The Royal College of Pathologists
Pathology: the science behind the cure

Supplement for authors writing cancer datasets

1 Background

The College publishes its cancer datasets electronically and is encouraging software companies to develop products for local pathology systems, so there is a need for a consistent format and style across all datasets. This document describes the 'College style' and the process that must be applied to new datasets and to existing datasets as they are revised. As the practical issues within the datasets are likely to vary from site to site, dialogue between the College and authors is essential to ensure that this style is achieved. Advice is available from the Chair of the College's Working Group on Cancer Services (WGCS) and from the Clinical Effectiveness Department (CED).

The series is entitled 'Standards and Datasets for Reporting Cancers'. The term 'minimum dataset' is no longer used. This does not mean the series now comprises 'maximum' datasets, but authors are nonetheless encouraged to describe the key features required for optimal patient management.

2 Purpose of cancer datasets

Each dataset contains **core data items**, which are required by the National Cancer Outcomes and Services Dataset (COSD) for the staging and grading of cancers, and which published evidence indicates are required for optimal patient management and prognosis. Other **non-core data items** that fall outside the core definition are also described. These may be included to provide a comprehensive report or to meet local clinical or research requirements. All items should be clearly defined to allow the unambiguous recording of data.

Authors should be aware that datasets are likely to be read by trainees, general pathologists, specialist pathologists, clinicians and service commissioners. Secondary use by Cancer Registries should be noted, and the health benefits associated with implementation of the guidelines should be emphasised. Standardised cancer reporting ensures that clinicians have all of the relevant pathological information required for tumour staging, management and prognosis. Failure to standardise cancer reporting carries the risk of incorrect or incomplete cancer diagnosis, which may result in the provision of incorrect advice to the clinicians who have direct responsibility for patient care, resulting in incorrect management decisions to the detriment of patient health. Clinicians who treat cancer patients will benefit from having tumour-specific prognostic information and thereby ensure that patients are assigned to appropriate treatment regimens in a timely fashion. Collection of standardised cancer-specific data provides information for healthcare providers and epidemiologists, and facilitates international benchmarking and research. The dataset should seek to deliver guidance with a reasonable balance between the differing needs and expectations of the different groups. They are not intended to cover all aspects of service delivery and reference should be made, where possible and appropriate, to guidance on other aspects of delivery of a tumour-specific service, e.g. cytology and molecular genetics.

3 Role of the Working Group on Cancer Services (WGCS)

The WGCS:

- commissions the development of new datasets and the revisions of existing datasets from lead authors and the writing group
- must agree a timetable for producing the dataset with the lead author (see also general process shown in the appendix)
- is involved in the review of draft datasets
- should keep the Clinical Effectiveness Department (CED) informed about changes in pathology practice and the delivery of cancer services that might impact upon the content of datasets.

4 Standard structure of the dataset document

All datasets should follow the numbered structure outlined below. If any sections are not required in an individual dataset, a statement to that effect should be included.

1 The dataset should include:

- title
- date
- details of author/s (including their appointments and brief details of their expertise)
- contents page.

2 Foreword, which includes:

- a list of the stakeholders consulted (if and where appropriate)
- a brief description of the methods used to obtain the evidence for the dataset
- a brief description of the method used to evaluate the strength of the supporting evidence (classified in accordance with modified Scottish Intercollegiate Guidelines Network [SIGN] guidance), and, if relevant, the ways in which disagreements were resolved.

3 Introduction, to include**:**

- an explanation of the importance and clinical application of the dataset
- information highlighting any site-specific issues
- for new datasets, methods used to pilot or validate the dataset
- validation of any proposed changes or revisions of datasets
- a statement indicating the target users of the dataset (the primary users of the cancer datasets are cellular pathologists and trainee pathologists and, on their behalf, the suppliers of IT products to laboratories). Secondary use by surgeons and oncologists, cancer registries and the NCIN should be noted, and the health benefits associated with implementation of the guidelines should be emphasised. *[The Clinical Effectiveness Department will insert a generic statement into the introduction about the target users and health benefits of this guideline].*

4 Clinical information required on the specimen request form

5 Preparation of specimens before dissection

6 **Specimen handling and block selection** (relatively brief, evidence-based summary where applicable).

