NSW Oncology Group
Upper Gastrointestinal
Minimum Data Set Extension

Data Dictionary

DRAFT: Version 1

July 2007

Publication Details

Cancer Institute NSW

ISBN:
ISSN:

Suggested citation: NSW Cancer Institute.
NSW Upper Gastrointestinal Data Dictionary Version 1.0,
Sydney, Australia, 2007.

Inquiries:

Any inquiries about this publication should
be directed to:

Manager, Cancer Information Systems
Cancer Institute NSW
Level 1, Biomedical Building
Australian Technology Park
Eveleigh NSW 2015
Tel (02) 8374 5687
Fax (02) 8374 5700

Copyright information:

The work is copyright. It may be reproduced
in whole or in part for study or training purposes
subject to the inclusion of an acknowledgment
of the source and no commercial usage or sale.

© Cancer Institute NSW 2007

Publisher:
Published by the Cancer Institute NSW

Contents

1. Document Version Control ... 4
2. Introduction ... 5
3. Upper Gastrointestinal Cancers ... 6
4. List of Abbreviations ... 7
5. Data Dictionary Format Guide ... 8
   I. Intent of Treatment ... 10
   II. Date of First Symptoms ... 12
   III. Duration of Symptoms ... 14
   IV. Pain at Presentation ... 16
   V. Nutritional Status (Automated Risk Score) ... 18
   VI. Complications ... 21
   VII. Diagnostic Tests ... 23
   VIII. Reason for Palliative Care Referral ... 26

1. Document Version Control

Version	Date Issued	Change Description
0.1	6/02/2007	1st Draft - Developed by Ms Christine Erratt based on work by the Upper GI NSW Oncology Group
0.2	27/03/2007	2nd Draft - The following values were revised by the NSWOG Upper GI, Intent to treat, Referral to Palliative Care, Nutritional Status, Pain at Presentation, Rearrangement of Items.

2. Introduction

Population-based cancer registries in each Australian state and jurisdiction provide comprehensive information on cancer incidence. By matching, verifying and registering each case, incidence of each cancer type can be mapped by area of residence, age, sex and country of birth. Death notifications (and cause of death) are also matched and provide the definitive mortality and survival rates for cancer in NSW.

To enhance this epidemiological information, *clinical cancer registries* are designed to add the dimensions of stage, treatment and quality of care, allowing analyses of patterns of cancer care against best-practice guidelines. The Institute is funding Area-based clinical cancer registries in six Area Health Services.

By describing cancer stage and actual surgical, radiation and chemotherapeutic interventions, Areas and tumour streams can monitor access and quality of care. However, specific quality of care indicators for each cancer type requires collection of a more specific subset of data items. For instance, for breast cancer the receptor status (oestrogen, progesterone and HER2), together with disease stage dictates the appropriate drug treatment options. Other data items will support better monitoring of supportive care or enhance the prognostic value of the core dataset.

The NSW Oncology Group (NSWOG) was established by the Cancer Institute NSW and comprises cancer specialist doctors and nurses, consumers and patients. The aim of NSWOG includes the identification of best practice care guidelines, and of the data needed to monitor and improve cancer outcomes in NSW. NSWOGs also promote sub specialised training and education for each type of cancer, and clinical trials.

The NSW Minimum Dataset for Clinical Cancer Registration is being collected in many public hospitals in NSW. The *core cancer dataset* describes cancer type, stage, treatment and quality of care for each cancer patient. Concurrently, NSW Oncology Group is working to identify succinct *dataset extensions* as statewide standards, to complement the core dataset with additional measures and indicators specific to tumour streams.

Only data elements specific to upper gastrointestinal cancers are presented in this data dictionary. Additional data elements for all cancers are covered by the NSW Cancer Registry Data Dictionary and the NSW Clinical Cancer Data Dictionary.

The review of the data dictionary will be conducted after the collection has been piloted for a period of time so that the decisions concerning changes to the dataset can be based on feasibility, usability and experience. It is intended that this data dictionary defines the disease-specific dataset and will be used by a variety of NSW Clinical Cancer Registry stakeholders.

Specifically this data dictionary will have relevance to:

- Upper GI Cancer Clinicians
- Clinical Cancer Registry Staff
- Analysts of Clinical Cancer Registry data

The Data Set Specification, Cancer (Clinical): National Health Data Dictionary Version 12 Supplement published by the Australian Institute of Health and Welfare (AIHW) and the NSW Clinical Cancer Data Collection for Outcomes and Quality Data Dictionary published by the NSW Health Department were both reviewed prior to producing this publication.

