staff to elicit this information in a respectful and efficient manner;
- How to address the discomfort of registration/admission staff (hospitals and clinics) or call center staff (health plans) about requesting this information;
- How to address potential patient or enrollee pushback respectfully; and
- How to address system-level issues, such as changes in patient registration screens and data flow (IOM 2009: 127).

The report concludes that there are significant gains to be made from standardising data collection of race, ethnicity and primary language variables across the health care system and emphasises the importance of this in order to “inform point-of-care services, application of resources, and decisions in patient-provider interactions in ways that can assist in improving overall quality and reducing disparities” (IOM 2009: 10).

Improving ethnicity data in the United Kingdom

There has been variable collection of health sector ethnicity data in the United Kingdom in the past. Ethnicity data on hospital in-patients has been mandatory since the mid-1990s, with a number of other national collections recording the information since 2001 (DH/Health and Social Care Information Centre/NHS Employers 2005). Movements have been made in the United Kingdom in recent years to standardise collection of ethnicity in health and social services. In 2001, the Department of Health and the NHS moved to using a new set of 16 standard codes, designed to align health collection with the population census but also incorporate local needs. Some guidance material was produced to support the move.

In 2005, the Department of Health, Health and Social Care Information Centre, and NHS Employers published *A practical guide to ethnic monitoring in the NHI and social care*. The guidelines outlined information about how ethnic group data should be collected and used. This included discussion of the appropriate approach to collection, the standard ethnic group codes to be used, the processes for collecting and reporting ethnic group data, as well as links to other resources (DH/Health and Social Care Information Centre/NHS Employers 2005).

Summary of activities to improve ethnicity data in health systems internationally

Comparisons with Australia, the United Kingdom, and the United States identify similarities in the barriers and facilitators to high quality ethnicity data in health systems, as well as some consistency in the types of activities or strategies that have been engaged to improve data. This includes a general move towards increasing standardisation of ethnicity data collection, recording and use across the different jurisdictions, supported by the development and dissemination of best practice guidelines. Ethnicity data collection guidelines for health settings have existed for a number of years in the United Kingdom, but are more recent for Australia and the United States. In addition to overarching guidelines, recurring strategies identified to improve ethnicity data include provision of training and resources, the need for data quality assurance processes, and the incentivisation of data collection.

CONCLUSIONS

There is a strong rationale for improving and maintaining the quality of ethnicity data in the health and disability sector. This is important for the sector itself, but also in order to meet broader monitoring purposes and Crown obligations.

The health sector in New Zealand has a history of collecting ethnicity data, although practice has often been variable and non-standardised. There are now some clear policy and strategy imperatives behind the collection of health sector ethnicity data, and accompanying support for the improvement of data collections and evaluation of data quality. In particular, the ethnicity data protocols provide clear guidance on recommended best practice for the sector. In addition, the move towards electronic health records signalled in the National Health IT Plan (National Health IT Board 2010) means that consideration will need to be given to how to ensure that quality ethnicity data is integral to information technology developments in the health and disability sector. The Institute of Medicine (2009) identified the need for standards to both guide the inclusion of ethnicity (or race/ethnicity in the US context) in electronic health records and also how this information will be transferred between providers.

The significant advances made in the health sector towards standardised, high-quality ethnicity data should be recognised. Since the mid-1990s, the sector has promoted the use of the census ethnicity question in order to both standardise health sector data and improve comparability of data with other datasets, including official statistics. In contrast to other social sectors, the health sector has protocols for the collection, input, analysis and output of ethnicity data. In this respect, and in relation to the completeness and quality of ethnicity data for indigenous peoples in other jurisdictions, the sector in New Zealand is performing relatively well.

In New Zealand, the health sector has actively engaged with the official statistical practices, reviews and standards, including consulting within the sector about how official statistical standards will be implemented and operationalised. Recommended health sector practice aligns for the most part with the current Statistical Standard for Ethnicity (with the exception of decisions about the use of prioritisation for input and output of multiple ethnicities).

As is evident from the mapping of the current state of ethnicity data in the sector, issues remain in terms of gaps in collection, as well as data quality. While the evidence is somewhat limited, it suggests that variation in collection practices remain and misclassification of Māori continues to exist in both primary care and hospital settings, which results in undercounting of Māori. In addition, there is some evidence of quality issues with ethnicity data on national collections.

The centrality of the NHI to national collections means that supporting the improvement of the quality and completeness of ethnicity data on the NHI should be seen as a key priority.

Options to improve the collection of ethnicity data more broadly across the sector need to be considered. Gaps in our understanding of Māori health and ethnic disparities remain because we lack data on use of the full range of healthcare services in the sector, such as private hospitals, hospices, community-based rehabilitative and support services, for example.

