

USING THE INTERNATIONAL CLASSIFICATION OF DISEASES WITH HOIST

David Muscatello and Susan Travis
*Epidemiology and Surveillance Branch
NSW Department of Health*

This article provides assistance with using the International Classification of Diseases (ICD) with the Health Outcomes and Information Statistical Toolkit (HOIST), which is a collection of databases (data warehouse) maintained by the Epidemiology and Surveillance Branch of the NSW Department of Health. The article is for readers who have access to HOIST or who want to learn how to use HOIST. Please note that HOIST is only available to staff of the NSW public health system.

Many datasets available in HOIST include variables that allow you to select data according to the disease or condition that led to a patient being hospitalised or that caused a person's death. These datasets include the Australian Bureau of Statistics (ABS) cause of death data collection (see the HOIST 'DEATHS' library), the NSW Inpatient Statistics Data Collection (see the HOIST 'ISC' library), and the NSW Emergency Department Data Collection (see the HOIST 'EDDC' library).

The ICD commenced in 1893 and has been revised periodically by the World Health Organization (WHO) to accommodate new developments in the understanding and awareness of diseases. The ICD currently in use is the ICD-10; that is, the tenth revision of the ICD. Many of the datasets in HOIST still use ICD-9 (the ninth revision), because there is a time lag between the introduction of a new revision and the release of data collections that use the new standard.

The WHO ICD-10 (international version) is currently used by the ABS to code causes of death in the ABS mortality data collection. For deaths registered between 1979 and 1998 inclusive, the WHO ICD-9 was used. For deaths registered in 1997 and 1998, causes of death were coded using both the WHO ICD-9 and WHO ICD-10. From 1999, only the WHO ICD-10 was used.

CLINICAL MODIFICATION OF THE ICD

For coding of hospitalisation statistics, Australia uses an adaptation of the WHO ICD that has been modified for clinical use, often known as the *Clinical Modification*. These modified editions are

continued on page 290

CONTENTS

- 289 Using the International Classification of Diseases with HOIST
- 293 Time trends in emergency department presentations of children with acute severe asthma in NSW
- 296 The impact of a general practice staffed casualty service on overall primary medical services
- 299 Survey of public swimming pools in the Mid Western Area Health Service: 2000–2001
- 302 Development of a child and youth health report card for Central Sydney, 2000
- 308 **FactSheet: Listeriosis**
- 309 **Communicable Diseases: November 2001**
- 309 Trends
- 309 Influenza in NSW, 2001
- 309 Quarterly report: Australian Childhood Immunisation Register

very similar to the WHO versions, but may contain subtle differences for specific codes (see the 'Take Care' section below). The latest version of the Australian edition is called the *ICD-10 Australian Modification* (ICD-10-AM). In Australia, the modified editions are developed and maintained by the National Centre for Classification in Health (NCCH) at www.cchs.usyd.edu.au/ncch.

Prior to July 1995, the United States version of the *ICD-9 Clinical Modification* (ICD-9-CM) was used for coding hospitalisation statistics in NSW. From July 1995, an Australian version of the ICD-9-CM was used. From July 1998, the first edition of the ICD-10-AM was used; and from July 2000, the second edition of the ICD-10-AM was used. Further editions are likely to be developed in the future.

The modified editions include additional sections on codes for clinical procedures and tables of drugs and chemicals used in coding. There are alphabetic indexes of diseases and drugs and chemicals that help in finding the correct code.

WHAT DO THE CODES LOOK LIKE?

ICD-9 and ICD-9-CM diagnosis codes are numeric and of the form *rrr*, *rrr.n*, or *rrr.nn*. The first three numbers, *rrr*, represent a specific diagnosis. For example, code *038* in the ICD-9-CM represents septicaemia. Additional numbers after the decimal point, *n* or *nn*, allow more detail. For example, *038.4* represents septicaemia due to 'other gram-negative organisms', and *038.42* represents septicaemia due to the *Escherichia coli* organism.

