



Public Health
England

Protecting and improving the nation's health

Cancer Outcome and Services Data set

Pathology technical guidance v4.0.1

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-leading science, research, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: [@PHE_uk](https://twitter.com/PHE_uk)
Facebook: www.facebook.com/PublicHealthEngland



© Crown copyright 2019

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit [OGL](https://www.ogilive.com/). Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Published September 2019
PHE publications
gateway number: GW-621

PHE supports the UN
Sustainable Development Goals



Contents

Executive summary	4
Introduction	5
Help and support	6
General submission principles	7
General file formatting principles	8
Data set and XML record structure	10
The COSD element	10
The {*}Record element	10
Core and content group elements	11
Generating COSD Pathology XML	13
Reserved characters	15
File naming convention.....	15
Data submissions	17
Who will submit the data?.....	17
What data items should be submitted?.....	18
Validation.....	18
Reporting	20
Appendix 1: sample lung XML record	21
Appendix 2: file submission template	24

Amendment history

Version(s)	Date	Amendment History
COSD Pathology v4.0	25 June 2019	Draft to support v4.0 sent to DCB
COSD Pathology v4.0	4 July 2019	Final v4.0 sent to DCB
COSD Pathology v4.0	2 August 2019	Final amended document for publication
COSD Pathology v4.0.1	5 November 2019	Diagnosis (ICD Pathological) code reporting update pg7

Executive summary

The purpose of this document is to provide instruction to informatics personnel within provider organisation pathology laboratory departments, Laboratory Information Management System (LIMS) suppliers and other IT software suppliers (inhouse and commercial), regarding file creation and submission of the Cancer Outcomes and Services Data set (COSD) pathology data. It should be read in conjunction with the documents listed within.

This document describes the standards for file submission, including the XML construction and file naming to facilitate uploading onto the National Cancer Registration and Analysis Service (NCRAS) database (ENCORE). In addition, it provides assurances that the proposed approach supports the implementation of DCB1521 Amd 13/2019.

This is an update to an existing information standard DCB1521 Amd 74/2016 and is required to ensure that the data still meets the business objectives, scope and content of the standard and continues to be clinically accurate and relevant.

In order to maintain the clinical accuracy, it is important to regularly review COSD with clinical experts from across the NHS, including analysts at Public Health England (PHE) and NHS England. For this data set, extensive consultation and advice was also sought from the Royal College of Pathologists (RC Path) Working Group on Cancer Services.

Occasionally other information standards have specific data items which interact with COSD. Where this happens, liaison with the developers of those standards was concluded to ensure all data items remain accurate and are updated where necessary.

Introduction

This document provides technical guidance to support personnel within provider organisation pathology laboratory departments, Laboratory Information Management System (LIMS) suppliers and other IT software suppliers (inhouse and commercial), in the submission of the COSD Pathology data set v4.0. It should be read in conjunction with:

- information standards notice: reference DCB1521 Amd 13/2019
- COSD Specification
- xml pathology schema documentation v4.0
- COSD pathology user guide v4.0

Users may also wish to read the COSD Implementation Guide, which provides further support on implementation of the changes to the standard. See NCIN website¹ for access to these documents and other information.

Providers of pathology services are required to provide a monthly return on all cancer patients diagnosed from 1 January 2013, who then go on and have a pathological excision, sample or other pathological test using this data set. The data are stored in the National Cancer Registration and Analysis Service (NCRAS) database (ENCORE). Submissions are made by each provider to the relevant NCRAS branch office for uploading to ENCORE.

Data may be extracted from a number of different electronic sources and submitted as separate files. The required format for submissions for COSD is XML. These and other details of the submission should be included in the COSD data transfer agreement, agreed between the provider organisation and their local NCRAS office.

Data is collated and mechanisms for transmission of data from providers to NCRAS offices have been extended to carry the COSD data items. On receipt, patient level data is validated and linked with existing records as appropriate.

Reporting mechanisms exist to provide consistent feedback on submissions and additional reports to local providers². Both patient identifiable and anonymised data (where appropriate) will be made available for analysis and reporting purposes³.

¹ http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd

² <https://cancerstats.ndrs.nhs.uk/>

³ <https://www.cancerdata.nhs.uk/dashboard/>

Help and support

For technical queries relating to the creation of these files please contact your local NCRAS office in the first instance (see data transfer agreement for local details).