Activities to improve data quality have been undertaken in New Zealand and have probably fed in to improved data collection. In some cases, this is demonstrated by lower undercounting of Māori than historically has been the case, such as in the New Zealand Cancer Registry, and by increased coverage of data, as in the primary care sector. However, improvements in coverage of themselves do not give us information about the quality of the data. There is an absence of co-ordinated or routine evaluation or audit of ethnicity data. While some audits have been undertaken, they are mostly based in secondary care, and have usually been one-off activities or as part of a specific research project rather than a routine process. The development of audit tools for use in settings other than hospitals is a positive step, and is in line with recommendations made in the WAVE report in the early 2000s about the need for regular, ongoing audit of ethnicity data. In addition, where ethnicity data improvement or quality assurance activities have been undertaken in New Zealand, there is limited accessible information about the nature of these activities, and activities tend not to be formally evaluated (a situation that has also been identified overseas). This makes it difficult to share learnings about successful (or unsuccessful) methods to improve and maintain ethnicity data, and may result in unnecessary duplication of effort and resources.

While the literature on barriers and facilitators is also relatively limited, there is consistency in issues that have identified over time and across settings. Factors that are recurring in the literature include issues around the leadership, co-ordination and funding of ethnicity data quality activities, issues with information systems, the need for training and availability to support data collection and use, and awareness and understanding of the concept of ethnicity and purpose of data collection in the health sector among both data collectors and respondents.

The problems with ethnicity data quality are not new, nor are the barriers and facilitators. Much good work has been undertaken, and substantial progress has been made in the health sector over time, particularly in terms of movement towards increased standardisation and

better coverage. However, a concerted focused and resourced effort is still required to ensure continued improvements to data quality going forward, as well as to provide tools to support the maintenance of high quality datasets. It is likely that this will be achieved by incentivising data collection appropriately for health institutions and data collectors, coupled with the necessary checks and balance. It will also require strong, committed leadership at both national and local levels. The limitations of health and disability sector ethnicity data remain a significant barrier to a comprehensive understanding of Māori health and ethnic inequalities and, therefore, an impediment to our ability to address these effectively and in a timely manner to achieve goals of good health for all.

REFERENCES

- Ajwani S, Blakely T, Robson B, Tobias M, Bonne M. (2003). *Decades of disparity: ethnic mortality trends in New Zealand 1980–1999*. Wellington: Ministry of Health and University of Otago.
- Australian Institute of Health and Welfare. (2010). *National best practice guidelines for collecting Indigenous status in health data sets*. Canberra: AIHW.
- Bailey T. (1998). *The collection of ethnicity information: a consumer's perspective*. Wellington: Department of General Practice, Wellington School of Medicine.
- Baker D, Cameron K, Feinglass J, Thompson J, Georgas P, Foster S, Pierce D, Hasnain-Wynia R. (2006). 'A system for rapidly and accurately collecting patients' race and ethnicity'. *American Journal of Public Health*, 96(3): 532–537.
- Baker D, Cameron K, Feinglass J, Georgas P, Foster S, Pierce D, Thompson J, Hasnain-Wynia R. (2005). 'Patients' attitudes toward health care providers collecting information about their race and ethnicity'. *Journal of General Internal Medicine*, 20: 895–900.
- Blakely T, Atkinson J, Fawcett J. (2008). 'Ethnic counts on mortality and census data (mostly) agree for 2001-2004: New Zealand Census-Mortality Study update'. *New Zealand Medical Journal*, 121(1281): 58–62.
- Blakely T, Tobias M, Atkinson J, Yeh L-C, Huang K. (2007). *Tracking disparity: trends in ethnic and socioeconomic inequalities in mortality, 1981–2004*. Wellington: Ministry of health.
- Bramley D, Latimer S. (2007). 'The accuracy of ethnicity data in primary care'. *New Zealand Medical Journal*, 120(1262): U2779.
- Brown, P. (1983). *An investigation of official ethnic statistics*. Wellington: Department of Statistics.
- Canadian Institute for Health Information (CIHI). (2005). *The CIHI Data Quality Framework*. Ottawa, Ontario: Canadian Institute for Health Information.

Cormack D. (2010). *The practice and politics of counting: ethnicity data in official statistics in Aotearoa/New Zealand*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

Cormack D & Harris R. (2009). *Issues in monitoring Māori health and ethnic disparities: an update*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

Cormack D, Purdie G & Robson B. (2007). 'Cancer'. In B Robson & R Harris (eds). *Hauora: Māori Standards of Health IV: A study of the years 2000-2005*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

Cormack D, Robson B, Purdie G, Ratima M, Brown R. (2005). *Access to cancer services for Māori: a report for the Ministry of Health*. Wellington: Ministry of Health.

