Genomics England Cancer Model

Work stream	Informatics	**Status**	Final
Programme Director	Peter Counter	**Version**	V3.1.2
Document Owner	Amanda O'Neill	**Version Date**	15/11/2016

1 Document Management

Document Owner: Amanda O'Neill
Authors: Jim Davies, Clare Turnbull, Kathy Farndon, Nancy Horseman, Adam Milward, Steve Harris, Louise Jones, Amanda O'Neill, Kay Lawson

Version Control

Version	Date	Summary of Changes
1.0.1	13/12/2016	IIP Release
2.0.0	17/03/2016	Main Programme Release
2.0.1 Errata	01/02/2016	Errata
2.0.2 Errata	05/02/2016	Errata
3.0.0	13/10/2016	Changes for Gear 2, Pre-consultation release
3.1.1	26/10/2016	Post consultation release
3.1.2	15/11/2016	Final review comments completed

Reviewers

Name	Responsibility	Date
Amanda O'Neill	Director of Clinical Data	16 November 16 Version 3.1.2
Clare Craig	Lead GMC Liaison Officer	15 November 16 Version 3.1.2

Approved by

Name	Responsibility	Date
Tom Fowler	Deputy Chief Scientist	16 November 16 Version 3.1.2
Peter Counter	Programme Director	16 November 16 Version 3.1.2

Table of Contents

2 Introduction

Purpose

The purpose of this document is to describe the data that Genomic Medicine Centres are asked to supply to accompany the samples submitted for analysis. This document is intended to be read in conjunction with the User Guide and describe the classes of data expected, the association between diseases and the classes of data deemed relevant, and the datatypes employed.

Audience

This document is primarily written for informatics leads within the GMCs and those involved in the collection and submission of data for the UK 100,000 Genomes Project.

Related Documents

This document should be read in conjunction with:

- Genomics England Data Model Catalogue (Genomics England Model Catalogue)
- Schemas (XSDs)
- Example XML Files
- NHS England GMC Service Specification
- National Cancer Intelligence Network Cancer Outcomes and Services Dataset (COSD) Version 7.0 User Guide
- Appendix A, B, C, D, E, F and G (contained in this release pack)

How to use this document

This document is split into sections that describe the information we expect to receive within the context of each xml submission. The document is primarily split by schema. Within each schema we expect to receive a set of

classes of data. Within each class we expect to receive a number of data elements, each of which is associated with a name, a brief explanation, a multiplicity, a datatype and often a business rule.

NOTE: this document should be used in conjunction with the User Guidance for Cancer Data Specification Document.

In this version of the document, the data item identifiers have been included, to facilitate look-up in the current version of the on-line metadata catalogue. In addition you can click on the value link to view the full definition for the data type and the applicable constraints.

SCHEMAS

Define the classes and data elements included in each xml submission. Each schema describes the information within the corresponding xml document. Each class describes the data elements included within each section of the schema. Each data element has a data type and some have business rules. Click on the hyperlinks in the document to navigate to each of these.

DATA TYPES:

Types, rules and enumerations that constrain the value of a data element.

Rules are expressed as regular expressions and/or groovy code.

Enumerations are described by their code and description.

NOTE: Enumerations in grey are deprecated. Although deprecated codes will continue to be accepted they will re removed from the next release and users are encouraged to use other suitable codes.

BUSINESS RULES:

Contains business rules that apply across data elements within the context of a file submission.

Completeness:

From a data modelling perspective most of the classes are 'optional', with the exception of the registration and consent information, in that an event of that class may not yet have occurred, or may not yet have been reported, for a particular participant.

Event records or reports are required for all relevant clinical events or observations to date (for the core data).

For some classes of event, a report will be accepted only if additional classes are provided as part of the report and/or other values are supplied for some of the data items involved: these are the 'mandatory' items within those classes.

An item is 'mandatory' for a particular class if it has a multiplicity of 1..1 or 1..*

Where an item has a multiplicity of 0..1 it is considered 'non-mandatory' and a report for that class of event will be accepted even if no value has been supplied for that item.

From a contractual perspective, however, values are expected for all applicable data.

3 XML Submission

All submissions in XML format must include as metadata: the date and time upon which the XML file was generated; the name and version of the schema used for validation; and the organisation within the GMC responsible for the participant. A source system identifier and a local report identifier may also be included.

The data provided in XML format must include an event date and event reference for the report in question. This reference should be unique within the GMC. If a second submission is received against the same event reference then this will be treated as an update.

XML Schema (.XSD Files)

The XML files and the Cancer Model Data Specification v3 are generated from the Genomics England Data Model Catalogue. The XML schemas can be downloaded directly from the Model Catalogue within the Assets folder for the Model.

For access to the Cancer Data Model Catalogue, and to submit any comments, observations or issues please contact the Genomics England Contact Desk *(see Section 14 in this document regarding support and queries).*

XML Validation to take place before submission of file

The use of XML was mandated in the e-Government Interoperability Framework (eGIF) as a messaging standard between government organisations and has been adopted by NHS (http://systems.digital.nhs.uk/data/nhsdmds/ddcn/cr1345.pdf) and therefore, this programme. XML delivers some rigour to messaging by controlling the message structure, and the data element contents and format through an XML schema definition (XSD).

It is essential that NHS GMCs send files that comply with their corresponding schema. This is done by 'validation', an electronic process that compares an XML message against its XSD. There are a number of online tools that provide this service – including free tools such as: Notepad++ (https://notepad-plus-plus.org/), freeformatter.com (http://www.freeformatter.com/xml-formatter.html). There are also richer paid-for tools that provide a graphical view. The NHS recommends Altova (https://www.altova.com/) and GeL uses oXygen XML (https://www.oxygenxml.com/).

The Model Catalogue also offers XML validation functionality by selecting 'validate xml' from the asset menu.

All submissions not passing validation will be rejected and an email will be sent advising of the failure and the reason from the Data Acquisition and Management system to any individuals subscribed to warning messages for that GMC.

4 FAQs

Essential data

Patient/Ethnicity

'99 - Unknown' has been added in the latest release.

Core Data

RiskFactors/Breast Density

Although Breast Density is not routinely captured by all sites, it is considered a strong indicator of breast cancer and breast cancer recurrence. Therefore, it is hoped that GMCs will use this opportunity to capture this information.

Tumour imaging

Images should be captured according to the requirements outlined in the GeL Tissue Handling Protocol Draft July 2016 v2.2.

Images should be appropriately labelled with Participant ID and Tumour ID, indicating whether FF or FFPE samples (linked to individual sample IDs submitted). Images should not be marked with the patient's NHS number. Initially these images should be stored at GMCs and details for image upload to the Genomics England biorepository will be confirmed by Genomics England.

5 Essential Data

The GMC clinic must provide registration information (participant information, consent, diagnosis) via EDCT or sFTP (in XML format) before any samples for that participant are sent to the Biorepository. The GMC clinic is expected to establish eligibility and to validate NHS numbers before registration. The data items described below, where applicable, are essential to the subsequent management of the participant, the sample, and any results obtained from the sequencing process.

Schema 5.1 Registration and Consent [1..1]

One report containing Registration and Consent must be submitted for each participant.

5.1.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Registration and Consent report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

5.1.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland). Patients without an NHS number (or equivalent) should be allocated a temporary NHS number by the treating hospital.

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

5.1.2 Registration [1..1]

One report containing Registration must be submitted together with each Registration and Consent report.

The Registration Event Date is the date of registration.

Name	Description	Multiplicity	Data Type	Related To
Surname at Birth (12511@1.0.1)	The participant's surname at birth, if available and different from current surname	0..1	personFamilyNameAtBirth	PERSON FAMILY NAME (AT BIRTH) (CR0140 from Cancer Outcomes and Services Dataset)
Person Stated Gender (12509@1.0.1)	The participant's current gender	1..1	personStatedGenderCode **1** Male **2** Female **9** Indeterminate (Unable to be classified as either male or female) **X** Not Known (PERSON STATED GENDER CODE not recorded)	PERSON STATED GENDER CODE (CR3170 from Cancer Outcomes and Services Dataset)
Person Phenotypic Sex (12510@1.0.1)	The participant's sex classification at birth. 9 - Indeterminate, may only be used if the patients chromosomal sex at birth is ambiguous. Samples cannot be sequenced until the chromosomal sex is established and therefore all patients registered with a value of 9 must confirm that the patients chromosomal sex is ambiguous rather than unknown prior to sequencing.	1..1	personPhenotypicSexClassification **2** Female **1** Male **9** Indeterminate	
Ethnicity (14445@1.0.1)	The ethnicity of a PERSON, as specified by the PERSON. The 16+1 ethnic data categories defined in the 2001 census is the national mandatory standard for the collection and analysis of	1..1	ethnicCategory >10 enumerations, please click link above to view full list.	ETHNIC CATEGORY (CR0150 from Cancer Outcomes and Services Dataset)

	ethnicity.			
Recruiting Trust ID (14860@1.0.1)	ODS code of the recruiting trust – LDP (Local Delivery Partner) or main GMC trust	1..1	organisationSiteCode	
Clinical Trial Number (34570@3.1.2)	ISRCTN number(s) of any clinical trial(s) that the patient is enrolled in. This information can be entered at a later date through resubmission of the Registration data.	0..*	isrctNumber	

5.1.2.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Registration report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

5.1.2.2 Participant Contact Details [1..1]

One report containing Participant Contact Details must be submitted together with each Registration report.