This document was written by Ms Christine Erratt based on the work by the Upper GI NSW Oncology Group (NSWOG). Dr Ross Smith and Dr Jane Ingham, in particular, provided detailed support and review of some items.

3. Upper Gastrointestinal Cancers

The table below shows the Primary Site of Cancer ICD10AM 5th Edition codes that trigger the reporting of an Upper GI Minimum Data Set.

- C15 Malignant neoplasm of oesophagus
- C15.0 Malignant neoplasm of cervical part of oesophagus
- C15.1 Malignant neoplasm of thoracic part of oesophagus
- C15.2 Malignant neoplasm of abdominal part of oesophagus
- C15.3 Malignant neoplasm of upper third of oesophagus
- C15.4 Malignant neoplasm of middle third of oesophagus
- C15.5 Malignant neoplasm of lower third of oesophagus
- C15.8 Overlapping malignant lesion of oesophagus
- C15.9 Malignant neoplasm of oesophagus, unspecified
- C16 Malignant neoplasm of stomach
- C16.0 Malignant neoplasm of cardia
- C16.1 Malignant neoplasm of fundus of stomach
- C16.2 Malignant neoplasm of body of stomach
- C16.3 Malignant neoplasm of pyloric antrum
- C16.4 Malignant neoplasm of pylorus
- C16.5 Malignant neoplasm of lesser curvature of stomach, unspecified
- C16.6 Malignant neoplasm of greater curvature of stomach, unspecified
- C16.8 Overlapping malignant lesion of stomach
- C16.9 Malignant neoplasm of stomach, unspecified
- C23 Malignant neoplasm of gallbladder
- C24.0 Malignant neoplasm of extrahepatic bile duct
- C24.1 Malignant neoplasm of ampulla of Vater
- C24.8 Overlapping malignant lesion of biliary tract
- C24.9 Malignant neoplasm of biliary tract, unspecified
- C25.0 Malignant neoplasm of head of pancreas
- C25.1 Malignant neoplasm of body of pancreas
- C25.2 Malignant neoplasm of tail of pancreas
- C25.3 Malignant neoplasm of pancreatic duct
- C25.4 Malignant neoplasm of endocrine pancreas
- C25.7 Malignant neoplasm of other parts of pancreas
- C25.8 Overlapping malignant lesion of pancreas
- C25.9 Malignant neoplasm of pancreas, part unspecified

4. List of Abbreviations

AIHW	Australian Institute of Health and Welfare
ARS	Automated Risk Score
BPI	Brief Pain Inventory
CINSW	Cancer Institute New South Wales
Ed	Edition
GI	Gastrointestinal
HDD	Health Data Dictionary
ICD-10	International Statistical Classification and Related Health Problems, Tenth Revision
MDS	Minimal data set
NHDD	National Health Data Dictionary
NSWOG	New South Wales Oncology Group
SGA	Subjective Global Assessment

5. Data Dictionary Format Guide

Each data item is described in terms of its defining characteristics and its physical representation. In addition to this, certain administrative information is provided to inform users of the sources and the currency of the version of the individual item. The components included under these section headings are based on the NHDD standard, as described below:

Heading	Description
Defining Attributes	
Definition	A statement that expresses the essential nature of a data element and its differentiation from all other data elements.
Coverage	A description of the circumstances under which the data item should be collected and reported.
Guide for Use	
Data Domain	The set of possible values for the data item. This may take the form of a code set, or a description of the possible values. Domain values are only specified where size of the code set is small enough to be reasonably reproduced in the document. In other instances the domain may be indicated by reference to a source document.
Domain Definitions	The definitions of each domain category within the classification, where such definitions are warranted – that is more information that the domain descriptor is required to fully understand what is captured with the domain value.
Clarifying Points	These are comments designed to assist in further defining aspects of the data domain.
Collection Methods	This provides important comments concerning the actual capture of data for the particular data element.
Screen Prompts	This is suggested terminology to use in computer applications.
Validation Rules	These are included to assist in reducing input error. Where validation rules are known to exist, they have been included to assist with the programming.
Justification	The reason for collecting this data element.
Representation	
Data Element Type	There are four types of data elements, and this describes which of the element is. Definitions of each type are provided below. *Data Concept* - a concept which can be represented in the form of a data element, described independently of any particular representation. For example, hospital 'admission' is a process, which does not have any particular representation of its own, except through data elements such as 'Date of Admission', 'mode of admission' etc. *Data Element* – a unit of data for which the definition, identification, representation and permissible values are specified by means of a set of attributes. *Derived Data Element* – a data element whose values are derived by calculation from the values of other data elements. For example the data element 'length of stay' is