Curtis E, Wright C, Wall M. (2005). 'The epidemiology of breast cancer in Māori women in Aotearoa New Zealand: implications for ethnicity data analysis'. *New Zealand Medical Journal*, 118 (1209).

Department of Health/Health and Social Care Information Centre/NHS Employers. (2005). *A practical guide to ethnic monitoring in the NHS and social care*. London: Department of Health/Health and Social Care Information Centre/NHS Employers.

Department of Statistics. (1998). *Report of the Review Committee on Ethnic Statistics*. Wellington: Department of Statistics.

de Lusignan S, Hague N, van Vlymen J, Kumarapeli P. (2006). 'Routinely-collected general practice data are complex, but with systematic processing can be used for quality improvement and research'. *Informatics in Primary Care*, 14: 59–66.

Donaldson, H. (2002). *Ethnicity data collection in Dunedin Public Hospital: the process and impacting issues*. Dunedin: Te Rōpū Rangahau Hauora a Ngai Tahu, University of Otago.

Gomez S, Kelsey J, Glaser S, Lee M, Sidney S. (2004). 'Inconsistencies between self-reported ethnicity and ethnicity recorded in a Health Maintenance Organization'. *Annals of Epidemiology*, 15: 71–79.

Graham P, Jackson R, Beaglehole R, & de Boer G. (1989). 'The validity of Māori mortality statistics'. *New Zealand Medical Journal*, 102(864): 124–126.

- Gribben B, Coster G, Pringle M, Simon J. (2001). 'Non-invasive methods for measuring data quality in general practice'. *New Zealand Medical Journal*, 114: 30–32.
- Hadi H, Ayers R, Chang K, Doherty G, Forest A, Leuy B, Mardani J, Morrison M, Patel B, Shaw C. (1997). *An evaluation of ethnicity classification in hospital admissions systems: a review of the Napier and Hastings Hospitals*. Wellington: Wellington School of Medicine.
- Hall J, Tomlin A, Martin I, Tilyard M. (2002). 'A general practice minimum data set for New Zealand'. *New Zealand Medical Journal*, 115(1163): U200.
- Hamilton N, Edelman D, Weinberger M, Jackson G. (2009). 'Concordance between self-reported race/ethnicity and that recorded in a Veteran Affairs electronic medical record'. *North Carolina Medical Journal*, 70(4): 296–300.
- Harris R, Purdie G, Robson B, Wright C, Zhang J, Baker M. (2007). 'Estimating Māori hospitalisations and cancer registrations: Appendix 3'. In: B Robson & R Harris (3eds). *Hauora: Māori Standards of Health IV. A study of the years 2000–2005*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare, pp. 249–259.
- Harris R, Robson B, Reid P, Keefe V. (1997). *CHE cultural reviews: ethnicity data collection – error rates and training effectiveness (Report to the Central Regional Health Authority)*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.
- Hasnain-Wynia R & Baker D. (2006). 'Obtaining data on patient race, ethnicity, and primary language in Health Care Organizations: current challenges and proposed solutions'. *Health Services Research*, 41(4): 1501–1518.
- Hasnain-Wynia R, Pierce D & Pittman M. (2004). *Who, when, and how: the current state of race, ethnicity, and primary language data collection in hospitals*. The Commonwealth Fund.
- He Kamaka Oranga. (1998). *Collection of patient ethnicity audit*. Auckland: Māori Health Management Service.
- Health Information Strategy Steering Committee. 2005. *Health information strategy for New Zealand*. Wellington: Ministry of Health.
- Health Outcomes International Pty Ltd. *Methods and Systems Used to Measure and Monitor Occupational Disease and Injury in New Zealand: NOHSAC Technical Report 2*: Wellington, 2005.

Health Utilisation Research Alliance (HURA). (2006). 'Ethnicity data and primary care in New Zealand: lessons from the Health Utilisation Research Alliance (HURA) study'. *New Zealand Medical Journal*, 119(1231): U1917.

Hoque E, Hope V, Scragg R, Baker M, Shrestha R. (2004). 'A descriptive epidemiology of giardiasis in New Zealand and gaps in surveillance data'. *New Zealand Medical Journal*, 117(1205): U1149.

Horsburgh S, Norris P, Becket G, Crampton P, Arroll B, Cumming J, Herbison P, Sides G. (2010). 'The Equity in Prescription Medicines Use Study: using community pharmacy databases to study medicines utilisation'. *Journal of Biomedical Informatics*, in press.

Information Directorate. (2010). *National Booking Reporting System Data Dictionary*. Wellington: Ministry of Health.