For queries regarding the:

- data set, contact: COSDenquiries@phe.gov.uk
- Data Model and Dictionary Service, contact: datastandards@nhs.net
- schema, contact: COSD@phe.gov.uk
- further help and tutorials on XML can be found at:
www.w3schools.com/xml/default.asp

General submission principles

The following set of 10 principles should be used to support data submissions:

1. Pathology providers must submit all data relating to patients for whom they have either reported on and authorised the original pathology report or commented on (second opinion) as a specialist centre (creating a supplementary report).
2. Submitted files must be sent by secure file transfer methods as agreed with your regional NCRAS office.
3. Files must reach the regional NCRAS office by the twenty-fifth working day following the end of month for all 'registerable pathology reports' authorised.
4. Each file may include records for more than one tumour group.
5. Individual records must contain the section 'LinkagePatientID'.
6. Providers should aim to complete all the relevant data items as soon as possible, however as long as the mandatory fields within the 'LinkagePatientID' are completed the record can be submitted.
7. Records in each submission must include all applicable sections where possible.
8. Demographic section must be submitted by each provider on the first submission of a record.
9. For updated records, only the updated/amended sections and core linkage items need to be submitted.
10. The data submission files MUST comply with the COSD XML schema specifications pack*.

(Note: All data must be submitted to the NCRAS in XML format only.)

Note: Where the 'Diagnosis (ICD10 Pathological)' code only has 3 characters, for example C01, please add "X" as a 'packing digit' to meet the validation rules (such as C01.X, C07.X, C73.X). In addition, the reporting format excludes the decimal CXX.X or DXX.X, all xml reports must be recorded as CXXX or DXXX

General file formatting principles

The required format for submission of data for the COSD will be extensible markup language (XML) as specified in the COSD Pathology XML schema specification pack. This contains all of the schema documents listed below and all embedded schema referenced in the above documents. This schema pack also contains the data type schema, containing the formats for submitted data. The XML schema pack is available on request from the NHS Data Model and Dictionary Service or by downloading it direct from the TRUD website⁴.

XML schemas (XSD) have been designed for a generic core COSD pathology data set and for 11 tumour site specific data sets, which each include the core data items. In addition:

- these schemas define the expected structure of the XML submissions
- schema design is segmented into separate schema, defining the different sections of COSD pathology data specification
- these schemas are embedded in a hierarchical manner into a single schema
- these schemas contain information on the expected values for a data element

Data submitted in XML format will be required to conform to the schemas for the appropriate cancer site, or core for all registerable conditions.

Within the data set 'core' is a 4-letter word which describes:

- the subset of COSD pathology data items that are common to all site groups
- if you are submitting a lung record and need to provide the core pathology and lung specific data items, this is perhaps the more obvious definition
- 'CORE' is no longer the non-specific site group that is used when a record does not belong to any of the other well-defined groups

In v4 the word CORE is not used within the schema, rather the site-specific group sections and records are now known as a, BreastRecord, CNSRecord, UpperGIRecord, UrologyRecord etc., this is a change from v3.

For all other disease types (which do not have their own specific subsection in v4), will be recorded under 'OtherRecord'. This simplifies data recording when compared to v3.

⁴ <https://isd.digital.nhs.uk/trud3/user/guest/group/0/home>

The top-level schema in the hierarchical structure is: COSD_Pathology-v4-0.xsd.

For a Record which does not fall into any of the tumour groups, i.e. a Record with only core elements, then the default schema path for a Record is used which is defined in the master file: COSD_Pathology-v4-0.xsd. This will only allow a record to build using what is deemed as Core sections as per the dataset specification.

The following list outlines the tumour specific site and associated schema, embedded within the above schema:

Tumour site	Schema
Breast	COSD_Pathology-v4-0_BREAST.xsd
Central Nervous System (CNS)	COSD_Pathology-v4-0_CNS.xsd
Colorectal	COSD_Pathology-v4-0_COLORECTAL.xsd
Children, Teenagers and Young Adults	COSD_Pathology-v4-0_CTYA.xsd
Gynaecology	COSD_Pathology-v4-0_GYNAECOLOGICAL.xsd
Head and neck	COSD_Pathology-v4-0_HEADNECK.xsd
Lung	COSD_Pathology-v4-0_LUNG.xsd
Sarcoma	COSD_Pathology-v4-0_SARCOMA.xsd
Skin	COSD_Pathology-v4-0_SKIN.xsd
Upper GI	COSD_Pathology-v4-0_UPPERGI.xsd
Urology	COSD_Pathology-v4-0_UROLOGICAL.xsd

Character set: UTF-8: information on permitted formats are contained in the data type schema

Note: for clarity a sample of a single lung record is included in Appendix 1.