7 **Core data items** (an evidence-based list of items that are essential for prognosis or management)

7.1 Clinical (usually the clinical information related to the procedure and site of the tumour)

7.2 Pathological (to include macroscopic and microscopic items; do not ask for the same data under both macro and micro headings).

8 **Non-core data items**, to include:

- preferences of individual laboratories
- items for clinical research
- supplementary information that may contribute to prognosis, management or treatment decisions in individual cases.

9 **Diagnostic coding and staging** (recommended protocols, e.g. TNM, SNOMED), to include:

- a clear statement as to which version of the TNM classification is recommended.
- a standard statement on the use of SNOMED coding (see section 7 below).

10 **Reporting of small biopsy specimens**

11 **Reporting of frozen sections**

12 **Specific aspects of individual tumours not covered elsewhere**

13 **Criteria for audit of the dataset** (to include suggestions for criteria against which the successful implementation of the dataset could be audited. These might include, for example, audit of the completeness of recording of all data items in histopathology reports, audit of turnaround times, compliance with the key performance indicators (KPIs) of the College, and audit of lymph node retrieval, where relevant). *[The Clinical Effectiveness Department will insert a generic standard paragraph of the KPIs recommended by the RCPath].*

14 **References**

15 **Appendices**, in the following sequence:

1. TNM or other classification system for sites covered by the dataset
2. SNOMED CT or SNOMED T and M codes, also P codes if applicable
3. Draft specimen request form (if applicable); this might include diagrams and site-specific clinical data in a form that is suitable for local modification if necessary
4. Reporting proformas (see section 5 below)
5. Summary table – explanation of levels of evidence
6. AGREE compliance monitoring sheet.

5 Reporting proforma structure

The reporting proformas have four key purposes:

- to act as an *aide memoire* for reporting pathologists
- they may be incorporated into a surgical report to ensure all core data items are reported
- to provide a concise defined list of core data types
- to form the basis for the electronic collection of data by COSD – to support with data collection. The proformas are also given in list format to assist system suppliers with system design.

It is therefore important to remember these key principles:

- make the data items as clear as possible for non-pathologists to read; the proformas may be read by other clinicians (as part of a report) or by non-clinicians (particularly developers of Laboratory Information Systems)
- do not include non-core data items in the reporting proformas
- avoid unnecessary repetition of data items
- avoid open, free-text responses where possible; defined checkbox-type lists are much better for data collection and subsequent analysis
- ensure all checkboxes for a positive response have an associated checkbox for a negative response, rather than relying on no response to imply a negative response
- provide enough options in responses to a data item so that an answer should always be given by the pathologist; this may mean including options such as 'Cannot be assessed', 'Not applicable' and 'Uncertain'
- use the standard header for key demographic data (see 4.1 below).
- use the list of 'Standardisation of terminology used for responses for core data items in the reporting proformas of cancer datasets', as agreed by the WCGS (see 4.2 below).

5.1 Standard header for demographic data in reporting proformas (to copy and paste)

Reporting proforma for [title]

Surname: ……Forenames:……..Date of birth: ……Sex:……..
Hospital…….……Hospital no: ……..……NHS no:……..
Date of surgery: ……..…Date of report authorisation: ……Report no:…….
Date of receipt:……...Pathologist:…….……Clinician:…….....

5.2 Standardisation of terminology used for responses for core data items in reporting proformas

With few exceptions, when the following core data items appear in the reporting proformas of the cancer datasets, checkboxes should be provided to include the following standard choices:

Lymphovascular and perineural invasion
Present
Not identified
Uncertain
Cannot be assessed

Resection margins
Involved
Not involved
Uncertain
Not applicable

Necrosis
Present
Not identified
Uncertain
Cannot be assessed

Lymph nodes
Involved
Isolated tumour cells identified
Not involved
Cannot be assessed
Not applicable

Metastatic spread
Present
Not identified
Uncertain
Cannot be assessed

6 Supporting evidence

6.1 Provision of supporting key evidence

The evidence underpinning core data items that are required for cancer staging, grading and tumour classification has already been evaluated by international organisations such as UICC (Union for International Cancer Control), AJCC (American Joint Committee on Cancer), FIGO (International Federation of Gynecology and Obstetrics), WHO (World Health Organization), etc. Authors do not have discretion to alter internationally established staging, grading and classification systems, and are not expected to re-evaluate the supporting evidence used by international bodies in formulating these systems.

