



A general practice minimum data set for New Zealand

Jason Hall, Andrew Tomlin, Isobel Martin and Murray Tilyard

Abstract

Aim To assess the completeness of primary care data collected by the Royal New Zealand College of General Practitioners' (RNZCGP) Dunedin Research Unit, and assess the feasibility of creating a New Zealand national minimum data set for primary care.

Methods Patients from 42 practices contributing data to the Dunedin Research Unit made up the study population. A six-month sample of data was evaluated for completeness, and compliance to a minimum data set structure. Rates of recording patient identifiers, sex, ethnicity, community services card status, consultation identifier and date, prescriptions and Read codes were calculated for each practice and registered patient.

Results Patient demography, NHI and community services card status were all well recorded (date of birth 99.3%, sex 98.9%, NHI 94.8%, CSC 100.0%). Read codes and ethnicity were still poorly recorded, although there was wide variation between practices.

Conclusions The completeness of data collected by the Dunedin Research Unit appears to be improving, although there is wide variation between contributing practices. The capability to create a primary care national minimum data set exists, but this will not become a reality until suitable education programmes and support are supplied for general practitioners and other staff members who record patient information.

Accurate data from general practice is needed for good clinical management, audit, teaching, research, administration, financial control, general management, and statutory and legal needs.¹ An ideal clinical database would contain accurate information about all aspects of patient care, but such a gold standard is difficult to obtain. The most difficult data to capture is the codified record of clinical history, the physical examination, and the physician's impressions, assessment of risk and differential diagnosis.² Practicality leads us to conceive the smallest number of essential data items required among many potential users – a minimum basic data set.

Criteria for inclusion in a minimum basic data set include: finding data that can be readily collected with reasonable accuracy and economy; avoiding unnecessary duplication of data available from other sources; conforming to the limits of patient and provider confidentiality; and ensuring data collection is periodically considered for both its utility and cost.³

Results from a 2001 government survey have indicated that 57% of New Zealand general practitioners used patient management systems (PMSs) to record patient details and to store prescription and some clinical data. It is estimated that by 2004, 89% will be using PMSs.⁴ Patient-centred electronic health records supported by the

main PMS vendors include clinical notes. New Zealand is also in the unique position of having the National Health Index (NHI) database. The NHI is a register of users of the healthcare system in New Zealand that provides positive unique identification of an individual under the strict legislative control of privacy, the 1993 Privacy Act. In an encrypted form, it also provides the electronic link that enables the matching of primary care records to information contained in the National Minimum Data Set for secondary care.

There are no legal requirements in New Zealand to maintain patient records on paper as well as in electronic form. This means that if computers are used to store practice records, they are often used more comprehensively than in other countries.⁵ There is therefore sufficient electronic capability and infrastructure to provide a platform that will support the collection of electronic health records from general practice. A national minimum data set for primary care would provide detailed specifications for recording and categorisation of patient data in a standardised format.

The Dunedin Research Unit investigated the possibility of creating a minimum data set using data extracted in 1995 from 12 computerised general practices.⁶ This study demonstrated that the variability in the comprehensiveness of data recorded on PMSs was one of the major difficulties in achieving a minimum data set, even when considering information from practices accustomed to providing data for non-clinical reasons such as research. Such difficulties were deemed to be surmountable and a national minimum data set achievable using existing data sources. This current study revisits our earlier work, but reviews data from practices nationwide and also examines routinely recorded prescription data.

Methods

Forty two New Zealand general practices provided data for this study. All were members of the RNZCGP Computer Research Group and used the same PMS (Medtech-32: Health Technology Ltd). This PMS was chosen because it has the largest market share for a general practice PMS (2001 MediMedia PMS in general practice survey) and because the Research Unit has proven extraction routines for accessing its data.

Records are extracted from practice computers using programs written by the PMS software manufacturers in accordance with Research Unit specifications. These programs remove all patient identifying information such as name, address and telephone number. Each patient is allocated a unique code that is individuating but non identifying. If recorded, the National Health Index (NHI) of the patient is encrypted before transferral of data to the Dunedin Research Unit. The Research Unit does not have access to the decryption program.