Data set and XML record structure

The COSD element

The root element of a COSD XML file is `<COSD>` and only one is permitted and required per submission – such as in each individual xml file. There are 6 child elements that need to be provided within the root element, these are:

- `<Id root="uuid" />`
 - (the root attribute will be a universal unique identifier (UUID) for the submission, this takes the form of an 8-4-4-4-12 hexadecimal characters, for example DEAECC2-76AA-411E-B994-8FDD98C3FFFA)
 - (a UUID Library should be used to create it)
- `<OrgCodeSubmitter extension="orgcode"/>`
 - (the extension attribute should contain the NACS code of the submitting provider)
- `<RecordCount value="count" />`
 - (the value attribute should identify the number of `<COSDRecord>` elements being supplied with this submission)
- `<ReportingPeriodStartDate>` and `<ReportingPeriodEndDate>`
 - (these elements give the time period for the data submission, this is normally the “trigger event” date range and takes the (ISO) format YYYY-MM-DD)
- `<FileCreationDateTime>`
 - (this is the timestamp of when the submission was generated and takes the format YYYY-MM-DDTHH:MM:SS e.g. 1900-01-01T10:11:12)

The `{*}Record` element

The `<{*}Record>` element is a child element of the `<COSD>` element. This element is a choice of single tumour record within the submission, e.g. BreastRecord, CNSRecord etc. As with the `<COSD>` element it has an `<Id root>` element with a UUID value attribute. This is a unique identifier for the tumour record and does not need to be preserved across submissions – such as `{*}Records` pertaining to the same tumour in future submissions do not need to have the same UUID.

The choice of `{*}Record` determines what elements can be submitted for that type of record, for example PrognosticIndex for BreastRecords, Pretreatment Assessments for HeadNeck etc.

Core and content group elements

The choice of record element determines what can be submitted for that particular record. It will allow all core elements and any site specific elements where specified by the dataset.

For example, a Breast COSDRecord is represented in COSD Pathology XML like:

```
<BreastRecord>
  <Id root>
  <LinkagePatientID>
    <!-- LinkagePatientID Elements -->
  </LinkagePatientID>
  <Demographics>
    <!-- Demographic Elements -->
  </Demographics>
  </Pathology>
  <!-- Core Pathology Elements -->
  <PathologyBreast>
    <!-- Additional Breast Pathology Elements -->
  </PathologyBreast>
</Pathology>
<BreastRecord>
```

It should therefore become clear to the reader that 'Other; (non-specific site) records have no content group:

```
<OtherRecord>
  <Id root>
  <LinkagePatientID>
    <!-- LinkagePatientID Elements -->
  </LinkagePatientID>
  <Demographics>
    <!-- Demographic Elements -->
  </Demographics>
  </Pathology>
  <!-- Core Pathology Elements -->
</Pathology>
<OtherRecord>
```

Note: Site specific data sets don't just add additional data items to their content group (and sub sections), they also augment core group sections where necessary.

For example, a non specific 'Other' Pathology record may contain:

```
<OtherRecord>  
  <InvestigationResultDate>...</InvestigationResultDate>  
  ...  
  <MicrosatelliteInstabilityMsiTesting code="..."/>  
</OtherRecord>
```

Whereas a BreastPathology section may also contain:

```
<BreastRecord>  
  <Pathology>  
    <InvestigationResultDate>...</InvestigationResultDate>  
    ...  
    <MicrosatelliteInstabilityMsiTesting code="..."/>  
    <PathologyBreast>  
      <CoreBiopsyNode code="..."/>  
    </PathologyBreast>  
  </Pathology>  
</BreastRecord>
```

Note: for clarity a sample of a single lung pathology record is included in Appendix 1.