Institute of Medicine. (2009). *Race, ethnicity, and language data: standardization for health care quality improvement*. Washington, DC: The National Academies Press.

Iqbal G, Gumber A, Johnson M, Szczepura A, Wilson S, Dunn J. (2009). 'Improving ethnicity data collection for health statistics in the UK'. *Diversity in Health and Care*, 6: 267–285.

Iqbal, G., Gumber, A., Szczepura, A., Johnson, M., Wilson, S., & Dunn, J. (2008). *Improving ethnic data collection for statistics of cancer incidence, management, mortality and survival in the UK*.

Jack R, Linklater K, Hofman D, Fitzpatrick J, Moller H. (2006). 'Ethnicity coding in a regional cancer registry and in Hospital Episode Statistics'. *BMC Public Health*, 6(281).

Joshy G, Lawrenson R, Simmons D. (2008). 'Retention of patients in the Get Checked free annual diabetes review programme in New Zealand'. *New Zealand Medical Journal*, 121(1270): U2945.

Kamira, R. (2008). *Ethnicity standards consultation report for the Ministry of Health*. Puaa Interface Ltd.

Kandula N, Hasnain-Wynia R, Thompson J, Brown E, Baker D. (2009). 'Association between prior experiences of discrimination and patients' attitudes towards health care providers collecting information about race and ethnicity'. *Journal of General Internal Medicine*, 24(7): 789–794.

Kelagher M, Parry A, Day S, Paradies Y, Lawlor J & Solomon L. (2010). *Improving the identification of Aboriginal and Torres Strait Islander people in mainstream general practice*. Melbourne: The Lowitja Institute.

Kerr K. (2009). *The development of a data quality framework and strategy for the New Zealand Ministry of Health*. Department of Information Systems and Operations Management, University of Auckland: Unpublished working paper.

Kerr K. (2003). *Ethnicity data quality process: review of Ministry of Health national data collections*. Ministry of Health: Unpublished working paper.

Kilgour R & Keefe V. (1992). *Kia piki te ora: the collection of Māori health statistics*. Wellington: Department of Health.

Kirk, R., Barnett, P., Clayden, C., Cumming, J., Heckert, K., Penrose, A., Pitama, S., & Siataga, P. (2002). *Evaluation of the Pegasus Health Global Budget Contract*. Christchurch: New Zealand Health Technology Assessment (NZHTA).

Kljakovic M. (1993). 'Is it easy collecting ethnicity data in general practice?' *New Zealand Medical Journal*, 106: 103–104.

Kukutai T. *The dynamics of ethnicity reporting: Māori in New Zealand*. A discussion paper prepared for Te Puni Kokiri. Population Studies Centre, University of Waikato; 2003.

Kumarapeli P, Stepaniuk R, de Lusignan S, Williams R, Rowlands G. (2006). 'Ethnicity recording in general practice computer systems'. *Journal of Public Health*, 28(3): 283–287.

Lang K. (2002). *Measuring ethnicity in the New Zealand Population Census*. Wellington: Statistics New Zealand.

Langton J & Crampton P. (2008). 'Capitation funding of primary health organisations in New Zealand: are enrolled populations being funded according to need?' *New Zealand Medical Journal*, 121(1272): 47–58.

Latimer S. (2003). *Waitemata DHB Ethnicity Data Collection Baseline Review*. Waitemata: Waitemata DHB.

Lindsay G. (2005). 'The validity of primary care ethnicity data in New Zealand'. Master of Public Health dissertation. Unpublished: Section Epidemiology and Biostatistics, School of Population Health, University of Auckland.