Data for the six-month period, 1 July 2000 to 31 December 2000, was extracted from the study practices. Three main files were reviewed for validity of data values and comprehensiveness of patient information recorded: 1) patient details files, 2) consulting files and associated Read coding, 3) prescribing files (Table 1).

Patient details files were reviewed to determine the proportion of registered patients with a unique identifier and NHI, and the proportion with demographic data (date of birth, sex, ethnicity) and information relating to eligibility for government health subsidies (community card status and general medical subsidy). Attributes examined in the consultations files included the consultation date, the unique consultation identifier, the provider of care and Read coding of consultations. Prescription files were reviewed to determine the validity, utility and level of recording of prescription information, including the prescription date, the pharmaceutical prescribed, its form, strength and dosage, the quantity prescribed and the number of repeats. Descriptive statistics were calculated for each data item (mean, median and range of the percentage of records with a valid value).

Table 1. Proposed elements for the national minimum data set

Field name	Description	Purpose
Patient details file		
Patient identifier	A unique number assigned by the PMS	To calculate healthcare event rates
Practice identifier	A unique number assigned by the Research Unit to all records from a practice	To calculate and compare healthcare event rates by practice
National Health Index (NHI)	A unique alphanumeric code	To enable linking with other data sets and mapping of healthcare use
Sex	Male or Female	To calculate and compare healthcare event rates by sex
Date of birth	Date of birth: dd/mm/yyyy	To calculate and compare healthcare event rates by age
Ethnicity	Ethnicity	To calculate and compare healthcare event rates by ethnicity
Income group	Community card status	To calculate and compare healthcare event rates by income group
Subsidy eligibility	General Medical Subsidy	To calculate and compare healthcare event rates by subsidy eligibility
Consultations file		
Consultation date	Date of consultation: dd/mm/yyyy	To calculate and compare consultation rates by time
Consultation identifier	A unique number assigned by the PMS	To uniquely identify consultations
Read code	A morbidity code	To identify conditions treated in general practice
New Zealand Medical Council number	A provider index number	To calculate and compare consultation rates by different providers
Prescriptions file		
Prescription date	Date of prescription: dd/mm/yyyy	To calculate and compare prescribing rates by time
Pharmaceutical prescribed	Brand or generic name	To calculate and compare prescribing rates by therapeutic group
Form	Physical form eg tablet, syrup, cream	To evaluate treatment outcomes and expenditure
Strength	Various standardised units eg mg, mg/ml	To evaluate treatment outcomes and expenditure
Dosage	Dose/time	To evaluate treatment outcomes and expenditure
Quantity	Quantity prescribed (units)	To evaluate treatment outcomes and expenditure
Number of Repeats	0, 1 or 2	To evaluate treatment outcomes and expenditure

Results

The total registered population of the 42 study practices was 258 249 patients (range: 1246–16 496; median: 5454; mean: 6149). This is approximately 6.7% of New Zealand's population as at 31 December 2000. The practices provided an acceptable geographic distribution of patients, with 16.3% of the study population coming from the Central region, 24.1% from Midland, 26.7% from Northern and 32.9% from the

Southern region. The total number of consultations recorded for the six-month study period was 557 147 (range: 1432–61 738; median: 11 163; mean: 13 265) and there were 542 707 (range: 1321–39 433; median: 10 897; mean 12 970) prescriptions generated (Table 2).

Table 2. Practice compliance with national minimum data set standards

Data Item	Mean	Median	Range
Patient details file			
Total patients (n = 258 249)	6149	5454	1246–16 496
NHI %	94.8	96.1	77.0–99.5
Date of Birth %	99.3	99.6	96.4–100.0
Sex %	98.9	99.1	94.8–100.0
Ethnicity %	34.9	26.8	0.7–99.4
Subsidy eligibility %	100.0	100.0	99.7–100.0
Community card status %	100.0	100.0	100.0–100.0
NZMC number of GP %	93.4	99.8	0.0–100.0
Consultations file			
Total consultations (n = 557 147)	13 265	1163	1432–61 738
Read code %	22.8	16.3	0.0–81.5
Consultation provider %	87.9	97.6	0.3–100.0
Prescriptions file			
Total prescriptions (n = 542 707)	12 970	10 897	1321–39 433
Pharmaceuticals prescribed %	100.0	100.0	100.0–100.0
Form %	94.8	94.6	89.1–99.2
Strength %	99.5	99.5	98.6–99.8
Dosage %	99.4	99.6	96.6–100.0
Quantity %	85.5	91.0	40.9–100.0
Number of repeats %	100.0	100.0	100.0–100.0