Generating COSD Pathology XML

The hierarchical nature of the data set may lead providers to adopt an object orientated programming (OOP) approach to developing the COSD Pathology XML submission, but it is also possible to use a more procedural approach to generate the XML and to use condition logic to include or exclude specific data items.

This may be simpler to produce, but might be harder to maintain in the longer term and whilst OOP may be the more elegant solution, providers may need a pragmatic approach with constrained resources, such as existing skills/technologies.

A working example of the OOP paradigm is given below. The code is far from complete but demonstrates the key principle of inheritance. The use of CoffeeScript is merely for convenience; to all intents and purposes consider it pseudo code:

```
class CoreGenerator
  constructor: () ->

  buildCOSDRecord: (pathology) ->
    @buildCore(pathology)

  # Core section
  buildCore: (pathology) ->
    @buildLinkagePatientId(pathology.patient)
    @buildDemographics(pathology.patient)
    @buildPathology(pathology) for pathology in tumour.pathologies

  buildLinkagePatientId: (patient) ->
    alert "build LinkagePatientId"

  buildDemographics: (patient) ->
    alert "build Demographics"

  buildPathology: (pathology) ->
    alert "build Pathology "

class BreastGenerator extends CoreGenerator
  # Core section

  buildPathology: (pathology) ->
    super pathology
    alert "build BreastPathology/BreastPathology"
```

TODO: add site specific fields here

```
generator = new BreastGenerator  
generator.buildCOSDRecord(...)
```

Reserved characters

XML has reserved characters which should not be used in data submissions, these are:

Reserved character	Meaning	Entity reference
>	Greater than	>
<	Less than	<
&	Ampersand	&
%	Percent	%

Note: If it is unavoidable that these reserved characters are used in data submission, they should either be replaced by the corresponding entity reference or encapsulated with the <![CDATA[{data text}]]> tag. Particular care should be taken with the pathology report text field within the COSD Pathology v4.0.

File naming convention

The submission file must be named using the following convention:

XML file:

- COSD_<FILE SOURCE>_<Submitting Org>_<Reporting Period Start Date>_<Reporting Period End Date>_<Date of file creation>.xml

Where:

- <COSD> is a fixed value
- <FILE SOURCE> is MDT or PAS or PATH or RIS (or other source description as agreed with NCRAS)
- <Submitting Org> is the Organisation Code (e.g. AB3) for the submitting organisation
- <Reporting Period Start Date> **must** always be in the format CCYY-MM-DD
- <Reporting Period End Date> **must** always be in the format CCYY-MM-DD
- <Date and time of file creation> Timestamp when the file was created in the format CCYY-MM-DDThh:mm:ss

Example

the file name for organisation (X09) submitting its own pathology data for activity month July 2020 on the 5th September 2020 at 10:30:22 AM will be:

- COSD_MDT_X09_2020-07-01_2020-07-31 2020-09-05T10_30 22.xml

Files may be zipped prior to transmission, in this case the file extension .zip will be acceptable.

Note: Each file submitted must have a unique filename as generated by the above method.

Data submissions

XML data should be validated against the schema prior to submission.

XML data submissions should be given a new UUID in the <COSD> element, where submissions are altered and re-submitted a new UUID should be applied. Each tumour record in the submission should have a unique UUID as the root attribute for the <COSDRecord> element. New UUID must be created for resubmissions of data.

All data submissions MUST be transmitted either between nhs.net to nhs.net email accounts (or alternative email accounts that are accepted as part of the nhs.net framework), or in some cases the pathology report may be submitted by direct files transfer (via an API) immediately after the report is authorised. The regional NCRAS offices will agree with providers the relevant recipient account in the data transfer agreement.

Monthly data submissions are required from each provider within 25 working days of the relevant month end. A schedule of submission deadlines is available on the NCIN website.

Trusts should complete the COSD submission template which should be attached to each email submission, along with the data submission files themselves. This provides a summary overview of their submission and alerts the NCRAS to any special or extenuating circumstances which may have affected the submission that month. See Appendix 2.

Who will submit the data?

Trusts were mandated to collect and submit Pathology data in XML from January 2016, with an agreed delayed start until July 2016 for September 2016 submissions.

1. Files should be submitted by NHS providers, where the pathology specimens are authorised and reported or where a second opinion has been sought and a supplementary report is created.
2. The submission files should relate to a single provider and only for data that they own.
3. Providers must clarify arrangements or changes for submitting the data with their local NCRAS office.

