



Guidance on writing cancer datasets

1. Background

This document provides working groups and authors with the necessary guidance to produce acceptable and timely datasets. It is a working document that will be reviewed regularly by the Working Group on Cancer Services in the light of experience and feedback from authors.

As the College publishes the cancer datasets electronically and is encouraging software companies to develop products for local pathology systems, there is a need for a consistent format and style across all datasets. This document describes the 'College style' and the process that should 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 these aims are achieved. Informal advice is available from the Chair of the Working Group and the Professional Standards Unit at all stages of development of the datasets.

The dataset series will continue to evolve. The datasets can be amended whenever new information is available, and authors will be invited to review the content periodically following publication. Authors should consider using feedback from specialist pathology groups and users of software systems.

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, and authors are encouraged to describe the key features required for optimal patient management.

2. Purpose of cancer datasets

The datasets are **guidelines**. Guidelines are systematically developed statements to assist the decisions of practitioners and patients about appropriate health care for specific clinical circumstances and are based on the best available evidence at the time the dataset was prepared. It may be necessary or even desirable to depart from the guidelines in the interests of specific patients and special circumstances. Just as adherence to the guidelines may not constitute defence against a claim of negligence, so deviation from them should not necessarily be deemed negligent.

Each dataset contains **core data items** that are required by the National Cancer Data Set. Core data items are items that 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 data items should be clearly defined to allow the unambiguous recording of data.

Authors should be aware that datasets are likely to be read, *inter alia*, by trainees, general pathologists, specialist pathologists and clinicians, and by service commissioners. The dataset should seek to deliver guidance with a reasonable balance between the differing needs and expectations of the different groups. The datasets 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.

3. Role of the working group on cancer services

- The Group commissions new datasets and revisions of existing datasets from key authors, taking advice from the sub-specialty advisors.
- The Group will agree a timetable for producing the dataset with the lead author.
- The members of the Group will be involved in the processes of review and consultation on draft datasets (see Section 10).
- The Group should keep the Professional Standards Unit informed about changes in pathology practice and the delivery of cancer services that might impact upon the content of the datasets.

4. Role of the author(s)

- To agree the scope of the dataset.
- To involve necessary stakeholders (see Section 7).
- To agree a timetable for producing the dataset with the Working Group and the Professional Standards Unit.
- To assess the evidence, offer expert opinion and translate the evidence into written recommendations for practice.

5. Standard structure

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

1. Introduction

This section should:

- give details of the author(s)
- explain the importance and clinical application of the dataset
- highlight any site-specific issues
- briefly describe the methods used to obtain the evidence base for the dataset e.g. which search terms were used to search which databases.
- briefly describe the methods used to evaluate the evidence used and the ways in which disagreements were resolved
- list the clinical stakeholders consulted as part of generating the dataset (see section 7)
- briefly describe the methods used to pilot/validate the dataset and proforma. For revisions of datasets this will largely be through experience with previous versions of the dataset. Authors should consider how to validate any proposed changes to datasets.

The introduction will include the following statements (modified as necessary for each dataset):

The dataset was reviewed by the Cancer Services Working Group and was placed on the College website for consultation with the membership between xx-xx. All comments received from the Working Group and the memberships have been addressed by the authors to the satisfaction of the Chair of the Working Group and the Director of the Professional Standards Unit.

Each year, the authors of the dataset, in conjunction with the sub-specialty advisor to the College, will consider whether or not the dataset needs to be revised.

This dataset was developed without external funding to the dataset writing group/lead author (alternatively, sources of funding should be specified). The remit of The Royal College of Pathologists is to promote the quality of pathology services through training and education. It has no remit to negotiate the terms and conditions of employment for pathologists.

The College requires the authors of datasets to provide a list of potential conflicts of interest; these are monitored by the Director of the Professional Standards Unit and are available on request.

2. Clinical information required on the specimen request form

3. Preparation of specimens before dissection

4. Specimen handling and block selection (relatively brief, evidence-based summary)

4.1 Section staining and use of levels when cutting blocks

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

5.1 Clinical (usually the clinical information on the site of the tumour).

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

6. Non-core data items

To include a complete description of a case, preferences of individual laboratories, items for clinical research and supplementary information that may contribute to management or treatment decisions in individual cases.

7. Diagnostic coding

Recommended protocols (TNM, SNOMED). The dataset should include a clear statement as to which version of the TNM classification is recommended. The default is TNM6.

8. Reporting of small biopsy specimens

9. Reporting of frozen sections

10. Specific aspects of individual tumours not covered elsewhere

11. Criteria for audit of the dataset

This section should 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, and audit of lymph node retrieval.

12. References (see Section 6)

13. Appendices

TNM classification for sites covered by the dataset.

SNOMED T codes.

SNOMED M codes.

Reporting proforma.

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.

Form for synoptic reports (if applicable).

6. Editorial style

The College aims to produce all its publications to a uniformly high standard. All material should therefore be factually accurate and up to date, fully referenced, grammatically correct, clearly and consistently written, in the appropriate style for the target audience and intelligible to all potential readers.

The dataset will be fully edited before publication (see Section 10), but to aid the process, please ensure it adheres to the following 'house style' rules.