6.2 Evaluation of the strength of supporting evidence: the view of the WGCS

However, given the above, authors are expected to cite key evidence that underpins the staging and grading systems for educational purposes and to inform clinical discussion. The grade (quality) of any evidence is a continuum (low to high), so any categorisation is arbitrary. The College aims for an approach that indicates whether evidence is either of high quality and unlikely to change in the near future, or is more uncertain and thus likely to be subject to future modification.

Importance of recommendations

For the cancer datasets, the importance of the recommendations is indicated by whether a data item forms part of the core data (strongly recommended) or the non-core data (less strongly recommended). The WGCS considers that, for simplicity, further categorisation of the importance of the recommendations is not required.

Quality of evidence

Any evidence evaluated will normally be directly applicable to the relevant cancer type considered in the dataset. For some rare cancers, extrapolation of evidence from more common cancers may be justified in order to provide a uniformity of approach to the pathology. (Such evidence will never achieve grade A.)

For traditional histopathological data items, e.g. prognosis is related to tumour type or size, the strongest evidence is likely to be provided by high-quality cohort studies (grade B). For pathological markers of predictive value, in particular molecular and genetic markers, it is likely that there will be evidence of clinical impact from randomised controlled trials (grade A) to justify inclusion as core data. The WGCS considers it unlikely that data items based on grade C evidence would be included in lists of core data.

In the cancer dataset documents, the grade of supporting evidence should be recorded against each item in the core data list.

General guidance for the inclusion and exclusion of evidence

The inclusion of core data items is dictated by international cancer staging, grading and classification systems. Dataset authors are required to cite key evidence that supports these. Inclusion criteria for evidence that supports each of the data items are:

- meta-analyses of the prognostic or therapeutic value of the data
- review articles summarising several studies
- cohort studies that have been validated on independent cohorts by other authors
- smaller studies if several papers describe consistent findings.

Evidence that should be rejected comprises:

- evidence that is not directly applicable to the cancer type under consideration
- reports that are not available in English
- single-case reports or single reports of small series of patients.

Gaps in evidence or missing evidence may be identified and submitted by members during the consultation process. The authors are responsible for evaluating the submitted evidence and, if appropriate, including it in the final draft of the dataset that is submitted for publication.

Achieving consensus

Because the core data items (guideline recommendations) are dictated by internationally agreed staging, grading and tumour classification systems, it is uncommon to encounter significant disagreements or variant opinions. However, dataset authors often provide guidance about dealing with areas of diagnostic difficulty, how to take measurements or assess other specific tumour parameters. Disagreements occasionally arise in relation to the latter, and to some non-core data items. Any disagreements are resolved by consensus expert review, i.e. comments and disagreements that are logged during the consultation process are reviewed by the expert author group, who consult with experts from a range of specialist histological and allied professional clinical groups as appropriate. The experts must identify evidence to support the inclusion or omission of a specific data item, or should revert to custom and practice and amend the dataset to indicate that there is uncertainty (listing the opposing views or options) and provide detailed explanatory text and references if available.

7 The use of SNOMED coding

SNOMED topography should be recorded for the site of the tumour. SNOMED morphology codes should be recorded for the diagnosis/tumour morphology.

Versions of SNOMED prior to SNOMED CT will cease to be licenced by the International Health Terminology Standards Development Organisation from 26 April 2017. It is recognised that versions of SNOMED 2, SNOMED 3/RT and SNOMED CT are in use in the UK; these are therefore currently considered acceptable.

SNOMED Procedure codes (P codes in SNOMED 2/3/RT) should be recorded for the procedure. P codes vary according to the SNOMED system in use in different organisations, therefore local P codes should be recorded and used for audit purposes.

A list of applicable SNOMED morphology and topography codes should be provided in an appendix.

Appendix A Process for the annual abridged review of cancer datasets