Please include all available contact details for the participant. At least one set of the contact details field MUST be supplied.

Name	Description	Multiplicity	Data Type	Related To

Participant Email Address (12529@1.0.1)	Email address of participant	0..1	emailAddress	
Participant Home Telephone (12532@1.0.1)	If available, the participant's home telephone number	0..1	ukTelephoneNumber	
Participant Mobile Telephone (12533@1.0.1)	If available, the participant's mobile telephone number	0..1	ukTelephoneNumber	
Address line 1 (12822@1.0.1)	Premises ID and/or house name, e.g. 'Flat 1', 'The Old Schoolhouse'	0..1	addressLine	
Address line 2 (12823@1.0.1)	House number, dependent thoroughfare name and descriptor without commas, e.g. '23 Mill Lane'	0..1	addressLine	
Address line 3 (12824@1.0.1)	Dependent locality/village, e.g. 'Boxgrove'	0..1	addressLine	
Address line 4 (12825@1.0.1)	Post town, e.g. 'Leeds'	0..1	addressLine	
Address line 5 (12826@1.0.1)	County (if present), e.g. 'Hampshire', 'Hants'	0..1	addressLine	
Postcode (12827@1.0.1)	The UK format Postcode, 8 character string, as per BS7666. The 8 characters field allows a space to be inserted to differentiate between the inward and outward segments of the code, enabling full use to be made of Royal Mail postcode functionality.	0..1	Postcode	

5.1.2.3 Disease Information (Tumour Sample) [1..*]

One or more reports containing Disease Information (Tumour Sample) must be submitted for each Registration report.

Disease type and subtype refer to the sample submitted to Genomics England for sequencing. This is to enable high-level grouping and analysis of the tumour type. It is understood that this may not be available at Registration or may change between Registration and submission of Core data therefore this can be provided during Core Data submissions. As this information will initially be provided in the clinic rather than the laboratory, it is included in the participant information.

Name	Description	Multiplicity	Data Type	Related To
Disease Type (12834@3.1.2)	The cancer type of the tumour sample submitted to Genomics England. The list of disease types will be validated against the types contained in Appendix A. These may be subject to change and GMCs are requested to ensure that data capture systems are flexible enough to accommodate future changes to the list of diseases contained in Appendix A. If this is unknown at registration, it can be updated as part of the patient information in the core data submissions.	1..1	xs:string	
Disease Subtype (12835@3.1.2)	The subtype of the cancer in question, recorded against a limited set of supplied enumerations. The list of disease subtypes will be validated against the subtypes contained in Appendix A. These may be subject to change and GMCs are requested to ensure that data capture systems are flexible enough to accommodate changes to the list of disease contained in Appendix A. This is to enable high-level grouping and analysis of the tumour type. It is understood that this may not be available at Registration or may change between Registration and submission of Core data. Note that the enumeration 'not_available', although available at Registration, should not be submitted for Core Disease Information Updates. If the diagnosis is not listed as a subtype it can be entered under	1..*	xs:string	

	"other". A tumour comprised of more than one subtype should be entered as follows. The predominant tumour subtype in the sample sent for whole genome sequencing should be entered first. The remaining subtypes should be entered in descending order with the most prevalant subtype in the whole tumour listed second. It is helpful to include "mixed tumour type" as a subtype but this should not be entered alone.			

5.1.2.4 Consultant Details [1..1]

One report containing Consultant Details must be submitted together with each Registration report.

Include details of the consultant responsible for the patient's clinical care, including receipt of clinical reports and communications with Genomics England. This should be completed for all participants including unaffected relatives, as the results may have individual clinical relevance for all participants. Please include the consultant's GMC number to ensure the accuracy of this record.

Name	Description	Multiplicity	Data Type	Related To
Full Name of Responsible Consultant (12774@1.0.1)	Nominated person responsible for patients clinical care and recipient of clinical reports and communications for Genomics England	1..1	xs:string	
Consultant GMC number (31254@1.0.1)	GMC number of consultant with responsibility for the patient's clinical care	1..1	consultantCode	CONSULTANT CODE (TREATMENT) (CR0660 from Cancer Outcomes and Services Dataset)
Full Name not Consultant (4495@1.0.1)	Full name of person entering data on behalf of consultant	0..1	xs:string	
Contact number (14520@1.0.1)	Phone number for the consultant.	0..1	ukTelephoneNumber	
Hospital of	ODS code of the hospital to	0..1	organisationSiteCode	

Responsible Consultant (12516@1.0.1)	which the consultant is contracted under their MAIN SPECIALTY for the purposes of the current work.			

5.1.3 Consent [1..1]

One report containing Consent must be submitted together with each Registration and Consent report.

This section reports information obtained at consent for cancer participants, including the overall consent status (consent given) and the individual questions and responses relating to the participant's options regarding additional findings (Consent Details (29742.1)).

Additional mandatory fields include full name of the person taking consent, and details of the version of the consent form and information sheet used for participation in the 100,000 Genomes Project.

The Assent form can be used for children to sign to indicate their assent to join the project, but they must legally also have a parental consent form completed. Further, full consent should be sought from a child on their 16th birthday.

No further data should be entered if the answer to the 'Consent Given' question is 'No'.

The Consent Event Date is the date of consent.

Name	Description	Multiplicity	Data Type	Related To
Name and Version of Consent Form (34549@3.1.2)	Name and Version of form used. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	1..1	xs:string	
Consent Given (12545@1.0.1)	Yes no answer to consent given	1..1	yesNo **yes** Yes **no** No	
Consent Form (12546@1.0.1)	File name of uploaded PDF copy of consent form - requested format [ParticipantId]_consent_[TimeStamp].pdf	0..1	xs:string	
Person Taking Consent	The full name of the person taking consent	1..1	xs:string	

(12547@1.0.1)				
Name and Version of Participant Information Sheet (4454@1.0.1)	Name and Version of information sheet presented. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	1..1	xs:string	
Name and Version of Assent Form (34552@3.1.2)	Name and Version of Cancer Assent form used. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	0..1	xs:string	
Assent Form (34543@1.0.1)	File name of the uploaded PDF copy of the assent form. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	0..1	xs:string	
Additional optional consent materials (40373@3.1.2)	Names and versions of consent additional consent materials used. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	0..1	xs:string	

5.1.3.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Consent report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference	Unique identifier for local record	1..1	xs:string	

(14858@1.0.1)	of clinical event or observation			

5.1.3.2 Consent Details [1..1]

One report containing Consent Details must be submitted together with each Consent report.

Details corresponding to the questions and responses on the consent form.

Name	Description	Multiplicity	Data Type	Related To
Health Related Additional Findings (34544@1.0.1)	Health-related additional findings: Does the participant want these looked for and fed back to their clinical team?	1..1	yesNo **yes** Yes **no** No	
Reproductive Additional Findings (34546@1.0.1)	Reproductive additional findings: Does the participant want these looked for and fed back to their clinical team?	0..1	yesNoNotRelevant **yes** yes **no** no **not_relevant** not relevant	

Schema 5.3 Withdrawals

Details related to a patient's withdrawal from the programme. Current protocol can be summarised as follows:

Partial withdrawal
The participant will no longer be contacted by 100,000 Genomes Project to request further samples or information. However, existing samples can still be used and information from the participant's information can still be stored and updated by Genomics England.

Full withdrawal
The participant will no longer be contacted, all samples will be destroyed and all data will be put beyond further use except for audit purposes and no further clinical information will be gathered.
Patients who have had registration data submitted but are ineligible as no tumour sample can be provided.

Withdrawal Event Date is the Date of Withdrawal of Consent on the withdrawal form.

5.3.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Withdrawals report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

5.3.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

5.3.2 Withdrawal [1..*]

One or more reports containing Withdrawal must be submitted for each Withdrawals report.

A report of withdrawal of consent.