- Lopez L, Sexton K, Heffernan H. (2010). *Tuberculosis in New Zealand: Annual Report 2009*. Wellington: ESR.
- Loring B & Curtis E. (2009). 'Routine vaccination coverage of 11 year olds, by ethnicity, through school-based vaccination in South Auckland'. *New Zealand Medical Journal*, 122(1291): U3509.
- Love T & Crampton P. (2001). 'Deprivation profiles in Wellington IPA practices'. *New Zealand Family Physician*, 28(5), 327–331.
- Maizlish N & Herrera L. (2006). 'Race/ethnicity in medical charts and administrative databases of patients served by community health centers'. *Ethnicity & Disease*, 16: 483–487.
- Malcolm L. (2010). 'Towards a reliable and accurate ethnicity database at district and national levels: progress in Canterbury'. *New Zealand Medical Journal*, 123(1314): U4108.
- Malcolm, L., & Barnett, R. (2009a). 'Utilisation of inpatient cardiology services including by Māori: a study of hospital discharges for patients enrolled with Partnership Health practices for the 2 years ending June 2007'. *New Zealand Medical Journal*, 63(122): 63–71.
- Malcolm, L., & Barnett, R. (2009a). 'Variation in Partnership Health general practice enrolment data related to need and hospital national records on ethnicity'. *Journal of Primary Health Care*, 1(4): 297–301.
- Mandelblatt J, Yabroff R & Kerner J. (1999). 'Equitable access to cancer care: a review of barriers to quality care'. *Cancer*, 86(11): 2378–2390.
- Marshall R, Zhang Z, Broad J, Wells S. (2007). 'Agreement between ethnicity recorded in two New Zealand health databases: effects of discordance on cardiovascular outcome measures (PREDICT CVD-3)'. *Australian and New Zealand Journal of Public Health*, 31(3): 211–16.
- McAlpine D, Beebe T, Davern M, Call K. (2007). 'Agreement between self-reported and administrative race and ethnicity data among Medicaid enrollees in Minnesota'. *Health Services Research*, 42(6): 2373–2388.
- McLeod D, Mansell J, Harris R, Bailey T, Dowell A, Robson B, Reid P. (2000). 'The collection of patient ethnicity data: a challenge for general practice'. *New Zealand Family Physician*, 27(3): 51–57.

Minister of Disability Issues. (2001). *New Zealand Disability Strategy*. Wellington: Ministry of Health.

Ministry of Health. (2010a). *Tatau kahukura: Māori health chart book 2010, 2nd edition*. Wellington: Ministry of Health.

Ministry of Health. (2010b). *Cancer: New Registrations and Deaths 2007*. Wellington: Ministry of Health.

Ministry of Health. (2009a). *National Health Index Data Dictionary: version 5.3*. Wellington: Ministry of Health.

Ministry of Health. (2008). *The New Zealand health and disability system: organisations and responsibilities*. Briefing to the Minister of Health. Wellington: Ministry of Health.

Ministry of Health. (2007). *Capitation enrolment and register handling guidelines*. Wellington: Ministry of Health.

Ministry of Health. (2006). *Whakatataka Tuarua: Māori Health Action Plan 2006–2011*. Wellington: Ministry of Health.

Ministry of Health. (2004a). *Ethnicity Data Protocols for the Health and Disability Sector*. Wellington: Ministry of Health.

Ministry of Health. (2004b). *Enrolment requirements for primary health organisations*. Wellington: Ministry of Health.

Ministry of Health. (2002). *He Korowai Oranga*. Wellington: Ministry of Health.

Ministry of Health. (2001a). *Monitoring ethnic inequalities in health*. Wellington: Ministry of Health.

Ministry of Health. (2001b). *The Primary Health Care Strategy*. Wellington: Ministry of Health.

Ministry of Health. (2000). *The New Zealand Health Strategy*. Wellington: Ministry of Health.

Ministry of Health & Long Term Care (MOHLTC). (2006). *Building a data quality management framework for Ontario*. Ontario: Ministry of Health & Long Term Care.

- Moala A. (1999). *Recording of ethnicity information at Capital Coast Health: Final report on Pacific Health Sector Fellowship to the Ministry of Health*. Wellington: Unpublished.
- National Health Board Business Unit. (2010). *National Minimum Dataset (Hospital Events) data dictionary*. Wellington: Ministry of Health.
- National Health IT Board. (2010). *National Health IT Plan*. Wellington: National Health IT Board.
- New Zealand Health Information Service. (1995). *Multiple Ethnicity System Specification*. Wellington: New Zealand Health Information Service.
- Norris P, Horsburgh S, Padukkage P, Baik N, Kim D, Fussell A, Hutchinson S, Ragupathy R. (2010). 'Coverage and accuracy of ethnicity data on three Asian ethnic groups in New Zealand'. *Australian and New Zealand Journal of Public Health*, 34(3): 258–261.
- Pasupati M, Chapman J, Brown G, Moore P, Clark K, Bluett D, Loveridge G, Read D, Reeve B, Wiren P. (1980). *Ethnic classification in New Zealand health statistics*. Wellington: Wellington Clinical School of Medicine.
- Paua Interface Ltd. (2007). *Ethnicity data training & review report to the Ministry of Health*. Paua Interface Ltd.
- Perot R & Youdelman M. (2001). *Racial, ethnic, and primary language data collection in the health care system: an assessment of federal policies and practices*. Xxx: The Commonwealth Fund.
- Pinheiro P, Sherman R, Fleming L, Gomez-Marin O, Huang Y, Lee D, Penedo F. (2009). 'Validation of ethnicity in cancer data: which Hispanics are we misclassifying?' *Journal of Registry Management*, 36(2): 42–47.
- Priest, P., & Jackson, R. (1994). The validity of Māori hospital discharge statistics for ischaemic heart disease in New Zealand. *New Zealand Medical Journal*, 170: 3–4.
- Pringle M & Rothers I. (1996). 'Practicality of recording patient ethnicity in general practice: descriptive intervention study and attitude survey'. *BMJ*, 312: 1080–1082.
- Regenstein M & Sickler D. (2006). *Race, ethnicity and language of patients: hospital practices regarding collection of information to address disparities in health care*. The Robert Wood Johnson Foundation and National Public Health and Hospital Institute.