A practice identifier is a numeric code assigned to all records on delivery in the Research Unit and this is therefore always present. Data elements automatically created internally by PMSs include the alphanumeric patient identifier, the consultation identifier, consultation date and prescription date. These data elements were 100% complete for all practices. Date of birth (99.3% of patients) and sex (98.9%) were well recorded and general medical subsidy and community services card eligibility was recorded for all patients.

There were varying degrees of recording of other data elements and wide variation between practices. Recording of the NHI ranged from between 77.0% and 99.5% of patients (mean 94.8%). The New Zealand Medical Council number of the patient's physician was recorded for between 0.0% and 100.0% of patients (mean 93.4%) and ethnicity recorded for between 0.7% and 99.4% of patients (mean 34.9%). Only eight of the 42 study practices recorded ethnicity for over 80% of their patients.

All consultation and prescription records could be linked to patient register details. Read codes were recorded for between 0% and 81.5% of consultations (mean 22.8%) with only six of the practices providing more than 50% of consultations with at least one associated Read code.

Although the name of the pharmaceutical prescribed was available on all prescription records, and form (94.8% of all prescription records), strength (99.5%), dosage (99.4%) and quantity of drug prescribed (85.5%) information was available for most scripts, in many cases these quantities had to be derived implicitly from other prescription data.

Discussion

This study indicates there are still difficulties involved in constructing a general practice national minimum data set. These primarily relate to the infrequent coding of patient ethnicity and morbidity and the ongoing administration and validation of patient registers by practice staff. Such problems may be overcome, however, and do not preclude the creation of a minimum data set. At least one practice, for example, coded ethnicity for 99% of patients, while another provided Read codes for 81% of consultations indicating that complete recording of these variables is achievable in the practice setting.

The utilisation of practical methods such as reduced Read code sets for recording patient morbidity requires support from government agencies, clinicians and practice staff. Morbidity recording in general practice is problematic and is the main hurdle in creating a general practice national minimum data set. In existing PMSs the diagnosis is primarily contained inside the textual consultation notes, although all practices had access to the full database of Read codes available via their PMS. Analysis of the text notes is extremely time consuming, as current information processing requires searching for strings relating to diagnoses and manual checking for accuracy. The Read Clinical Classification (RCC) was introduced in order to capture diagnoses in a computer readable code, but due to the lack of incentive for health professionals and the significant effort involved, most do not routinely code their consultations. A reduced Read code set for primary care has been proposed as a means of capturing diagnostic information.⁷ This would provide doctors with a more accessible and “user friendly” way of assigning Read codes and thus help to achieve a higher degree of consistency in morbidity coding by general practitioners nationwide.

One potentially useful approach may be that of Auckland Public Health, which since October 2000 has been collaborating with a group of Auckland general practitioners to develop a system (GPSURV) that monitors illness trends. GPSURV focuses on the incidence rates of four chronic diseases and four acute or episodic conditions. Evaluation of the system has revealed, however, that GPSURV encounters similar problems to other methods of morbidity recording including the incomplete recording of diagnosis codes, patient migration and changing providers.⁸

There has been a strong call for accurate data on patient ethnicity.⁴ This study reveals that ethnicity recording in general practice is still poor, although there are some encouraging signs of improvement. It has been proposed that general practice is ideally positioned to provide this health information by collecting ethnicity as part of patient registration if barriers to data collection are addressed.⁹ It is possible, using the NHI, to link ethnicity codes in the secondary care National Minimum Data Set to general practice patients who have accessed secondary health care in New Zealand. A study by the Wellington Independent Practitioner Association (WIPA) matched 99.2% of all records with an NHI number to an ethnicity code, although 54.7% of these were either “not stated” or “other”. When these unassigned codes were excluded

a similar ethnic distribution to that in the 1996 census was found.¹⁰ WIPA concluded that this is likely to be an improvement on base level of ethnicity coverage, but suggest that in the long term, direct recording of ethnicity in general practice is likely to be a more accurate and complete method of collecting this information.