Name	Description	Multiplicity	Data Type	Related To
Withdrawal Form	Filename of uploaded copy of scanned withdrawal form pdf - requested format is	0..1	xs:string	

(12730@1.0.1)	[ParticipantId]_withdrawal_[TimeStamp].pdf			
Withdrawal Option (12728@1.0.1)	Indicating full or partial withdrawal	1..1	consentWithdrawalOptions **full_withdrawal** — OPTION 2: FULL WITHDRAWAL: No further use **partial_withdrawal** — OPTION 1: PARTIAL WITHDRAWAL: No further contact	
Name and Version of the Withdrawal Form Used (12729@1.0.1)	Name and Version of form used - list of names and versions available from genomicsengland.co.uk/library-and-resources/	1..1	genomicsEnglandConsentWithdrawalForms **6a** — Withdrawal information and form – for adult or child participants (6a) **6b** — Consultee declaration of advice regarding adult participant withdrawal information – for consultees (withdrawal) (6b)	
Person Reporting Withdrawal (12731@1.0.1)	Full name, including forenames and surname, of person reporting withdrawal.	1..1	xs:string	

5.3.2.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Withdrawal report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To

Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6 Core Data

The core data required for a cancer participant consists of reports on clinical events in the existing medical history, and should be supplied within eight weeks of sample collection.

All of the clinical events correspond to items in a specific longitudinal record model, in which each piece of data reported is associated with a date, usually the date of a relevant clinical event and one or more relevant diagnoses.

Every submission must include a set of participant identifiers.

Schema 6.1 Disease Information Update (Tumour Sample)

Disease type and subtype of the sample submitted to GEL for sequencing. This is to enable high-level grouping and analysis of the tumour type. It is understood that this may not be available at Registration or may change between Registration and submission of Core data therefore this can be provided during Core data submissions.

The Disease Information Update Event Date will be the date that the information was updated.

Multiple instances of this section can be provided where multiple tumour samples are submitted with different disease types.

Name	Description	Multiplicity	Data Type	Related To
Disease Type (12834@3.1.2)	The cancer type of the tumour sample submitted to Genomics England. The list of disease types will be validated against the types contained in Appendix A. These may be subject to change and GMCs are requested to ensure that data capture systems are flexible enough to accommodate future changes to the list of diseases contained in Appendix A. If this is unknown at registration,	1..1	xs:string	

	it can be updated as part of the patient information in the core data submissions.			
Disease Subtype (12835@3.1.2)	The subtype of the cancer in question, recorded against a limited set of supplied enumerations. The list of disease subtypes will be validated against the subtypes contained in Appendix A. These may be subject to change and GMCs are requested to ensure that data capture systems are flexible enough to accommodate changes to the list of disease contained in Appendix A. This is to enable high-level grouping and analysis of the tumour type. It is understood that this may not be available at Registration or may change between Registration and submission of Core data. Note that the enumeration 'not_available', although available at Registration, should not be submitted for Core Disease Information Updates. If the diagnosis is not listed as a subtype it can be entered under “other”. A tumour comprised of more than one subtype should be entered as follows: The predominant tumour subtype in the sample sent for whole genome sequencing should be entered first. The remaining subtypes should be entered in descending order with the most prevalant subtype in the whole tumour listed second. It is helpful to include “mixed tumour type” as a subtype but this should not be entered alone.	1..*	xs:string	

Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.1.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Disease Information Update (Tumour Sample) report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)

Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.1.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.1.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Disease Information Update (Tumour Sample) report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

Schema 6.2 Risk Factor Assessment

Clinical event corresponding to the assessment of risk factors for a cancer participant.
The Risk Factor Assessment Event Date will be the date that the risk factors were recorded.

6.2.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Risk Factor Assessment report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.2.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To

NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.2.2 General Risk Factors [0..*]

Multiple reports containing General Specific Risk Factors can be submitted together with each Risk Factors report.

Risk factors for each participant.

Name	Description	Multiplicity	Data Type		Related To
Smoking (14446@3.1.2)	Specify the current smoking status of the patient	0..1	smokingStatus		SMOKING STATUS (LU10190 from Cancer Outcomes and Services Dataset)
			1	Current smoker	
			2	Ex smoker	
			3	Non-smoker - history unknown	
			4	Never smoked	

			Z	Not Stated (PERSON asked but declined to provide a response)	
			9	Unknown	
Alcohol Consumption (14447@1.0.1)	The ALCOHOL WEEKLY UNITS reported by the patient.	0..1	xs:nonNegativeInteger		
Height (4531@1.0.1)	Person height / length in metres to 2 decimal places. Height and weight to be used to calculate BMI as an indicator of the patient being overweight or obese. Provide the most relevant information that will inform this. Will relate to new data item in COSD v7, CR6430 PERSON OBSERVATION HEIGHT IN METERS	0..1	personHeightInMetres		
Weight (14760@1.0.1)	Weight in kg. Height and weight to be used to calculate BMI as an indicator of the patient being overweight or obese. Provide the most relevant information that will inform this. Will relate to new data item in COSD v7, CR6440 PERSON OBSERVATION (WEIGHT)	0..1	personObservationWeight		

6.2.2.1 Event Details [1..1]

One report containing Event Details must be submitted together with each General Risk Factors report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.2.3 Cancer Specific Risk Factors [0..*]

Multiple reports containing Cancer Specific Risk Factors can be submitted together with each Risk Factors report.

Submission of cancer specific risk factors is optional however, if risk factors are submitted, one of the specific risk factors must be provided

6.2.3.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Cancer Specific Risk Factors report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.2.3.2 Risk Factors [1..1]

One report containing Risk Factors must be submitted together with each Cancer Specific Risk Factors report.

Choice of risk factors for specific cancers:

6.2.3.2.1 Risk Factors for Ovarian Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Age of Menarche (14473@3.1.2)	The age in years at the first menstrual period.	0..1	Age	
Age of Menopause (14474@3.1.2)	The age in years at which menstruation ceased.	0..1	Age	
Duration of OCP (14475@3.1.2)	The duration of oral contraceptive use, in years.	0..1	durationInYears	
Duration of HRT (14476@3.1.2)	The duration of hormone replacement therapy, in years.	0..1	durationInYears	
Number of Pregnancies (14477@3.1.2)	The total number of pregnancies (including live births, still births, miscarriages and termination of pregnancies)	0..1	xs:nonNegativeInteger	
Number of Births (14478@3.1.2)	This is the number of registrable live births by the participant	0..1	xs:nonNegativeInteger	
Endometriosis (14479@3.1.2)	Medical diagnosis of endometriosis made.	0..1	positiveNegativeUnknown **unknown** unknown **negative** negative **positive** positive	
Previous Tubal Ligation (14480@3.1.2)	Previous Tubal Ligation	0..1	yesNoUnk **yes** Yes **no** No **unknown** Unknown	
Use of IUD (14481@3.1.2)	Intrauterine Device (IUD) ever used for duration of over 1 month	0..1	yesNoUnk **yes** Yes **no** No **unknown** Unknown	
Use of Non-Steroidal Anti Inflammatory Drugs (14482@3.1.2)	Any episode of chronic use = NSAIDS used more than half the days of the week, more than half the weeks of the year for over 1 year	0..1	yesNoUnk **yes** Yes **no** No	

			unknown Unknown	
Number of Children Breastfed (29104@3.1.2)	Number of children breastfed over 3 months duration	0..1	xs:nonNegativeInteger	
Number Cycles IVF (29105@3.1.2)	Number of cycles of IVF	0..1	xs:nonNegativeInteger	

or in the case of,

6.2.3.2.2 Risk Factors for Breast Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Age of Menarche (14473@3.1.2)	The age in years at the first menstrual period.	0..1	Age	
Age of Menopause (14474@3.1.2)	The age in years at which menstruation ceased.	0..1	Age	
Duration of OCP (14475@3.1.2)	The duration of oral contraceptive use, in years.	0..1	durationInYears	
Duration of HRT (14476@3.1.2)	The duration of hormone replacement therapy, in years.	0..1	durationInYears	
Number of Pregnancies (14477@3.1.2)	The total number of pregnancies (including live births, still births, miscarriages and termination of pregnancies)	0..1	xs:nonNegativeInteger	
Number of Children Breastfed (29104@3.1.2)	Number of children breastfed over 3 months duration	0..1	xs:nonNegativeInteger	
Number Cycles IVF (29105@3.1.2)	Number of cycles of IVF	0..1	xs:nonNegativeInteger	
Breast Density (29108@3.1.2)	Breast density at most recent available pre-surgical mammogram: based on percentage of fibroglandular tissue relative to total area on the two view mammogram.	0..1	breastDensity **birads_0** additional imaging evaluation and/or comparison to prior mammogram is	

				needed
			birads_1	glandular tissue is less than 25%
			birads_2	scattered fibroglandular densities (25-50%)
			birads_3	heterogeneously dense (50-75%)
			birads_4	extremely dense breast (75-100%)

or in the case of,

6.2.3.2.3 Risk Factors for Glioma Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Radiotherapy in Childhood (38862@3.1.2)	Was radiotherapy received in childhood for a Central Nervous System (CNS) or non CNS tumour?	1..1	radiotherapyInChildhood **cns** CNS **non_cns** non_CNS **none** none **unknown** unknown	

or in the case of,

6.2.3.3 Risk Factors for Endometrial Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Age of Menarche (14473@3.1.2)	The age in years at the first menstrual period.	0..1	Age	
Age of Menopause	The age in years at which	0..1	Age	