Heading	Description
	derived by calculating the number of days from the 'Date of Admission' to the 'Date of Separation' less the number of 'total leave days'. *Composite Data Element* – a data element whose values represent the grouping of the values of other data elements in a specified order.
Data Type	The type of symbol or character, or other designation used to represent the data element. For example numeric, alphanumeric, alphabetic or integer.
Form	Describes whether the valid values for the data item take the form of a code set, free text. If the form is described as "Code" the relevant code set or sets will be specified in the Domain section.
Minimum Size	The minimum number of characters allowable to represent the data element.
Maximum Size	The maximum number of characters allowable to represent the data element.
Layout	A generic example of what the data element should look like in the unit record. For example, dates should be represented in the format of DDMMYYYY where DD represents, the day, MM represents the month, and YYYY represents the four-digit numeric for the year. "N" is used to represent numeric values and "A" is used to represent alphabetic and alphanumeric values (the Data Type indicates whether it is alphabetic or alphanumeric).
Administrative Information	
Version	This is the version number of the individual data element as it exists in the New South Wales Health Data Dictionary only. The version number may differ from the version number of the NSW HDD publication, as data elements may be revised independently of the periodic review of the document.
Effective Date	The date from which this version of the data element is to be used for reporting.
References	
Related Elements	Data elements that have some direct relationship with the data element being described.
References	Documents listed here have been used as references when designing the specified item. The item as it is presented in the NSW HDD is not necessarily identical to the item in the source document. The name of the organisation(s) that developed the source document(s) or provided advice on the data item.

Intent of Treatment

Defining Attributes

Definition:

The intention of the primary phase of treatment for cancer for the particular patient presenting for this episode.

Coverage:

This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
0	Did not have treatment
1	Curative
2	Non-curative or palliative
9	Not stated

Domain Definitions:

Code	Definition
0	**Did not have treatment:** This is used when the patient did not have treatment as part of the management plan.
1	**Curative:** This is used when treatment is given for control of the disease.
2	**Non-curative or palliative**: This is used when the cure is unlikely to be achieved and treatment is given primarily for the purpose of symptom control. Other benefits of the treatment are considered secondary contributions to the patient's quality of life.
9	**Not stated:** This is used when the patient had treatment for cancer but the intention was not stated.

Clarifying Points: n/a

Collection Methods: This information should be sought from the patient's medical record notes.

Validation Rules: Must = 0, 1, 2 or 9

Justification: This information is collected for the purpose of:

- correlating outcome with the intent of treatment

Representation

Data Element Type: Data Element

Data Type: Numeric

Form: Code

Minimum Size: 1

Maximum Size: 1

Layout: N

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data: n/a

References: Australian Institute of Health and Welfare (2004), *Data Set Specification Cancer (Clinical) National Health Data Dictionary version 12 Supplement*, Canberra: AIHW.

Date of First Symptoms

Defining Attributes

Definition: When the patient first experienced symptoms of their upper gastrointestinal cancer.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
DD	Day of month (use leading zeros for 1 to 9 e.g. “01”, “02”)
MM	Month of the year (use leading zeros for 1 to 9 e.g. “01”, “02”)
YYYY	Year (use 4 digit format e.g. “2005”, “2006”)

Domain Definitions: ddmmyyyy

Clarifying Points: These symptoms are usually the reasons for presentation to a cancer specialist

Collection Methods: This information should be sought from the patient’s medical record notes, referral letter or medical practitioner.

Validation Rules: Date of first symptoms must be:

- greater than or equal to Date of birth
- less than or equal to Date of diagnosis
- less than or equal to Date of death

Justification: This information is collected for the purpose of:

- Analysis of health service usage, epidemiological studies and monitoring of specific disease entities and conditions.

Representation

Data Element Type: Data Element

Data Type: Date

Form: Date

Minimum Size: 8

Maximum Size: 8

Layout: DDMMYYYY

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data:

- Duration of Symptoms

References:

Australian Institute of Health and Welfare (2004), *Data Set Specification Cancer (Clinical) National Health Data Dictionary version 12 Supplement*, Canberra: AIHW.

Developed by the NSW Upper Gastrointestinal Oncology Group (NSWOG Upper GI). July 2007.