- All abbreviations must be written out **in full** the first time they are used.
- Please type in Arial font, size 11 – especially on the dataset proformas.
- Different levels of headings and sub-headings should be clear; please use bold/larger font to differentiate between them.
- On dataset forms, boxes must be to the **right** of the text, i.e. Yes No
- **Copyright:** if any part of the document requires permission for use and/or credits, please either provide the Professional Standards Unit with a copy of the permission or inform the Professional Standards Unit that the College still needs to pursue permission.
- **Figures and tables:**
 - must be referenced within the body of the text
 - must have concise titles
 - any embedded figures must also be provided separately as high-quality JPG files.
- **References/footnotes**
 - References must be consistent and complete. They should be presented in exactly the same format as the following examples.
For a journal article: Kearsley JH, Thomas S. Prognostic markers in cancers of the head and neck region. *Anticancer Drugs* 1993;4:419–429.
For a book: Sobin LH, Wittekind C. *TNM Classification of Malignant Tumours (6th edition)*. New York: Wiley-Liss, 2002.
 - Please **do not** use the Microsoft Word footnote facility. Instead, please reference by inserting superscript numbers **after** the punctuation (e.g. “xxx.⁴”) and give a full numbered reference list in a separate ‘References’ section.
 - If there are multiple authors, please give the first six names, followed by ‘*et al.*’.
 - Use **full** number ranges of pages to avoid ambiguity, e.g. 1023–1026 (not 1023–6 or 1203–26).

7. Stakeholder involvement

Specialist societies, organisations and individuals (including patient representatives) may have an interest in the preparation of the dataset. Key stakeholders should be encouraged to contribute to its development at an early stage.

A list of stakeholders (including email addresses) who have either been involved in the preparation of the dataset or who are to be consulted on the final version must be provided to the Professional Standards Unit on submission of the text for consultation. A list of the stakeholders should be included in the introduction.

The views of stakeholders will also be sought in parallel with the College-wide consultation process (see Section 10).

8. Interpretation of evidence

To produce a sound guideline, it is essential that evidence is identified and interpreted appropriately.

Authors must briefly describe the methods used in the introduction to the dataset. The description should include how the authors have:

- identified relevant and up-to-date evidence
- identified gaps in evidence
- used a consistent method to interpret and assess the strength of the evidence
- consulted with, or propose to consult with, stakeholders to obtain their comments on the acceptability of the evidence and its interpretation.

9. Further information

To audit the standard of production and content of the datasets, the College will use the internationally agreed process for the ‘Appraisal of Guidelines for Research and Evaluation’

(AGREE) – www.agreecollaboration.org. Authors may wish to consult this document during the development and revision of the datasets.

Help in devising a search strategy for acquiring information can be found in ‘SIGN guidance on search filters’ – www.sign.ac.uk/methodology/filters.html

10. Consultation and publication policy

The dataset must be submitted to the Working Group on Cancer Services via the Professional Standards Unit. The draft document will be circulated electronically to the Working Group. Comments will be collated by the Professional Standards Unit and reviewed by the Chair of the Working Group. Any major concerns will be discussed with the lead author before the dataset is placed on the website. Minor issues will be added to the list of comments received during the consultation with fellows of the College.

Once the draft dataset is approved by the Chair of the Working Group, the Professional Standards Unit will check the document, especially the references. It is then given to the College’s Managing Editor – Publications, who edits and formats the document as necessary. If anything is unclear, the Managing Editor may email the author directly for clarification.

The edited dataset is then put on the password-protected ‘Discussion documents’ area of the College website for consultation (usually for four weeks. For full details see the College’s publications policy, *RCPATH Bulletin* 2005;131:16–17). All stakeholders who have been identified by the author will also be consulted by email on the final version. The stakeholders will respond directly to the Publication Department and, in the absence of specific concerns, be asked to indicate their support for the content of the dataset.

Once the deadline has passed, the Managing Editor collates the feedback and, after consultation with the Chair of Working Group to address any generic issues, sends it back to the author(s).

The author is then asked to annotate the feedback and update the edited version of the dataset as necessary. Please bear in mind that members have a right to see the annotated feedback when the dataset is published.. It is then returned to the Managing Editor, who ensures that the Director of Publications is satisfied with the author’s annotations and updates, adds information about the consultation procedure to the front page and may do another round of editing if necessary. Once the Chair of the Working Group and the Director of Publications approve the finalised dataset, it is published on the College website.

Each dataset will contain a generic introduction. This introduction will be provided as a separate document on the website and will be referenced to each dataset by the Publications Department.

Sample time schedule

ACTION	RESPONSIBILITY OF	TIME
Check draft text and references	Professional Standards Unit	2–3 weeks
Circulate draft to Working Group on Cancer Services for approval through the Chair of the Working Group	Professional Standards Unit and Working Group	1 week
Editing of draft for online consultation	Publications Department	2 weeks
Online consultation for all fellows of the College and email consultation for stakeholders	Publications Department, Fellows and Stakeholders	4 weeks

Chair of Working Group reviews feedback and addresses any generic issues	Chair of Working Group	1 week
Author receives and annotates feedback collated from consultation, and updates dataset if necessary.	Author	2 weeks
Author submits annotated comments and final text is sent to Publications Department which will be copied to the Professional Standards Unit and Chair of Working Group	Professional Standards Unit and Working Group	2 weeks
Final edit, Chair of Working Group and Director of Publications approve final document, dataset published online.	Publications Department	1–2 weeks
Timescale		14 -16 weeks

Professional Standards Unit

The Royal College of Pathologists

August 2008