ICD-10 and ICD-10-AM diagnosis codes start with a letter of the alphabet followed by a two-digit number and are of the form *Xrr*, *Xrr.n* or *Xrr.nn*. There may be a further two numbers after the decimal point. For example, in the ICD-10-AM, *A41* represents 'Other septicaemia', *A41.5* represents septicaemia due to 'other gram-negative organisms', and *A41.51* represents septicaemia due to the *Escherichia coli* organism.

At the level of the first three characters (for example, *038* or *A41*), codes are usually consistent between the Australian versions that have been clinically modified and the WHO versions. It is only when you get to more detailed codes that differences may become apparent.

EXTERNAL CAUSES OF MORBIDITY AND MORTALITY

The ICD includes supplementary sections that permit the coding of circumstances that led to the occurrence of a disease or injury. These codes, known as 'external causes', are particularly useful for analysing the causes of injury or poisoning. In the ICD-9, external causes referred to injury and poisoning only. In the ICD-10, the meaning was broadened to include external causes of disease or disability (morbidity) and death (mortality). External causes in the ICD-9 had codes that started with 'E' and

were in the range *E800–E999*. In the ICD-10 these codes are constructed in the same way as diagnosis codes, but are in the range *V01–Y98* at the three-character level.

Like diagnosis codes, additional numbers after the decimal point provide more detail. For example, ICD-9 code *E965* represents assault by firearms or explosives, and *E965.0* represents assault by a handgun. The ICD-10 is different, with more detailed three-character codes, a fourth character to classify the place of occurrence of the event, and a fifth character for the activity the person was engaged in when the event happened. For example, ICD-10 code *X93* represents assault by handgun discharge, and *X93.72* represents assault by handgun discharge occurring on a farm while engaged in working for income.

For most deaths caused through injury or poisoning, the underlying cause of death code in the ABS cause of death collection records an external cause code obtained from the death certificate. For example, a person who died from head injuries as a result of a vehicle collision might have a code of *E812—Other motor vehicle traffic accident involving collision with another motor vehicle* as the underlying cause of death, not a diagnosis code for head injury such as *800—Fracture of vault of skull*.

Since 1997, the ABS has coded all causes contributing to the person's death (multiple cause coding) that were recorded on the death certificate. This meant that for a death caused through injury or poisoning, the diagnosis became available in addition to the cause. In HOIST, these additional codes are saved in separate datasets. Consult HOIST staff within the Epidemiology and Surveillance Branch at the NSW Department of Health if you intend analysing these (see the contact details below).

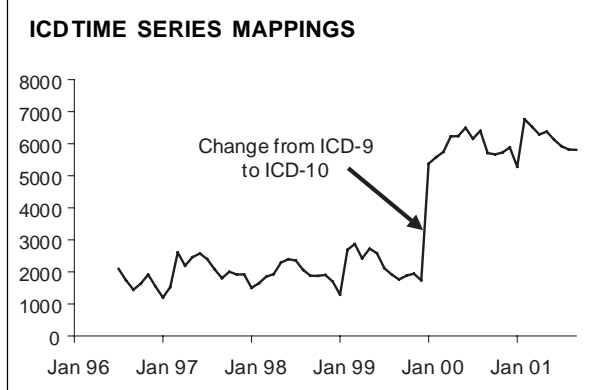
For underlying causes other than injury or poisoning, a diagnosis is coded as the underlying cause. For example, a person who died of a heart attack might have an ICD-9 diagnosis code of *410—Acute myocardial infarction*.

For hospital inpatient datasets, both diagnoses and external causes are coded and appear as separate variables in HOIST. For Emergency Department datasets, external cause codes are not available, only diagnoses.