NEW DATA ITEM

Duration of Symptoms

Defining Attributes

Definition: Duration of symptoms is the period from the date of first symptoms of their upper gastrointestinal cancer until the date of its diagnosis.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
0	< 2 weeks
1	>= 2 weeks
2	>= 3 weeks
3	>= 1 month
4	>= 2 months
5	>= 3 months
6	>= 6 months
7	>= 9 months
8	>= 1 year
9	Not stated or unknown

Domain Definitions: Select the description that best describes the patient's duration of symptoms.

Clarifying Points: n/a

Collection Methods: This information can be sought from the patient's medical record or the Referral letter.

Validation Rules: Must = 0, 1, 2, 3, 4, 5, 6, 7, 8 or 9

Justification: This information is collected for the purpose of:

- analysis of health service usage, epidemiological studies and monitoring of specific disease entities and conditions.

Representation

Data Element Type: Data Element - derived

Data Type: Numeric

Form: Code

Minimum Size: 1

Maximum Size: 1

Layout: N

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data:

- Date of first Symptoms

References: Developed by the NSW Upper Gastrointestinal Oncology Group (NSWOG Upper GI). July 2007.

NEW DATA ITEM

Pain at Presentation

Defining Attributes

Definition: The pain experienced by the patient at presentation to cancer specialist.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
0	No pain
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	Pain as bad as you can imagine

Domain Definitions: The BPI scale rates pain on a scale of 0-10.

Clarifying Points: Pain experienced by the patient should be rated by the pain scale of 0-10 that best describes their pain at its worst in the last 24 hours.

Collection Methods: This information should be sought from the patient's medical record notes or medical practitioner.

Validation Rules: Must = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10.

Justification: This information is collected for the purpose of:

- survival analysis adjusted by stage at diagnosis and distribution of cancer cases by type and stage.

Representation

Data Element Type: Data Element

Data Type: Numeric

Form: Code

Minimum Size: 1

Maximum Size: 2

Layout: NN

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data:

- Duration of Symptoms

References:

Cleeland, CS (1991), The Brief Pain Inventory. [Online]. Available: http://www.mdanderson.org/pdf/bpilong.pdf [accessed 10 July 2007].

Nutritional Status (Automated Risk Score)

Defining Attributes

Definition: The nutritional status (automated risk score) of the patient at the Date of Consultation with Cancer specialist.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
0	Not assessed
1	Score of 1
2	Score of 2
3	Score of 3
4	Score of 4
5	Score of 5
6	Score of 6
9	Unknown

Domain Definitions:

Code	Definition
0	**Not assessed**: Patient was not assessed.
1	**Score of 1**:
2	**Score of 2**:
3	**Score of 3**:
4	**Score of 4**:
5	**Score of 5**:
6	**Score of 6**: Patient was assessed and had any of the following, occurrence of a wound, poor oral intake, malnutrition-related admission diagnosis.
9	**Unknown**: It is unknown whether the patient had been assessed for a nutritional score.

Clarifying Points:

The ARS_B is derived by adding 1 for each blood value which did not reach the threshold values (31.5 g/L for plasma albumin, 1,202 $X10^9$/L for lymphocyte count and 99.5 g/L for haemoglobin). The maximal ARS_B was 3.

The full ARS included the score for ARS_B and an extra score of 1 if the patient was assessed to have any of; occurrence of a wound, poor oral intake or malnutrition-related admission diagnosis. The maximal score for ARS is 6.

Collection Methods:

This information should be sought from the patient's attending medical clinician or medical practitioner.

Validation Rules:

Must = 0, 1, 2, 3, 4, 5, 6 or 9.

Justification:

This information is collected for the purpose of:

The Automated Risk Score (ARS) has been shown to be a very useful predictor of complications It has a higher sensitivity and smaller negative Likelihood Ratio than the Subjective Global Assessment (SGA).

Representation

Data Element Type: Data Element

Data Type: Numeric

Form: Code

Minimum Size: 1

Maximum Size: 1

Layout: N

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data: n/a

References: Brugler I, Stankovic AK, Schlefer M, Berstein L, (2005), 'A simplified screen for hospitalised patients using readily available laboratory and patient information,' *Nutrition,* vol. 21, no. 6, pp. 650-658.

Complications

Defining Attributes

Definition: Post operative complications of the patient's cancer treatment.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
0	No surgery
1	Infection or inflammation
2	Mechanical
3	Haemorrhage
4	Peritonitis
5	Unplanned return to theatre
6	Leak of anastomosis
8	Specified, not elsewhere classified
9	Unknown

Domain Definitions: See data domain.

Clarifying Points: This data item refers to post-operative complications only.