In some cases, the files may relate to a provider who are contracted to report the specimens on behalf of the originating Trust. This may not be the Trust where the excisions or biopsies are performed.

What data items should be submitted?

All applicable data items specified as either mandatory or required in the data set and XML schema documentation should be submitted as soon as available.

The mandatory, required or optional (M/R/O) column indicates the recommendation for the inclusion of data. This applies specifically to the XML files but should also be used to decide on data to be included through other message formats.

M = Mandatory

This data item is mandatory; the record or part of the record cannot be submitted if the mandatory data items are not completed. The file will be rejected if the mandatory linkage items are absent and/or sections will be omitted where mandatory items are missing, even if other data items are completed.

R = Required

This data item is required as part of NHS business rules and must be included where available or applicable, however, the section can be submitted without completing all the required items.

O = Optional

This data item can be included at the discretion of the submitting organisation and their commissioners as required for local purposes.

Validation

The data will be validated in ENCORE according to a set of rules. If the data validation rules are not met, the whole or relevant parts (data set sections or records) of the extract may be rejected and returned to the provider.

The provider will normally be expected to resolve any errors/issues or add missing data and re-extract the file for sending to the NCRAS within 5 working days, for re-validation. The turnaround time for validation, any re-submission and subsequent re-validation is necessarily short as delays will adversely affect the timeliness and quality of the data and the validity of conformance reporting.

A record of rejected files will be kept by the NCRAS as an audit trail and to support conformance monitoring; original files may be retained in order to make a comparison with subsequent files received.

It is good practice to identify the areas which require attention prior to submission in the form of check reports, outlining if there are key (Mandatory) fields missing. This can be in the form of an error report or other explanation, allowing correction to prevent the records from being rejected. Equally a warning message could be created and displayed on the screen, where there is an attempt to save a record on a system without the mandatory data items.

Reporting

The NCRAS has developed standardised reports which are available to all providers submitting data through the NCRAS conformance portal (CancerStats)⁵, these are only available through the NHS Health and Social Care Network (HSCN) Network.

The HSCN is a new data network for health and care organisations which replaced N3. It provides the underlying network arrangements to help integrate and transform health and social care services by enabling them to access and share information more reliably, flexibly and efficiently.

Analytical reports are being developed separately⁶.

Providers should continue to contact their regional NCRAS office to request any data they require which is not made available via standardised reporting.

⁵ <https://cancerstats.ndrs.nhs.uk/>

⁶ <https://www.cancerdata.nhs.uk/dashboard/>

Appendix 1: sample lung XML record

```

<LungRecord>
  <Id root="FEF90fef-314a-AfCE-0ADd-e4f8627Fb6C5"/>
  <LinkagePatientId>
    <NhsNumber extension="6288305843"/>
    <NhsNumberStatusIndicatorCode code="02"/>
    <PersonBirthDate>2019-05-31</PersonBirthDate>
    <OrganisationIdentifierCodeOfProvider extension="QEPHK"/>
  </LinkagePatientId>
  <Demographics>
    <PersonFamilyNameAtBirth>miAwzlsKBRJ4zKszv4TedZGq3olwuAWREMg</PersonFamilyNameAtBirth>
    <PersonGivenName>BNN6Kk9nXvwUB3l6kwyeb7zD9HxYD8QfrtJ</PersonGivenName>
    <Address>
      <UnstructuredAddress>
        <Streetaddressline>FDHRrp8pxWTSSPvB5Zne6TB9YmTEGfzxvHnxxcyr3sT0cV5TW8</Streetaddressline>
      </UnstructuredAddress>
    </Address>
    <PostcodeOfUsualAddressAtDiagnosis>Oh9zxBYm</PostcodeOfUsualAddressAtDiagnosis>
    <PersonStatedGenderCode code="2"/>
  </Demographics>
  <Pathology>
    <InvestigationResultDate>2019-05-31</InvestigationResultDate>
    <ServiceReportIdentifier extension="fqOUfEBdg48SbfHoWQfnRdEX3Q7DVKldrCfo"/>
    <PathologyObservationReportIdentifier extension="iC41rWwpBYoh78e4JyRWhJYnKcoCavpjx34G"/>
    <ServiceReportStatus code="2"/>
    <ConsultantPathologyTestRequestedBy>
      <ProfessionalRegistrationIssuerCode-ConsultantPathologyTestRequestedBy code="02"/>
      <ProfessionalRegistrationEntryIdentifier-ConsultantPathologyTestRequestedBy>xmytZ8xIkPzAsDD1n3DrQxk
        t7evzwhv</ProfessionalRegistrationEntryIdentifier-ConsultantPathologyTestRequestedBy>
    </ConsultantPathologyTestRequestedBy>
  </Pathology>
</LungRecord>