(14474@3.1.2)	menstruation ceased.			
Duration of OCP (14475@3.1.2)	The duration of oral contraceptive use, in years.	0..1	durationInYears	
Duration of HRT (14476@3.1.2)	The duration of hormone replacement therapy, in years.	0..1	durationInYears	
Number of Pregnancies (14477@3.1.2)	The total number of pregnancies (including live births, still births, miscarriages and termination of pregnancies)	0..1	xs:nonNegativeInteger	
Tamoxifen use age (39005@3.1.2)	If treated with tamoxifen, age in years at which treatment started.	0..1	Age	

or in the case of,

6.2.3.3.1 Risk Factors for Renal Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Dialysis Duration (39011@3.1.2)	Number of years dialysis received	1..1	xs:nonNegativeInteger	

or in the case of,

6.2.3.3.2 Risk Factors for Malignant Melanoma [1..1]

Name	Description	Multiplicity	Data Type	Related To
Childhood Chronic Exposure (39012@3.1.2)	Number of years spent living in a country with high UV light between 0 and15 years.	0..1	childhoodChronicExposure	
Sunbed use age (39014@3.1.2)	Age, in years, when sunbed first used.	0..1	Age	
Skin type (39016@3.1.2)	Skin type according to the Fitzpatrick Scale. Link http://archderm.jamanetwork.com/article.aspx?articleid=549509	0..1	Skintype **i** I Always burns, never tans	

			ii	II Usually burns, tans minimally
			iii	III Sometimes mild burn, tans uniformly
			iv	IV Burns minimally, always tans well
			v	V Very rarely burns, tans very easily
			vi	VI Never burns, never tans
			unknown	unknown

or in the case of,

6.2.3.3.3 Risk Factors for Testicular Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Cryptorchidism (39019@3.1.2)	Presence or history of cryptorchidism (the absence of one or both testes from the scrotum).	0..1	yesNoUnk **yes** Yes **no** No **unknown** Unknown	
Cryptorchidism Age (39097@3.1.2)	If participant has a history of cryptorchidism, age at its correction in years.	0..1	Age	

or in the case of,

6.2.3.3.4 Risk Factors for HPB Cancer [1..1]

Name	Description	Multiplicity	Data Type	Related To
Hepatitis B infection (39025@3.1.2)	History of hepatitis B infection	0..1	infectionHistory **none** none **previous** previous **current** current **unknown** unknown	
Hepatitis C infection (39028@3.1.2)	History of hepatitis C infection	0..1	infectionHistory **none** none **previous** previous **current** current **unknown** unknown	
Cirrhosis (39030@3.1.2)	History of cirrhosis, total duration in years.	0..1	durationInYears	

or in the case of,

Schema 6.3 Investigations

Investigation events may be associated with multiple diagnosis events.

Investigations include imaging, sample investigation, including biopsies, with associated pathology, and tumour markers, and other investigation (e.g. blood tests).

The Investigation Event Date will be the date of the reported investigation.

If investigations aren't imaging investigations or sample investigations please use the generic 'Investigations - Other' class to record the appropriate results and metadata.

6.3.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Investigations report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.3.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.3.2 Imaging [0..*]

Multiple reports containing imaging can be submitted together with each investigations report.

If the investigation involves imaging, the following should be included:

Name	Description	Multiplicity	Data Type	Related To
Imaging Modality	*IMAGING CODE (NICIP) and/or *IMAGING CODE	0..1	cancerImagingModality >10 enumerations, please	CANCER IMAGING MODALITY (CR0330 from Cancer

(29094@3.1.2)	(SNOMED CT) and/or *CANCER IMAGING MODALITY and IMAGING ANATOMICAL SITE and ANATOMICAL SIDE (IMAGING) is required. The type of imaging procedure used during an Imaging or Radiodiagnostic Event for a Cancer Care Spell. NB: PET Scan also includes PET-CT Scan.		click link above to view full list.	Outcomes and Services Dataset)
Anatomical Site (12753@3.1.2)	A classification of the part of the body that is the subject of an Imaging Or Radiodiagnostic Event. The coding frame used is the OPCS-4 'Z' coding, plus two additional local codes: Whole body CZ001 Multiple sites CZ002	0..1	imagingAnatomicalSite	IMAGING ANATOMICAL SITE (CR0340 from Cancer Outcomes and Services Dataset)
Anatomical Side (33444@3.1.2)	The side of the body that is the subject of an Imaging or Radiodiagnostic Event.	0..1	anatomicalSideImaging **L** Left **R** Right **M** Midline **B** Bilateral **8** Not applicable **9** Not Known	ANATOMICAL SIDE (IMAGING) (CR3000 from Cancer Outcomes and Services Dataset)
Imaging Report Reference (29096@3.1.2)	This is an internal reference that will allow your centre to retrieve the imaging report associated with this imaging event	1..1	xs:string	
Image File Reference (14897@3.1.2)	If not possible to submit Image File, please supply Image File local reference, according to local imaging guidance.	0..*	xs:string	
Ultrasound Examination Result (42053@3.1.2)	Relates to COSD v7, CR6000. Result of the ultrasound examination. For example in Breast Cancer, this will	0..1	ultrasoundExaminationResult **U1** Normal	

	normally be the result of the ultrasound examination of the breast undertaken at the first outpatient appointment at the breast clinic. If the patient attends more than one breast clinic, the result of each ultrasound examination of the breast should be recorded.		**U2** Benign **U3** Indeterminate/probably benign **U4** Suspicious of malignancy **U5** Highly suspicious of malignancy	
Imaging Report Text (42266@3.1.2)	This is the full text provided in the imaging report.	0..1	xs:string	IMAGING REPORT TEXT (CR0160 from Cancer Outcomes and Services Dataset)
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.3.2.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Imaging report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the

same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.2.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Imaging report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.3.2.3 Sample Details [0..1]

A maximum of one report containing Sample Details can be submitted together with each Imaging report. The Imaging report can be submitted without this information.

All sample investigations relating to germline or tumour molecular genetics should have the following:

Name	Description	Multiplicity	Data Type	Related To
Local Sample ID (12762@1.0.1)	The local identifier for the source sample	0..1	xs:string	
Sample Taken Date (12760@1.0.1)	The date upon which the sample was taken	0..1	xs:date	SAMPLE COLLECTION DATE (CR1010 from Cancer Outcomes and Services Dataset)
Sample Receipt Date (12761@1.0.1)	The date upon which the sample was received at the laboratory	0..1	xs:date	SAMPLE RECEIPT DATE (CR0770 from Cancer Outcomes and Services Dataset)

6.3.2.4 Imaging Code [1..1]

One report containing Imaging Code must be submitted together with each Imaging report.

Choice of SNOMED CT or NICIP imaging codes.

Name	Description	Multiplicity	Data Type	Related To
Imaging Code (SNOMEDCT) (12752@3.1.2)	*IMAGING CODE (NICIP) and/or *IMAGING CODE (SNOMED CT) and/or *CANCER IMAGING MODALITY and IMAGING ANATOMICAL SITE and ANATOMICAL SIDE (IMAGING) is required. IMAGING CODE (NICIP) is the National Interim Clinical Imaging Procedure Code Set code which is used to identify both the test modality and body site of the test	1..1	snomedCt	IMAGING ANATOMICAL SITE (CR0340 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Imaging Code (NICIP)	*IMAGING CODE (NICIP) and/or	1..1	imagingCode(NICIP)	IMAGING CODE (NICIP) (CR1610

(33449@3.1.2)	*IMAGING CODE (SNOMED CT) and/or *CANCER IMAGING MODALITY and IMAGING ANATOMICAL SITE and ANATOMICAL SIDE (IMAGING) is required. IMAGING CODE (NICIP) is the National Interim Clinical Imaging Procedure Code Set code which is used to identify both the test modality and body site of the test.			from Cancer Outcomes and Services Dataset) PROCEDURE DATE (CANCER IMAGING) (CR0320 from Cancer Outcomes and Services Dataset)

6.3.2.5 Cancer Specific Imaging [0..1]

A maximum of one report containing Cancer Specific Imaging can be submitted together with each Imaging report. The Imaging report can be submitted without this information.