Riddell T, Lindsay G, Kenealy T, Jackson R, Crengle S, Bramley D, Wells S, Marshall R. (2008). 'The accuracy of ethnicity data in primary care and its impact on cardiovascular risk assessment and management – PREDICT CVD-8'. *New Zealand Medical Journal*, 121(1281): U3239.

B Robson & R Harris (eds). (2007). *Hauora: Māori Standards of Health IV: A study of the years 2000–2005*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

Robson B, Purdie G, Cormack D. (2006). *Unequal impact: Māori and non-Māori cancer statistics 1996–2001*. Wellington: Ministry of Health.

Robson B, Reid P, Keefe-Ormsby V, Pere L, Malcolm G, Scott N, Ormsby C. (1996). *A method for auditing the collection of ethnicity data in hospital statistics: a report to Te Kete Hauora*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

Shaw C, Atkinson J, Blakely T. (2009). '(Mis)classification of ethnicity on the New Zealand Cancer Registry: 1981-2004'. *New Zealand Medical Journal*, 122(1294): 10–22.

Sheerin I, Hamilton G, Humphrey A, Scragg A. (2007). 'Issues in the assessment of cardiovascular risk in selected general practices in Canterbury, New Zealand'. *New Zealand Medical Journal*, 120(1261): U2714.

Siegel B, Regenstein M, Jones K. (2007). *Enhancing public hospitals' reporting of data on racial and ethnic disparities in care*. The Commonwealth Fund.

Smedley B, Stith A, Nelson A. (eds). (2002). *Unequal treatment: confronting racial and ethnic disparities in health care*. Washington: The National Academies Press.

Southern Cancer Network. (2009). Available on: http://www.westcoastdhb.org.nz/links/southern_cancer_network/newsletters/SCN_newsletter_2009_12.pdf, accessed October 2010.

Sporle A & Pearce N. 'Impact of changes in the death registration process upon Maori mortality statistics'. *New Zealand Medical Journal*, 1999(112): 411–12.

Statistics Act 1975, Accessed on February 10 2009, Available from: <http://www.legislation.govt.nz/act/public/1975/0001/latest/viewpdf.aspx>

Statistics New Zealand. (2010). *Demographic trends: 2009*. Wellington: Statistics New Zealand.

- Statistics New Zealand. (2009). *Final report of a review of the official ethnicity statistical standard 2009*. Wellington: Statistics New Zealand.
- Statistics New Zealand. (2005). *Statistical Standard for Ethnicity 2005*. Wellington: Statistics New Zealand.
- Statistics New Zealand. (2004). *Report of the Review of the Measurement of Ethnicity*. Wellington: Statistics New Zealand.
- Statistics New Zealand. (1999). *Measuring Māori ethnicity in the New Zealand Census*. Wellington: New Zealand.
- Statistics New Zealand. (1996). *Ethnic Death Statistics Workshop Report*. 6 August 1996, Royal Society Rooms, Wellington.
- Stevens W, Stevens G, Kolbe J, Cox B. (2008). 'Comparison of New Zealand Cancer Registry data with an independent lung cancer audit'. *New Zealand Medical Journal*, 121(1276): 29–41.
- Swan, J., Lillis, S., & Simmons, D. (2006). 'Investigating the accuracy of ethnicity data in New Zealand hospital records: still room for improvement'. *New Zealand Medical Journal*, 119(1239): U2103.
- Tan L, Blakely T, Atkinson J. (2010). 'Ethnic counts on mortality and census data 2001-06: New Zealand Census-Mortality Study update'. *New Zealand Medical Journal*, 123(1320): 37–44.
- Te Puni Kokiri. (2001). *Follow up Review of the Accident Compensation Corporation's Service Deliver to Māori*. Wellington: Monitoring and Evaluation Social Development Directorate, Te Puni Kokiri.
- Te Puni Kokiri. (1998). *Review of ACC: service delivery to Māori*. Wellington: Te Puni Kokiri.
- Te Puni Kokiri. (1997). *ACC's service delivery: Māori claimant and provider perspectives (key findings)*. Wellington: Te Puni Kokiri.
- Te Rōpū Rangahau Hauora a Eru Pōmare (TRRHAEP). (2000). 'Counting for nothing: understanding the issues in monitoring disparities in health'. *Social Policy Journal of New Zealand*, 14: 1–16.