TIME SERIES

When analysing data over a number of years, you might need to select HOIST variables that contain ICD codes that were coded using different versions of the ICD. You may also need to combine these into one variable for your analysis. At the 'Downloads' page of the NCCH Web site (www.cchs.usyd.edu.au/ncch), you can obtain details describing the concordance between the ICD-9-CM and ICD-10-AM (first edition) and between the first and second editions of the ICD-10-AM. These details are known as 'mappings'. You should check the relevant ICD books, however, because the mappings may not be quite what you expect. Mapping errors will probably show up

FIGURE 1



as discontinuities in the time series graph where the ICD changeovers occurred, as illustrated below.

In 1997, the ABS introduced an automated system for coding causes of death. This caused a break in time series for some causes because of different standards used in the automated system. With the introduction of the ICD-10 for coding deaths that were registered in 1999, the automated system was used to recode deaths registered in 1997 and 1998 using ICD-10. For more information about breaks caused by this changeover, see the ABS publication *Causes of Death Australia* (Catalogue no. 3303.0). More detailed information can be obtained directly from the ABS.

TAKE CARE

- You have to take care to use the version of the ICD that applies to the HOIST dataset you are using, or you can end up producing inaccurate reports. When analysing hospital statistics, you even have to be careful to use the correct edition of the clinical modification, because clinical modifications can change quite significantly between editions. Do not

use the clinical modification to select codes for cause of death analysis.

For example, in the ICD-9-CM, the external cause code *E850.1* represents accidental poisoning by methadone, whereas in the WHO ICD-9, *E850.1* represents accidental poisoning by salicylates.

- Be aware that the SAS format used in HOIST for each variable based on the ICD may not accurately reflect underlying codes. The formats are usually based on a particular version of the clinical modification and, while correct for most purposes, may on occasions give incorrect results. This is usually not a problem at the three-digit code level.
- Disease codes in hospital data are only as good as the clarity and completeness of the medical records from which they were coded, the tools available for helping clinical coders code the data, and the skill and experience of the clinical coder. Sometimes many specific categories of diseases end up in the 'other' category at the three, four, or five digit code level because not enough information is in the medical record. So you should err on the side of inclusiveness when selecting codes, but then report the resulting uncertainties when writing up the results.
- Disease codes in ABS death data are only as good as the accuracy of the death certificates from which they were coded. The ABS now uses an automated coding system for improving consistency and is rigorous in applying coding standards. Causes of death can be particularly unclear for older people who may have multiple contributing causes.
- Coding standards may not always be obvious. For example, under the ICD-9, deaths from opiate overdose were coded by the ABS mainly using codes *304.0—Morphine type drug dependence*; *304.7—Combinations of morphine type drug with any other*; and *305.5—Morphine type nondependent use of drugs*. The code *E850.0—Accidental poisoning by opiates and related narcotics* was only used when it was unclear whether the death resulted from

TABLE 1

EXAMPLES OF HOW TO SUBSET DATA IN HOIST ACCORDING TO DISEASES OR GROUPINGS OF DISEASES

Select what?	HOIST-SAS code
Deaths caused by stroke from the DEATHS.NSWRES dataset using the ICD-9	<i>where '430' <=: ICD9_DTH <=: '438'</i>
Hospitalisations with a principal diagnosis of cerebrovascular disease—stroke from the ISC.ISC9798 dataset using ICD-9-CM (Australian version)	<i>where '430' <=: ICD1 <=: '438'</i>
Hospitalisations with a principal diagnosis of cerebrovascular disease—stroke from the ISC.ISC9899 dataset using ICD-10-AM (1 st edition)	<i>where 'I60' <=: ICD10D1 <=: 'I69'</i>
Deaths from suicide by hanging from the DEATHS.NSWRES dataset using the ICD-9	<i>where ICD9_DTH =: 'E953.0'</i>
Hospitalisations for attempted suicide by hanging from the ISC.ISC9798 dataset using ICD-9-CM (Australian version)	<i>where EXTCAUS =: 'E953.0'</i>
Hospitalisations for attempted suicide by hanging from the ISC.ISC9899 dataset using ICD-10-AM (first edition)	<i>where ICD10EX1 =: 'X70'</i> <i>or if you want to be really thorough:</i> <i>where ICD10EX1 =: 'X70' or</i> <i>ICD10EX2 =: 'X70' or ICD10EX3 =: 'X70'</i>