Submission of cancer specific risk imaging is optional however, if cancer specific imaging is submitted, one of the following must be provided:

6.3.2.5.1 Imaging (Breast) [1..1]

Priority COSD data items from BREAST - IMAGING. To carry imaging mammogram, ultrasound and axilla ultrasound details for breast cancer.

Name	Description	Multiplicity	Data Type	Related To
Mammogram result (38883@3.1.2)	Result of the mammogram. This will normally be the result of the mammogram taken at the first outpatient appointment at the breast clinic. If the patient attends more than one breast clinic, the result of each mammogram should be recorded.	1..1	mammogramResult **R1** Normal **R2** Benign **R3** Uncertain **R4** Suspicious **R5** Malignant	MAMMOGRAM RESULT (BR4050 from Cancer Outcomes and Services Dataset)

or in the case of,

6.3.2.5.2 Imaging (CNS) [1..1]

Priority COSD data items from CNS - IMAGING. To carry imaging details for CNS cancer.

Name	Description	Multiplicity	Data Type	Related To
Lesion location (radiological) (38930@3.1.2)	Radiologically determined anatomical location of lesion (largest lesion if more than one) or where centred. This is recorded prior to treatment.	0..1	locationSurgical >10 enumerations, please click link above to view full list.	LESION LOCATION (RADIOLOGICAL) (BA3000 from Cancer Outcomes and Services Dataset)
Number of lesions (radiological) (38931@3.1.2)	Radiologically determined number of lesions. (From UPPER GI - STAGING - LIVER HCC)	0..1	numberOfLesionsRadiological	NUMBER OF LESIONS (RADIOLOGICAL) (UG14540 from Cancer Outcomes and Services Dataset)
Principal diagnostic imaging type (38934@3.1.2)	Indicate the principal imaging procedure undertaken to diagnose the tumour. NB: PET Scan also includes PET-CT Scan	0..1	principalDiagnosticImagingType **1** CT Scan **2** MRI Scan **3** PET Scan	PRINCIPAL DIAGNOSTIC IMAGING TYPE (BA3050 from Cancer Outcomes and Services Dataset)
Features of largest lesion (radiological) (38933@3.1.2)	Radiologically identified features of the largest lesion such as density, necrosis recorded pre treatment. This may involve selection of more than one value.	0..1	featuresOfLargestLesionRadiological **01** Contrast-enhancement **02** Calcification **03** Mass effect **04** Hydrocephalus **05** Haemorrhage **06** Cystic/multi-cystic **07** Dural tail **08** Brain oedema **09** Cord signal change **10** Cord compression	FEATURES OF LARGEST LESION (RADIOLOGICAL) (BA3040 from Cancer Outcomes and Services Dataset)
Lesion Size (Radiological)	Radiological estimate in millimetres of the maximum	0..1	lesionSizeRadiological	LESION SIZE (RADIOLOGICAL) (BA3030

(38932@3.1.2)	diameter of the tumour measured prior to treatment (largest lesion if more than one). Record as "0" to indicate not assessable for diffuse tumours (e.g. gliomatosis cerebri).			from Cancer Outcomes and Services Dataset)

6.3.3 Sample Pathology [0..*]

Multiple reports containing sample pathology can be submitted together with each Investigations report.

Clinical event corresponding to a pathology report.

We expect to receive at least one or more Sample Pathology records for each participant and for each Diagnosis.

The Sample Pathology Event Date will be the date of final authorisation of the corresponding pathology report.

The pathology report, as well as Event Details, should include Morphology, Topography, and, where applicable, Cancer Specific Tumour Markers, Cancer Specific Pathology and Cancer Specific Grading information.

Name	Description	Multiplicity	Data Type	Related To
Primary Diagnosis (ICD Pathological) (38876@3.1.2)	PRIMARY DIAGNOSIS (ICD PATHOLOGICAL) is the PRIMARY DIAGNOSIS based on the evidence from a pathological examination.	0..1	primaryDiagnosisIcd	PRIMARY DIAGNOSIS (ICD PATHOLOGICAL) (CR0810 from Cancer Outcomes and Services Dataset)
Primary Diagnosis (SNOMED CT Pathological) (42275@3.1.2)	PRIMARY DIAGNOSIS (SNOMED CT PATHOLOGICAL) is the PRIMARY DIAGNOSIS based on the evidence from a pathological examination. Multiple SNOMED CT codes may be provided.	0..*	snomedCt	
Pathology Investigation Type	The type of pathology investigation carried out. Although this item	1..1	pathologyInvestigationType >10 enumerations, please click link above to	

(14903@3.1.2)	is based on COSD CR0760, an additional value 'BM' for Bone Marrow Aspirate has been added for the purposes of this project in order to collect haematological bone marrow aspirate samples.		view full list.	
Excision Margin (14904@3.1.2)	An indication of whether the excision margin was clear of the tumour and if so, by how much. Where there is more than one measurement, record the closest or closest relevant margin. Where actual measurements are not taken use options 01, 05 or 06 as applicable.	0..1	excisionMargin >10 enumerations, please click link above to view full list.	EXCISION MARGIN (CR0880 from Cancer Outcomes and Services Dataset)
Grade of Differentiation (14905@3.1.2)	GRADE OF DIFFERENTIATION (PATHOLOGICAL) is the definitive grade of the Tumour based on the evidence from a pathological examination. Not applicable for CNS, Haematology, Melanoma and Sarcoma. Either Grade of Differentiation and/or individual cancer specific grading should be provided, where available. Where applicable, both may be provided, e.g. in the case of urothelial cancers. Where cancer specific grading is available and provided, GX should be	1..1	gradeOfDifferentiation(pathological) **G4** Undifferentiated / anaplastic **G3** Poorly differentiated **G2** Moderately differentiated **G1** Well differentiated **GX** Grade of differentiation is not appropriate or cannot be assessed	GRADE OF DIFFERENTIATION (PATHOLOGICAL) (CR0860 from Cancer Outcomes and Services Dataset)

	recorded for general Grade of Differentiation.			
Pathology Report (14907@3.1.2)	The full text from the pathology report (uploaded copy of pathology report)	0..1	pathologyReportText	PATHOLOGY REPORT TEXT (CR1020 from Cancer Outcomes and Services Dataset)
Number of Nodes Examined (14908@3.1.2)	Number of nodes examined, where applicable	0..1	numberOfNodesExamined	NUMBER OF NODES EXAMINED (CR0890 from Cancer Outcomes and Services Dataset)
Number of Nodes Positive (38882@3.1.2)	The number of local and regional nodes reported as being positive for the presence of Tumour metastases (in this specimen report only)	0..1	numberOfNodesPositive	NUMBER OF NODES POSITIVE (CR0900 from Cancer Outcomes and Services Dataset)
Pathology Image File Reference (14912@3.1.2)	Image of the section from block submitted to GEL	0..*	xs:string	
Tumour Type (14721@3.1.2)	The type of the tumour sampled and sent for sequencing For haematological cancers only 'primary' is applicable.	0..1	tumourType **primary** — Primary; source of cancer tumour sample **recurrence_of_primary_tumour** — Recurrence; a tumour has returned at the site of the original cancer **metastatic_recurrence** — Metastatic (different cancer site) which developed	

			and was sampled after presentation	
			metastases	Metastatic (different cancer site) which was present and sampled at diagnosis instead of the primary tumour
Pre-operative Therapy (35534@3.1.2)	Has the patient received pre-operative therapy?	1..1	yesNoUnk **yes**: Yes **no**: No **unknown**: Unknown	
Investigation Result Date (40107@3.1.2)	The date on which an investigation was concluded e.g. the date the result was authorised.	0..1	xs:date	INVESTIGATION RESULT DATE (UG14500 from Cancer Outcomes and Services Dataset)
Service Report Identifier (39088@3.1.2)	Priority COSD data item from CORE - PATHOLOGY DETAILS. A unique identifier of a SERVICE REPORT. max an18	0..1	serviceReportIdentifier	SERVICE REPORT IDENTIFIER (UG14510 from Cancer Outcomes and Services Dataset)
Cancer vascular or lymphatic invasion (38881@3.1.2)	An indication of the presence or absence of unequivocal tumour in lymphatic and/or vascular spaces.	0..1	cancerVascularOrLymphaticInvasion **NU**: No - vascular/lymphatic invasion not present **YU**: Yes - vascular/lymphatic invasion present **YV**: Vascular invasion only present **YL**: Lymphatic invasion only present	CANCER VASCULAR OR LYMPHATIC INVASION (CR0870 from Cancer Outcomes and Services Dataset)

			YB Both lymphatic and vascular invasion present" **UU** Uncertain whether vascular/lymphatic invasion is present or not **XX** Cannot be assessed **99** Not Known	
Tumour Size (29075@3.1.2)	Maximum dimension of the largest tumour in mm on the histopathology report.	0..1	diameterInMm	
Pre-invasive Elements (14872@3.1.2)	Description of atypia or in situ disease, if present. Input needs to be surrounded by double quotes i.e. "xxxx, xxxx xxxx"	0..1	xs:string	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16	1..*	tumourID	

	alphanumeric characters i.e. RN3_A098BC			

6.3.3.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Sample Pathology report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.3.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Sample Pathology report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis	Optionally, provide the related cancer diagnosis as SNOMED CT	0..*	diagnosisCode(snomedCt)	

(SNOMEDCT) (35539@3.1.2)	code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.			