The design of medical software packages requires active management of patient registers. Patients who move away, die or cease to be patients of the practice must be removed from the system by the practice staff. It is therefore difficult to make inferences about patients who did not consult. Furthermore, some patients may see more than one general practitioner over the study period and so some consultations may not appear on the database. Researchers using primary care data must take into account the fact that the data has been created by users concerned with the care of their patients and the administration of their business and, in the main, not concerned with the use of their records for general practice research.

Patient prescriptions are generated from the practice's PMS and may result from a consultation, a phone call from the patient or an after-hours visit by the doctor. Prescribing data is readily available and comprehensive data sets may be reviewed to compare treatment regimens and outcomes, monitor drug interactions and evaluate levels of usage and expenditure. As such, it should be included in a national minimum data set but standardisation of units for recording the quantity and strength of medicines and the treatment dosage should be applied.

All practices in this study used the same version of Health Technology software and thus provided data in the same format. A number of other PMS programs exist in New Zealand and all have different ways of structuring the medical record. Conversion programs would have to be run to incorporate these different file structures into any minimum data set. The Research Unit has written routines to handle this problem with two other PMS extraction programs and further development is currently underway.

In addition to the data required for this minimum data set model, clinical records on immunisations, laboratory results, ACC related contacts, screening information and patient measurements are collected by the Dunedin Research Unit. Much of this information is uncoded but can be retrospectively coded into usable form, however significant time and resources are required to achieve this.

Much of the data required to form a general practice minimum data set appears to be present in practices contributing data to the Dunedin Research Unit, although morbidity coding and ethnicity recording continue to be problematic. There are signs of improved ethnicity recording in some practices but whether this is so in the wider general practice community is unknown.

Author information: Jason Hall, Junior Research Fellow; Andrew Tomlin, Junior Research Fellow; Isobel Martin, Senior Lecturer and Director RNZCGP Dunedin Research Unit; Murray Tilyard, Elaine Gurr Professor of General Practice, Department of General Practice, Dunedin School of Medicine, University of Otago

Acknowledgements: We thank the participating general practitioners, practice staff and research officers who managed data collection. The RNZCGP Dunedin Research Unit receives financial support from the New Zealand Health Information Service.

Correspondence: Jason Hall, Department of General Practice, Dunedin School of Medicine, University of Otago, P O Box 913, Dunedin. Fax: (03) 477 2056; email: jhall@gpotago.ac.nz

References:

1. Tonks A, Smith R. Information in practice. *BMJ* 1996;313:438.
2. Tierney WM, McDonald CJ. Practice databases and their use in clinical research. *Stat Med* 1991;10:541–57.
3. Trevino FM. Uniform minimum data sets: in search of demographic comparability. *Am J Public Health* 1988;78:126–7.
4. Ministry of Health. From Strategy to Reality – The WAVE Project. Wellington: Ministry of Health; 2001
5. Tilyard MW, Dovey SM. The role of sentinel networks in general practice research in New Zealand and internationally. *NZFP* 2000;26:48–51.
6. Tilyard MW, Munro N, Walker SA, Dovey SM. Creating a general practice national minimum data set: present possibility or future plan? *NZ Med J* 1998;111:317–20.
7. Cunningham W, Tilyard MW. Making Read codes easy and useful in New Zealand General Practice. *NZFP* 2000;26:52–5.
8. Auckland District Health Board. Public Health Quarterly Report, Vol 8, Issue 1. Auckland: Auckland District Health Board; 2002.
9. McLeod D, Mansell J, Harris R et al. The collection of patient ethnicity data: a challenge for general practice. *NZFP* 2000;27:51–7.
10. Love T, Crampton P. Deprivation profiles in Wellington IPA practices. *NZFP* 2001;28:327–32.