TABLE 2

COMMONLY USED DATASETS AND VARIABLES IN HOIST THAT USE ICD CODING

Dataset–variable	Valid period	ICD version	Comments
DEATHS.NSWRES (deaths of NSW residents)			
ICD7_DTH	Deaths registered 1964–1967	ICD-7 (WHO)	Principal cause of death. Contains an external cause code for deaths from injury or poisoning unless drug addiction was mentioned on the death certificate for an accidental overdose). Check the relevant ICD book when selecting codes.
ICD8_DTH	1968–1978	ICD-8 (WHO)	
ICD9_DTH	1979–1998	ICD-9 (WHO)	
ICD10DTH	1997–	ICD-10 (WHO)	
ISC.ISC8889 onwards (NSW Inpatient Statistics Collection, by financial year of hospital separation)			
ICD1	Separations during 1988–89 to 1997–98	ICD-9-CM (The US CM was used prior to July 1995. The Australian version was used from July 1995.)	Principal diagnosis. Main admission diagnosis. This is the variable most used for analysis of diagnoses. Has been back-translated from ICD-10-AM for 1998–99, but probably better to use ICD-10-AM codes where available.
ICD2	1993–94 to 1997–98	ICD-9-CM as for ICD1	Stay diagnosis. Condition contributing to the length of stay. Usually the same as the principal diagnosis and not usually used.
ICD3-21	1993–94 to 1997–98	ICD-9-CM as for ICD1	Additional diagnoses. Not usually used because they may be comorbidities, pre-existing conditions, or historical conditions. There is no way to identify which.
EXTCAUS	1988–89 to 1997–98	ICD-9-CM as for ICD1	Main external cause of injury and poisoning (E-code). Has been back-translated from ICD-10-AM for 1998–99, but probably better to use ICD-10-AM codes where available.
EXTCAUS2	1996–97 to 1997–98	ICD-9-CM (Australian)	Additional external cause.
EXTCAUS3	1997–98	ICD-9-CM (Australian)	Additional external cause.
ICD10D1	Separations during 1998–99 (The 1st edition was used July 1998–June 2000. The 2nd edition was used July 2000.)	ICD-10-AM	Principal diagnosis.
ICD10D2	Separations during 1998–99	ICD-10-AM	Stay diagnosis. Condition contributing to the length of stay. Usually the same as the principal diagnosis and not usually used.
ICD10D3-21	1998–99	ICD-10-AM	Additional diagnoses. Not usually used because they may be comorbidities, pre-existing conditions, or historical conditions. There is no way to identify which.
ICD10EX1-ICD10EX3	1998–99	ICD-10-AM	External cause codes. The place of occurrence and activity codes were in the fourth and fifth characters of the code in the first edition of ICD-10-AM. In the second edition, new codes were added for these (Y92 and Y93).
PROC1	1988–89 to 1997–98	ICD-9-CM as for ICD1	Principal clinical procedure.
PROC2-20	1988–89 to 1997–98	ICD-9-CM as for ICD1	Additional clinical procedures.
MBS_EP1	1998–99 to present	ICD-10-AM	Principal clinical procedure.
MBS_EP2-20	1998–99	ICD-10-AM	Additional clinical procedure.
EDHIE.ED199607 onwards (NSW Emergency Department Data Collection, by month of attendance)			
ICD1	1996–97	ICD-9-CM (probably based on the Australian version of July 1996)	Primary Emergency Department diagnosis. Some hospitals do not provide data in ICD format and these will have blank codes. ICD-10-AM hasn't been implemented in Emergency Departments yet (at time of writing).
ICD2-5	1996–97	ICD-9-CM (probably based on the Australian version of July 1996)	Additional Emergency Department diagnoses. Best to not use as these may be even less consistently recorded than ICD1.

overdose. Check with the relevant authority (details below) if you are uncertain.