6.3.3.3 Sample Details [0..1]

A maximum of one report containing Sample Details can be submitted together with each Sample Pathology report. The Sample Pathology report can be submitted without this information.

All sample investigations relating to germline or tumour molecular genetics should have the following:

Name	Description	Multiplicity	Data Type	Related To
Local Sample ID (12762@1.0.1)	The local identifier for the source sample	0..1	xs:string	
Sample Taken Date (12760@1.0.1)	The date upon which the sample was taken	0..1	xs:date	SAMPLE COLLECTION DATE (CR1010 from Cancer Outcomes and Services Dataset)
Sample Receipt Date (12761@1.0.1)	The date upon which the sample was received at the laboratory	0..1	xs:date	SAMPLE RECEIPT DATE (CR0770 from Cancer Outcomes and Services Dataset)

6.3.3.4 Morphology [1..*]

One or more reports containing Morphology must be submitted for each Sample Pathology report.

Choice of ICD03 or SNOMED morphology codes.

Name	Description	Multiplicity	Data Type	Related To
Morphology (ICD) (14871@3.1.2)	The morphology code for the diagnosed cancer as defined by ICDO3. This can be recorded as well as or instead of	1..1	morphology(icd)	MORPHOLOGY (ICDO3) (CR0180 from Cancer Outcomes and Services Dataset)

	MORPHOLOGY (SNOMED).			
Or in the case of,				
Morphology (SNOMEDCT) (31244@3.1.2)	The morphology code for the diagnosed cancer as defined by SNOMED CT. This can be recorded as well as or instead of MORPHOLOGY (ICD).	1..1	morphology(snomedCt)	MORPHOLOGY (SNOMED CT) (CR3070 from Cancer Outcomes and Services Dataset) MORPHOLOGY (SNOMED) (CR0850 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Morphology (SNOMEDRT) (31243@3.1.2)	The morphology code for the diagnosed cancer as defined by SNOMED RT. This can be recorded as well as or instead of MORPHOLOGY (ICD).	1..1	morphology(snomed)	MORPHOLOGY (SNOMED) (CR0850 from Cancer Outcomes and Services Dataset)

Or in the case of,

6.3.3.4.1 Morphology (SNOMED) [1..1]

This is the morphology of the tumour as categorised by SNOMED and the version of SNOMED.

Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content.

Name	Description	Multiplicity	Data Type	Related To
Morphology (SNOMED) (42048@3.1.2)	This is the morphology of the tumour as categorised by SNOMED International / SNOMED CT Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content	1..1	snomed	
SNOMED version	The version of SNOMED used to encode MORPHOLOGY	1..1	snomedVersion	

(42049@3.1.2)	(SNOMED) and TOPOGRAPHY (SNOMED) Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content		**01** SNOMED II **02** SNOMED 3 **03** SNOMED 3.5 **04** SNOMED RT **05** SNOMED CT **99** Not Known	

6.3.3.5 Topography [0..1]

A maximum of one report containing Topography can be submitted together with each Sample Pathology report. The Sample Pathology report can be submitted without this information.

Choice of ICD03 or SNOMED topography codes.

Name	Description	Multiplicity	Data Type	Related To
Topography (ICD) (31228@3.1.2)	This is the topographical site of the tumour as categorised by ICD03	1..1	topographyIcdo3	Morphology (ICD) (Cancer Model) MORPHOLOGY (ICDO3) (CR0180 from Cancer Outcomes and Services Dataset) TOPOGRAPHY (ICDO3) (CR0480 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Topography (SNOMEDCT) (14876@3.1.2)	This is the topographical site of the tumour as categorised by SNOMED CT.	1..1	topographySnomedCt	TOPOGRAPHY (SNOMED CT) (CR3060 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Topography (SNOMEDRT)	This is the topographical site of the tumour as categorised by	1..1	topographySnomed	TOPOGRAPHY (SNOMED) (CR0530 from Cancer Outcomes and

(31227@3.1.2)	SNOMED RT			Services Dataset)

Or in the case of,

6.3.3.5.1 Topography (SNOMED) [1..1]

This is the topographical site of the tumour as categorised by SNOMED International / SNOMED CT.

Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017.

Name	Description	Multiplicity	Data Type	Related To
Topography (SNOMED) (42052@3.1.2)	This is the topographical site of the tumour as categorised by SNOMED International / SNOMED CT Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017.	1..1	snomed	
SNOMED version (42049@3.1.2)	The version of SNOMED used to encode MORPHOLOGY (SNOMED) and TOPOGRAPHY (SNOMED) Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content	1..1	snomedVersion **01** SNOMED II **02** SNOMED 3 **03** SNOMED 3.5 **04** SNOMED RT **05** SNOMED CT	

			99	Not Known	

6.3.3.6 pTNM [0..1]

A maximum of one report containing pTNM can be submitted together with each Sample Pathology report. The Sample Pathology report can be submitted without this information.

Record the details of the Union for International Cancer Control (UICC) pathological Tumour, Node and Metastasis (TNM) staging for cancer and the UICC version used.

Name	Description	Multiplicity	Data Type	Related To
TNM Version (14909@3.1.2)	Relates to CR2070 v7.0 - The AJCC (Skin) or UICC edition number used for Tumour, Node and Metastasis (TNM) staging for cancer diagnosis.	1..1	tnmEditionNumber	Integrated TNM Version (Cancer Model) TNM EDITION NUMBER (CR2070 from Cancer Outcomes and Services Dataset)
pT (14910@3.1.2)	T CATEGORY (PATHOLOGICAL) is the Union for International Cancer Control (UICC) code which classifies the size and extent of the primary Tumour based on the evidence from a pathological examination. (COSD User Guidance: not applicable for CNS, Gynaecology, Haematology, stageable Skin and most CTYA diagnosis. Please see site specific datasets for further information on collecting this data item, including the site specific values to be used.) See COSD User Guidance for recommended staging to be collected for individual cancer sites.	0..1	tCategoryPathological	T CATEGORY (PATHOLOGICAL) (CR0910 from Cancer Outcomes and Services Dataset)
pN (14911@3.1.2)	N CATEGORY (PATHOLOGICAL) is the Union for International Cancer Control (UICC) code which classifies the absence or presence and extent of regional	0..1	nCategoryPathological	N CATEGORY (PATHOLOGICAL) (CR0920 from Cancer Outcomes and Services Dataset)

	lymph node metastases based on the evidence from a pathological examination.			
pM (29098@3.1.2)	M CATEGORY (PATHOLOGICAL) is the Union for International Cancer Control (UICC) code which classifies the absence or presence of distant metastases based on the evidence from a pathological examination.	0..1	mCategoryPathological	M CATEGORY (PATHOLOGICAL) (CR0930 from Cancer Outcomes and Services Dataset)

6.3.3.7 Cancer Specific Tumour Markers [0..1]

A maximum of one report containing Cancer Specific Tumour Markers can be submitted together with each Sample Pathology report. The Sample Pathology report can be submitted without this information.

Cancer specific markers for colorectal or breast or lung or ovarian or childhood tumours.