```

```

</ConsultantPathologyTestRequestedBy>
<OrganisationSiteIdentifierPathologyTestRequestedBy extension="80QDT"/>
<SampleCollectionDate>2019-05-31</SampleCollectionDate>
<SampleReceiptDate>2019-05-31</SampleReceiptDate>
<OrganisationIdentifierOfReportingPathologist extension="38L"/>
<ConsultantPathologist>
  <ProfessionalRegistrationIssuerCode-ConsultantPathologist code="03"/>
  <ProfessionalRegistrationEntryIdentifier-ConsultantPathologist>QY2OdbFwp9d5nZcSd2o
    wqlecNvKI9z4a</ProfessionalRegistrationEntryIdentifier-ConsultantPathologist>
</ConsultantPathologist>
<SpecimenNature code="2"/>
<TopographyMorphologySnomed>
  <SnomedVersionPathology code="02"/>
  <TopographySnomedPathology code="WQ0SXOOWA"/>
</TopographyMorphologySnomed>
<DiagnosisIcdPathological code="A2efz6"/>
<TumourLateralityPathological code="M"/>
<PathologyInvestigationType code="FE"/>
<PathologyReportText>73Br4dlIV4OQqBeGYWb1OSRHXiD4N3b2R6QBMKKQZZOpFWaXc2pGD9vVCjRS
  NZkT67Nql7PAAJkMwUSpHhj1pX2W7RuaT</PathologyReportText>
<LesionSizePathological value="5.56"/>
<GradeOfDifferentiationPathological code="G4"/>
<CancerVascularOrLymphaticInvasion code="XX"/>
<Excisionmargin code="07"/>
<SynchronousTumourIndicator code="N"/>
<NumberOfNodesExamined value="891"/>
<NumberOfNodesPositive value="382"/>
<TnmCodingEdition code="3"/>
<TnmVersionNumberPathological>y5</TnmVersionNumberPathological>
<TCategoryPathological>saX5CVeZ8PkZLhX</TCategoryPathological>

```

```
<NCategoryPathological>B0gySWeX0N2Zj6M</NCategoryPathological>
<MCategoryPathological>mmsW4NoU9jTHwK6</MCategoryPathological>
<TnmStageGroupingPathological>5p8B8Fg46Z3twGd</TnmStageGroupingPathological>
<NeoadjuvantTherapyIndicator code="9"/>
<Ki-67Indicator code="3"/>
<Ki-67Result value="99"/>
<Mlh1NuclearExpressionIntact code="F"/>
<Pms2NuclearExpressionIntact code="F"/>
<Msh2NuclearExpressionIntact code="N"/>
<Msh6NuclearExpressionIntact code="N"/>
<MicrosatelliteInstabilityMsiTesting code="H"/>
<PathologyLung>
  <ExtentOfAtelectasis code="5"/>
  <ExtentOfPleuralInvasion code="2"/>
  <PericardialInvasion code="N"/>
  <DiaphragmInvasion code="N"/>
  <InvasionIntoGreatVessel code="N"/>
  <InvasionIntoHeart code="Y"/>
  <MalignantPleuralEffusion code="Y"/>
  <InvasionIntoMediastinum code="Y"/>
  <SatelliteTumourNodulesLocation code="9"/>
</PathologyLung>
</Pathology>
</LungRecord>
```

Appendix 2: file submission template

Generic file source	File name	Number of records	Any reasons for variation from number expected YES/NO	Extenuating circumstances if previous column contains 'YES'	Other comments
MDT					
PATH					
PAS					
RIS					

Note: generic file names as listed should align with the sources identified in the data transfer agreement. Any other file sources should be referenced using consistent terminology as agreed with the NCRAS.