Collection Methods: This information should be sought from the patient's medical record notes.

Validation Rules: This is a one-to-may data item. That is, a patient record could hold: either 0 or 9 or any combination of 1-6 and 8.

Justification: This information is collected for the purpose of:

- survival analysis, adjusted by stage at diagnosis and distribution of cancer cases by stage and type and
- quality of care outcomes

Representation

Data Element Type: Data Element

Data Type: Numeric

Form: Code

Minimum Size: 1

Maximum Size: 1

Layout: N

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data:

- Duration of Symptoms
- Nutritional Status
- Pain at Presentation

References:

National Centre for Classification in Health (2006), *The International Statistical Classification and Related Health Problems, Tenth Revision, Australian Modification (5th ed)*, Sydney: NCCH.

Diagnostic Tests

Defining Attributes

Definition: All tests undertaken that contribute towards the patient's cancer diagnosis.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
00	None
01	Histology
02	Cytology including FNA
03	X-Ray with or without contrast
04	Ultrasound
05	CT
06	ERCP
07	MRI
08	PET
98	Other imaging (includes angiography)
99	Unknown

Domain Definitions:

Code	Definition
00	None: no diagnostic tests were conducted
01	Histology:
02	Cytology: Including Fine Needle Aspiration (FNA)
03	X-Ray: includes with or without contrast
04	Ultrasound:
05	CT: Computed Tomography
06	ERCP: Endoscopic Retrograde Cholangiopancreatography
07	MRI: Magnetic Resonance Imaging
08	PET: Positron Emission Tomography
98	Other imaging: includes angiography
99	Unknown: it is unknown and not stated whether any tests were conducted

Clarifying Points: n/a

Collection Methods: This information should be sought from the patient's medical record notes or attending medical clinician.

Validation Rules: This is a one-to-many data item. That is, a patient record could hold: either 0 or 9 or any combination of 1 – 8.

Justification: This item is collected for the purpose of analysing pattern of care against best-practice guidelines.

Representation

Data Element Type: Data Element

Data Type: Numeric

Form: Code

Minimum Size: 2

Maximum Size: 2

Layout: NN

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data: n/a

References: Australian Institute of Health and Welfare [METeOR Metadata Online Registry] [Online] 1 March 2005 – last updated. Available: http://meteor.aihw.gov.au/content/index.phtml/itemId/270511 [10 July 2007].

Developed by the NSW Upper Gastrointestinal Oncology Group (NSWOG Upper GI). July 2007.

NEW DATA ITEM

Reason for Palliative Care Referral

Defining Attributes

Definition: Reason the patient was referred to the palliative care during their course of treatment.

Coverage: This item should be reported for:

☑ All Upper Gastrointestinal cancers

Guide for Use

Data Domain:

Code	Description
0	Not referred
1	Assessment
2	Care Planning
3	Implementation of Care
4	Evaluation
5	Pain assessment and control
6	Other
8	Not Stated
9	Not known

Domain Definitions:

Code	Definition
0	**Not referred:** Patient was not referred to Palliative care.
1	**Assessment:** Patient was referred to a Palliative Care Physician for a palliative care assessment.
2	**Care Planning:** Patient was referred to a Palliative Care Physician for a palliative care planning.
3	**Implementation of Care:**
4	**Evaluation:**
5	**Pain assessment and control**: includes symptom control
6	**Other**: Includes all those patients referred to a palliative care service for a reason that does not fit into codes 1-5 above.
8	**Not Stated:** The reason for palliative care referral is not recorded.
9	**Unknown**: The reason for palliative care referral is unknown.

Clarifying Points: n/a

Collection Methods: This information should be sought from the patient's medical record or referral letter.

Validation Rules: This is a one-to-many data item. That is, a patient record could hold: either 0, 8 or 9 or any combination of 1 – 6.

Justification: This information is collected for the purpose of:

- Monitoring levels and appropriateness of referrals to palliative care service.

Representation

Data Element Type: Data Element

Data Type: Numeric

Form: Code

Minimum Size: 1

Maximum Size: 1

Layout: N

Administrative Information

Version: 1

Effective date: 1 August 2007

Changes:

Related Information

Related Data: n/a

References:

Australian Institute of Health & Welfare, (2005), *National Palliative Care Performance Indicators: Report on the National Palliative Care Performance Indicator Data Collection 2006*. Canberra: AIHW.

Developed by the NSW Upper Gastrointestinal Oncology Group (NSWOG Upper GI). July 2007.

NEW DATA ITEM