Where applicable, and if available, one set of cancer specific tumour markers information should be submitted as part of the pathology information submission.
If submitted, one of the following must be submitted together with each Cancer Specific Tumour Markers report

6.3.3.7.1 Colorectal Tumour Markers [1..1]

Name	Description	Multiplicity	Data Type	Related To
MLH1 IHC (14916@3.1.2)	Indication of biomarkers presence	0..1	biomarkerPresent **present** present **not_tested** no tested **equivocal** equivocal **absent** absent	
MSH2 IHC (14917@3.1.2)	Indication of biomarkers presence	0..1	biomarkerPresent **present** present **not_tested** no tested **equivocal** equivocal	

			absent absent	
MSH6 IHC (14918@3.1.2)	Indication of biomarkers presence	0..1	biomarkerPresent **present** present **not_tested** no tested **equivocal** equivocal **absent** absent	
PMS2 IHC (14919@3.1.2)	Indication of biomarkers presence	0..1	biomarkerPresent **present** present **not_tested** no tested **equivocal** equivocal **absent** absent	

or in the case of,

6.3.3.7.2 Ovarian Tumour Markers [1..1]

Name	Description	Multiplicity	Data Type	Related To
WT1 IHC (14923@3.1.2)	Indication of biomarkers presence	1..1	biomarkerPresent **present** present **not_tested** no tested **equivocal** equivocal **absent** absent	

or in the case of,

6.3.3.7.3 Lung Tumour Markers [1..1]

Name	Description	Multiplicity	Data Type	Related To

Epidermal growth factor receptor mutational status (38900@3.1.2)	Relates to COSD v7, LU10090. Epidermal Growth Factor Receptor Mutational Status	1..1	epidermalGrowthFactorReceptorMutationalStatus *1* *Wild type* *2* *Mutation* **3** Failed analysis **4** Not assessed **5** Wild type/non-sensitising mutation **6** Sensitising/activating mutation	

or in the case of,

6.3.3.7.4 Childhood Tumour Markers [1..1]

Name	Description	Multiplicity	Data Type	Related To
Molecular subgroup (medulloblastoma) (38992@3.1.2)	molecular subgroup (medulloblastoma)	0..1	molecularSubgroupMedulloblastoma **not_tested** Not tested **shh** SHH **wnt** WNT **non_shh** Non-SHH **non_wnt** Non-WNT	
TP53 (medulloblastoma) (38993@3.1.2)	TP53 (medulloblastoma)	0..1	alkBlastomaMarker **not_tested** not tested **wild_type** wild-type **mutant** mutant	
MYC (medulloblastoma and neuroblastoma) (38994@3.1.2)	MYC (medulloblastoma and neuroblastoma)	0..1	blastomaMarker **not_tested** not tested **amplified** amplified **non_amplified** non amplified	

MYCN (medulloblastoma and neuroblastoma) (38995@3.1.2)	MYCN (medulloblastoma and neuroblastoma)	0..1	blastomaMarker **not_tested** not tested **amplified** amplified **non_amplified** non amplified	
ALK (neuroblastoma) (38998@3.1.2)	ALK (neuroblastoma)	0..1	alkBlastomaMarker **not_tested** not tested **wild_type** wild-type **mutant** mutant	
Chromosomal abnormality (neuroblastoma) (39000@3.1.2)	Chromosomal abnormality in neuroblastoma tumour	0..1	chromosomalabnormalityNeuroblastoma **not_tested** not tested **segmental** segmental **non_segmental** non-segmental	

6.3.3.8 Cancer Specific Grading [0..1]

A maximum of one report containing Cancer Specific Grading can be submitted together with each Sample Pathology report. The Sample Pathology report can be submitted without this information.

Cancer specific grades.

Where applicable, and if available, one set of cancer specific grading information should be submitted as part of the pathology information submission.

One of the following must be submitted together with each Cancer Specific Grading report

Name	Description	Multiplicity	Data Type	Related To
Fuhrman Grade (33070@3.1.2)	Specific Grading for Renal Cancer. Fuhrman grade according to RCP guidance: https://www.rcpath.org/Resources/RCPath/Migrated%20Resources/Documents/G/G037FINAL_AdultrenaldatasetNov06.pdf	1..1	fuhrmanGradingSystem **1** 1 **2** 2	

			3 3 **4** 4	
Or in the case of,				
Figo Grade (33065@3.1.2)	Specific Grading for Ovarian and Endometrial Cancer Figo Grade as per updated RCPath guidelines: http://www.rcpath.org/Resources/RCPath/Migrated%20Resources/Documents/C/CEU_FIGO1988vs2013ovarianstaging_Dec13.pdf	1..1	figoGrade **1** 1 **2** 2 **3** 3	
Or in the case of,				
Invasive Grade (Breast) (33062@3.1.2)	Specific Grading for Breast Cancer as defined by the modified Bloom–Richardson: https://www.rcpath.org/Resources/RCPath/Migrated%20Resources/Documents/P/PosterFinal.pdf The Bloom–Richardson grading system from 1957[1] refers to a breast cancer classification system to grade breast cancers, and was the precursor of the present criteria, the modified Bloom–Richardson–Elston grading system (also called the Nottingham system.)[2][3]	1..1	breastInvasiveTumourGrade **1** 1 **2** 2 **3** 3 **x** Not assessable	
Or in the case of,				
DCIS Grade (33063@3.1.2)	DCIS based on nuclear grade Please see Ref: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp58-low-resolution.pdf for further guidance	1..1	dcisTumourGrade **l** Low **i** Intermediate **h** High **x** Not assessable	
Or in the case of,				
Serum tumour markers (38869@3.1.2)	TESTICULAR ONLY. Based on serum tumour markers AFP, HCG and LDH. For Testicular Cancer S category is an additional prognostic factor.	1..1	SCATEGORY **SX** Marker studies not available or not performed **S0** Normal **S1** LDH (UNITS/LITRE) -	S-CATEGORY (UR15030 from Cancer Outcomes and Services Dataset)

				Less than 1.5 x normal, HCG (MILLIUNITS/MILLILITRE) - Less than 5,000, AFP (NANOGRAMS/MILLILITRE) - Less than 1,000	
			S2	LDH (UNITS/LITRE) - 1.5-10 x normal, HCG (MILLIUNITS/MILLILITRE) - 5,000-50,000, AFP (NANOGRAMS/MILLILITRE) - 1,000-10,000	
			S3	LDH (UNITS/LITRE) - Greater than 10 x normal, HCG (MILLIUNITS/MILLILITRE) - Greater than 50,000, AFP (NANOGRAMS/MILLILITRE) - Greater than 10,000	
Or in the case of,					
WHO tumour grade (CNS) (38941@3.1.2)	The grade of the tumour using WHO classification for tumours of the central nervous system. FOR INTRA AXIAL AND EXTRA AXIAL ONLY.	1..1	whoTumourGradeCns		WHO TUMOUR GRADE (CNS) (BA3160 from Cancer Outcomes and Services Dataset)
			1	I	
			2	II	
			3	III	
			4	IV	

Or in the case of,				
Glioma (WHO 2007) (39032@3.1.2)	Glioma (WHO 2007)	1..1	gliomagrading **i** I **ii** II **iii** III **iv** IV	
Or in the case of,				
Sarcomatoid change (39036@3.1.2)	As per core data items in RCPath minimum data set for renal cell carcinoma.	1..1	sarcomatoidGrading **present** present **absent** absent	
Or in the case of,				
Histopathological tumour grade sarcoma (39038@3.1.2)	The 3-grade French system, as per RCPath sarcoma minimum dataset.	1..1	frenchGradingSystem **g1** G1-Well differentiated (Low grade) **g2** G2-Moderately differentiated (Intermediate grade) **g3** G3-Poorly differentiated (High grade)	
Or in the case of,				
Leibovich score (39039@3.1.2)	The Leibovich score is a scoring algorithm to predict survival for patients with metastatic renal cell carcinoma. Please provide, if applicable. 0-11, not more than 12	1..1	leibovichScore	
Or in the case of,				
Tumour Grade (Ovarian Serous) (41993@3.1.2)	Specify the grade of the tumour. For serous tumours specify whether High or Low grade, as per RCPath dataset. Note that this data item relates to COSD GY7150, however 'I' =Intermediate is not applicable for serous tumours.	1..1	tumourGradeOvarianSerous **l** Low **h** High	

Or in the case of,				
Tumour grade (Urology) (38918@3.1.2)	BLADDER ONLY. Specify whether LOW, HIGH Grade or PUNLMP (Papillary Urothelial Neoplasm of Low Maligant Potential). Note that while punlmp is enumerated for COSD, punlmp is not eligible for collection for The Project.	1..1	tumourGradeUrology **L** Low **H** High **P** Punlmp **X** Not applicable	TUMOUR GRADE (UROLOGY) (UR15290 from Cancer Outcomes and Services Dataset)

Or in the case of,

6.3.3.8.1 Gleason Grade [1..1]

Name	Description	Multiplicity	Data Type	Related To
Gleason Grade (Primary) (33071@3.1.2)	What is the most extensive Gleason grade? Specific Grading for Prostate Cancer. Please see: Epstein JI et al Am J Surg Path 2005: 29: 1228-42 Pierorazio PM et al. BJU Int 2013: 111: 753-60 for further guidance.	1..1	gleasonGrade **1** 1 **2** 2 **3** 3 **4** 4 **5** 5	GLEASON GRADE (PRIMARY) (UR15210 from Cancer Outcomes and Services Dataset)
Gleason Grade (Secondary) (33512@3.1.2)	If additional grades are present, what is the highest grade (biopsy) or the second most extensive grade (TURP and radicals). Specific Grading for Prostate Cancer. Please see: Epstein JI et al Am J Surg Path 2005: 29: 1228-42 Pierorazio PM et al. BJU Int 2013: 111: 753-60 for further guidance.	0..1	gleasonGrade **1** 1 **2** 2 **3** 3 **4** 4 **5** 5	GLEASON GRADE (SECONDARY) (UR15220 from Cancer Outcomes and Services Dataset)

6.3.3.9 Cancer Specific Pathology [0..1]

A maximum of one report containing Cancer Specific Pathology can be submitted together with each Sample Pathology report. The Sample Pathology report can be submitted without this information.