Te Rōpū Rangahau Hauora a Eru Pōmare (TRRHAEP). (1996). *The classification of ethnicity in hospital discharge statistics: the experience at Hutt Valley Health 1995*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare.

The Waitangi Tribunal. (2001). *The Napier Hospital and Health Services Report*. Wellington: Legislation Direct.

Thomas D. (2001). 'Assessing ethnicity in New Zealand health research'. *New Zealand Medical Journal*, 114(1167): 84–88.

Tilyard M, Munroe N, Walker S, Dovey S. (1998). 'Creating a general practice minimum data set: present possibility or future plan?' *New Zealand Medical Journal*, 111: 317–320.

Tomlin A & Hall J. (2004). 'Linking primary and secondary healthcare databases in New Zealand'. *New Zealand Medical Journal*, 117(1191): U816.

Tomlin A & Martin I. (2003). 'Registered and unregistered patients in general practice: implications for primary health organisations'. *New Zealand Medical Journal*, 116(1186): U684.

Watson S. (2005). *Equity measures and systems reform as tools for reducing racial and ethnic disparities in health care*. The Commonwealth Fund.

WAVE Advisory Board to the Director-General of Health. (2001). *From strategy to reality: the WAVE project*. Wellington: Ministry of Health.

Wright C. (2009). *Comparison of ethnicity data in the PHO Register and New Zealand Health Survey 2006/07 (Unpublished report)*. Wellington: Health and Disability Intelligence Unit, Ministry of Health.

APPENDIX ONE: KEY EVENTS TIMELINE

Health sector	YEAR	Government sector
	1975	<i>Statistics Act 1975</i>
	1988	Review Committee on Ethnic Statistics
	1995	<i>Births, Deaths, Marriages and Relationships Act 1995</i>
Alignment of hospital ethnicity data with Census approach	1996	
	2000	Review of the Measurement of Ethnicity, Statistics New Zealand
WAVE project report released	2001	
<i>He Korowai Oranga: Māori Health Strategy released</i>	2002	
Ethnicity Data Improvement Project, MOH	2002/03	
<i>Ethnicity Data Protocols for the Health and Disability Sector</i>	2004	Final report on the Review of the Measurement of Ethnicity
<i>Implementation of the Protocols Health Information Strategy for New Zealand released</i>	2005	<i>Statistical Standard for Ethnicity 2005</i> released, Statistics New Zealand
<i>Whakatātaka Tuarua: Māori Health Action Plan released</i>	2006	Significant increase in 'New Zealander' responses in 2006 Census
	2007	Cabinet directs government agencies to align with 2005 Standard (SNZ)
Consultation with health sector on alignment with 2005 Statistical Standard (SNZ)	2008	
NCAMP changes to align with 2005 Standard (SNZ)		
Improving data quality for ethnicity project, MOH		
Discussion paper	2009	Updated Standard, Statistics New Zealand
Investment Logic Mapping (ILM) workshops Primary Care Audit Tool project	2010	

KEY EVENTS DETAILS

- **1975 Statistics Act**

The Statistics Act 1975 made it a statutory requirement for data on ethnicity (ethnic origin) to be collected in New Zealand in population censuses.

- **1998 Review Committee on Ethnic Statistics**

The Review Committee on Ethnic Statistics undertook a major review of ethnicity in official statistics in the 1980s. The Committee made a series of recommendations, including that Statistics New Zealand (then known as Department of Statistics) develop a statistical standard for ethnicity as a move towards standardising official ethnicity data.

- **Births, Deaths, Marriages and Relationships Act 1995**

The 1995 Act aligned the ethnicity collection on birth and death registrations with the Statistics New Zealand approach. Prior to the Act, data was only collected for Māori and Pacific registrations, and was based on proportions-of-descent criteria. Following the Act, ethnicity data was collected using the Statistics New Zealand census approach (initially the 1996 Census question, and currently the 2001 Census question), and was collected for all birth and death registrations.

- **Alignment of hospital ethnicity data collection with Statistics New Zealand approach**

In 1996, ethnicity data collection in public hospitals was officially aligned with the approach in the population census. Data collection was to be based on respondents self-identifying the ethnic group or groups they belonged to, although there is evidence that variability in data collection practices remain.

- **Review of the Measurement of Ethnicity (RME)**

Statistics New Zealand undertook a review of the measurement of ethnicity in 2000. The review involved the release of discussion papers, consultation with stakeholders, submissions on draft recommendations, and a final report (released in 2004).