- When using the Emergency Department Data Collection remember that diagnoses are often coded by Emergency Department staff as they do their clinical work. Clinical coders are not generally involved in Emergency Department coding. Accuracy of the coding depends on the pressure the staff are under, the ease of use of their Emergency Department information system and other workplace issues that may result in inconsistent or inaccurate codes. For these reasons Emergency Department diagnoses should only be used for coarse analyses and should be considered as indicative only.
- Copies of both the WHO ICD and the clinical modification should be part of your library or you should have access to these, if you are planning anything other than the most superficial analyses.
- You can use the footnotes of graphs or the Methods section of *The Health of the People of NSW—Report of the Chief Health Officer* to find out what codes to use if you want to conduct analyses of the same disease groups.

SAS EXAMPLES

Table 1 gives examples of how to create a subset of data in HOIST according to diseases or groupings of diseases. Table 2 lists the commonly used datasets and variables in HOIST that use ICD coding.

The : operator in SAS is useful for selecting ICD groupings. When you put the : after a SAS operator,

only the number of characters that you typed are compared. For example, if you type =: '038' then all codes beginning with 038 are captured, including '038', '038.4' and '038.42'. It is wise to use the : operator even when you want to select very specific codes also, because sometimes there can be sub-codes that you may not be aware of.

WHERE TO GET HELP

For hospital coding practices contact:

- your local clinical coder or Health Information Manager;
- Senior Clinical Data Consultant, NSW Department of Health by phoning (02) 9391 9684;
- the National Centre for Classification in Health at www.cchs.usyd.edu.au/ncch.

For cause of death coding standards, contact the Vital Statistics Unit of the Australian Bureau of Statistics by phoning 1800 620 963.

For help with HOIST, contact Alan Willmore, Epidemiology and Surveillance Branch, NSW Department of Health, by emailing awill@doh.health.nsw.gov.au or by phoning (02) 9391 9226.

ACKNOWLEDGMENTS

The authors would like to thank staff of the Australian Bureau of Statistics Cause of Death National Project Centre and Joanne Chicco, Senior Clinical Data Consultant, NSW Department of Health for their helpful comments on the draft manuscript. ☞

TIME TRENDS IN EMERGENCY DEPARTMENT PRESENTATIONS OF CHILDREN WITH ACUTE SEVERE ASTHMA IN NSW

Doug Lincoln and David Muscatello
Epidemiology and Surveillance Branch
NSW Department of Health

It has been observed that the number of admissions of children to NSW hospitals for asthma increases sharply during the month of February.¹ Recent anecdotal evidence suggested that this increase occurred again in 2001. Of greater concern, however, is the observation that the number of *severe* asthma episodes in children also rose sharply. This prompted an investigation to determine whether the reported peak was supported by the data routinely collected by the NSW Department of Health.

The data confirmed the report showing a seasonal increase of asthma presentations, and also highlighted an increasing trend in the proportion of asthma presentations that were severe. The Emergency Department Data Collection (EDDC) was the most up-to-date source of

routinely collected data available to investigate this phenomenon.² Despite some shortcomings, the results exemplify the usefulness of the EDDC as a public health surveillance tool. This article describes the methods used and the trends observed using the data from the EDDC.

METHODS

The EDDC is maintained by the NSW Department of Health and collects information on Emergency Department (ED) activity at 54 hospitals across NSW. It covers most urban hospitals and regional base hospitals. In this article, 'participating EDs' will refer to the EDs of these 54 hospitals. Information collected includes the patient's demographics, primary diagnosis, date and time of presentation, and the clinical response of the participating ED. Information collected at the participating EDs is periodically uploaded to a central data repository held at the NSW Department of Health. At