- **WAVE project report released**

The WAVE programme aimed to produce “an information and technology plan for the sector with the aim of improving health outcomes – through the effective use of information, at the least cost” (MOH 2001). The WAVE project emphasised the need for quality, accurate and timely information. The report commented on the lack of ethnicity data, issues with collection, the need for quality ethnicity data and the need to audit ethnicity data in health information systems annually.

- **He Korowai Oranga: Māori Health Strategy 2002**

He Korowai Oranga identified the need for quality information, including reliable ethnicity data, as part of meeting the goals of improving Māori health and reducing inequalities.

- **Ethnicity Data Improvement Project, Ministry of Health**

The Ministry of Health undertook a project focused on improving ethnicity data in the health and disability sector, which included the development of protocols for standardising ethnicity data across the sector.

- **Ethnicity Data Protocols for the Health and Disability Sector, 2004**

In 2004, the Ministry of Health released protocols to guide the standardisation of ethnicity data collection, recording and output across the health and disability sector.

- **Report on the Review of the Measurement of Ethnicity, 2004**

Statistics New Zealand released its report on the RME in 2004. The report signalled a number of changes to the official statistical standard, including a revised definition, changes to the official classification, changes to the coding of ‘New Zealander’-type responses, and changes to the recommended output of multiple ethnicities.

- **Implementation of the protocols**

Training around the Ethnicity Data Protocols was undertaken with some DHBs, using a Train-the-Trainer approach and resources that had been developed to aid the implementation of the Protocols.

- **Statistical Standard for Ethnicity 2005**

In 2005, Statistics New Zealand released a new official statistical standard for ethnicity. The Standard outlined the official approach to defining, collecting, and outputting ethnicity data. It was intended as a whole-of-government standard.

- **Health Information Strategy New Zealand**

The Health Information Strategy New Zealand (HIS-NZ) was released in 2005. It reaffirmed the need of the sector for timely, quality information to inform policy development and aid monitoring and evaluation.

- **Whakatātaka Tuarua: Māori Health Action Plan 2006**

As part of Whakatātaka Tuarua, a number of specific priorities around improving the collection of ethnicity data in the health sector were identified.

- **Increase in 'New Zealander' responses in the 2006 Census**

In the 2006 Census, there was a significant increase in the number of people who reported a 'New Zealander'-type response to the Census ethnicity question, representing 11% of the total population. This had impacts on data comparability, over time and between datasets.

- **Cabinet directive to align with the 2005 Statistical Standard, 2007**

The 2005 Statistical Standard for Ethnicity was intended to be a whole-of-government standard. In 2007, Cabinet directed that government agencies align with the Standard. As part of this, some changes were deemed mandatory, while others were nominated as voluntary.

- **Consultation with the health and disability sector on alignment with the Statistical Standard for Ethnicity, 2008**

In 2008, consultation was undertaken with the health sector on alignment with the Statistical Standard for Ethnicity 2005, and the potential implications of this. The consultation informed decisions about how the health sector would move toward alignment with the Standard.

- **National Collections Annual Maintenance Project, 2009**

Following on from the consultation in 2008, changes to the ethnicity codeset for the health sector were made as part of the 2009 National Collections Annual Maintenance Project, in order that the Level 2 codes aligned with the 2005 Statistical Standard. However, it was decided not to change to coding 'New Zealander' responses to a separate category (as in the Statistics New Zealand standard).

- **Improving data quality for ethnicity, Ministry of Health, 2009**

The Ministry of Health has a programme to improve ethnicity data quality. As part of this, an Investment Logic Mapping (ILM) process was commenced in 2009.

- **Discussion paper on issues in measuring and monitoring Maori health released, 2009**

A discussion paper on ethnicity data and issues in measuring and monitoring Maori health was released in late 2009. The discussion paper was part of a series of papers being produced for Te Kete Hauora to consider current and future ethnicity data issues.

- **Review of the Official Statistical Standard, 2009**

As part of the development of the 2011 Census, a review of the 2005 Statistical Standard was undertaken. The review focused principally on the issue of 'New Zealander'-type responses. Draft recommendations were reported in April 2009, with a final report in October 2009. The Review recommended retaining the 2006 Census question, introducing an alternate classification to group 'New Zealander' responses with European responses, and exempting particular data collections from coding 'New Zealander' responses to the 'Other' category.

- **Investment Logic Mapping, Ministry of Health, 2010**

As part of the ILM process, two workshops were held in 2010 with key stakeholders.

- **Primary care audit tool project, 2010**

The Ministry of Health recently funded an RFP to develop an ethnicity audit tool for the primary care sector.