Priority COSD data items for cancer specific pathology reports.

One of the following must be submitted together with each Cancer Specific Pathology report

6.3.3.10 Pathology (Bladder) [1..1]

Priority COSD data items from UROLOGY - PATHOLOGY - BLADDER. To carry the cancer pathology details for Bladder.

Name	Description	Multiplicity	Data Type		Related To
Detrusor muscle presence indicator (38917@3.1.2)	BLADDER ONLY Presence or absence of detrusor muscle in the specimen	1..1	detrusorMusclePresenceIndicator		DETRUSOR MUSCLE PRESENCE INDICATOR (UR15120 from Cancer Outcomes and Services Dataset)
			1	Present	
			2	Absent	
			9	Not known	

or in the case of,

6.3.3.11 Pathology (Breast) [1..1]

Priority COSD data items from BREAST - PATHOLOGY. To carry pathology details for breast cancer.

Name	Description	Multiplicity	Data Type		Related To
Core biopsy (Breast) (38887@3.1.2)	Needle core biopsy opinion.	0..1	coreBiopsyBreast		CORE BIOPSY (BREAST) (BR4260 from Cancer Outcomes and Services Dataset)
			B1	Normal	
			B2	Benign	
			B3	Uncertain malignant potential	
			B4	Suspicious	
			B5a	Malignant (In situ)	

			B5b Malignant (Invasive) **B5c** Malignant (Not assessable)	
Core biopsy (node) (38888@3.1.2)	Needle biopsy opinion on axillary lymph node.	0..1	coreBiopsyNode **B1** Normal **B2** Benign **B3** Uncertain malignant potential **B4** Suspicious **B5** Malignant	CORE BIOPSY (NODE) (BR4270 from Cancer Outcomes and Services Dataset)
Cytology (node) (38886@3.1.2)	Cytology opinion on axillary lymph node.	0..1	cytologyNode **C1** Inadequate/unsatisfactory specimen **C2** Benign **C3** Uncertain **C4** Suspicious of malignancy **C5** Malignant	CYTOLOGY (NODE) (BR4250 from Cancer Outcomes and Services Dataset)
ER ALLRED Score (14925@3.1.2)	ER ALLRED score (range 0-8)	0..1	allredScore	
PR ALLRED Score (14926@3.1.2)	Record the PR ALLRED score if ER status is negative. (Range 0-8)	0..1	allredScore	
HER2 Status (14927@3.1.2)	Human epidermal growth factor receptor 2	0..1	her2Status **P** Positive **N** Negative **B** Borderline **X** Not performed	HER2 STATUS (BR4280 from Cancer Outcomes and Services Dataset)
HER2 ISH Status	Record the result of the ISH	0..1	her2IshStatus	HER2 ISH STATUS (BR4310 from Cancer Outcomes and

(33079@3.1.2)	(in-situ hybridization) test. This is only required if the initial HER2 status is "Borderline".		**P** Positive **N** Negative	Services Dataset)
Distance to Margin (35533@3.1.2)	Distance to closest relevant margin (mm). Distance to nearest margin whether invasive or non invasive.	0..1	distanceToMargin	DISTANCE TO MARGIN (BR4210 from Cancer Outcomes and Services Dataset)

or in the case of,

6.3.3.12 Pathology (CNS) [1..1]

Pathology (Central Nervous System). Priority COSD data items from CNS - PATHOLOGY. To carry pathology details for CNS cancer.

Name	Description	Multiplicity	Data Type	Related To
Molecular diagnostics code (38940@3.1.2)	Chromosomal or genetic markers associated with the brain tumour. This may involve selection of more than one values for each tumour. Updated to reflect COSD v7. Enumerations deprecated in COSD v7 may still be supplied until full compliance with v7 is achieved.	1..1	molecularDiagnosticsCode >10 enumerations, please click link above to view full list.	MOLECULAR DIAGNOSTICS CODE (BA3070 from Cancer Outcomes and Services Dataset)

or in the case of,

6.3.3.13 Pathology (Endometrial) [1..1]

Priority COSD data items from GYNAECOLOGY - PATHOLOGY - ENDOMETRIAL. To carry pathology details for Gynae - Endometrial.

Name	Description	Multiplicity	Data Type	Related To

Involvement of cervical stroma (38948@3.1.2)	Is there microscopic involvement of cervical stroma?	0..1	yesNoNotAssessable		INVOLVEMENT OF CERVICAL STROMA (GY7240 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			X	Not Assessable	
			9	Not Known	
Distance to serosa (38947@3.1.2)	Specify the tumour free distance to the serosa	0..1	distanceToSerosa		DISTANCE TO SEROSA (GY7220 from Cancer Outcomes and Services Dataset)
Parametrium involvement (38952@3.1.2)	Is there microscopic involvement of parametrium?	0..1	yesNoNotAssessable		PARAMETRIUM INVOLVEMENT (GY7270 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			X	Not Assessable	
			9	Not Known	
Peritoneal washings (38953@3.1.2)	Were peritoneal washings submitted and if so were malignant cells seen?	0..1	peritonealWashings		PERITONEAL WASHINGS (GY7280 from Cancer Outcomes and Services Dataset)
			1	Positive	
			2	Negative	
			X	Not sent/Not assessable	
Myometrial invasion (38951@3.1.2)	Is there microscopic evidence of myometrial invasion?	0..1	myometrialInvasion		MYOMETRIAL INVASION (GY7260 from Cancer Outcomes and Services Dataset)
			1	None	
			2	Less than 50%	
			3	Greater than or	

			equal to 50%	

or in the case of,

6.3.3.14 Pathology (Gynaecology) [1..1]

Priority COSD data items from GYNAECOLOGY - PATHOLOGY. To carry pathology details for gynaecology.

Name	Description	Multiplicity	Data Type	Related To
Fallopian tube involvement (38942@3.1.2)	For endometrial and epithelial/ovarian cancers, is there microscopic involvement of fallopian tubes	0..1	tubeInvolvement **1** Not involved **2** Right involved **3** Left involved **4** Both involved **X** Not assessable	FALLOPIAN TUBE INVOLVEMENT (GY7050 from Cancer Outcomes and Services Dataset)
Ovarian involvement (38943@3.1.2)	For endometrial and fallopian cancers, is there microscopic involvement of ovaries	0..1	tubeInvolvement **1** Not involved **2** Right involved **3** Left involved **4** Both involved **X** Not assessabl	OVARIAN INVOLVEMENT (GY7120 from Cancer Outcomes and Services Dataset)

			e	
Serosal involvement (38944@3.1.2)	For endometrial, epithelial/ovarian and fallopian cancers, is there microscopic involvement of uterine serosa	0..1	yesNoNotAssessable **Y** Yes **N** No **X** Not Assessable **9** Not Known	SEROSAL INVOLVEMENT (GY7130 from Cancer Outcomes and Services Dataset)
Omental involvement (38945@3.1.2)	For endometrium, ovary, fallopian tube and primary peritoneum cancers, is there involvement of the omentum?	0..1	omentalInvolvement **1** Involved - deposit size not specified **2** Involved - deposit(s) 20mm or less **3** Involved - deposit(s) greater than 20mm **4** Not involved **X** Not assessable/Not sent	OMENTAL INVOLVEMENT (GY7100 from Cancer Outcomes and Services Dataset)
Nodes examined number (para-aortic) (38954@3.1.2)	The number of para-aortic nodes examined. Use 0 if nodes not sent.	0..1	nodesExaminedNumberParaAortic	NODES EXAMINED NUMBER (PARA-AORTIC) (GY7060 from Cancer Outcomes and Services Dataset)
Nodes positive number (para-	The number of para-aortic nodes reported as being	0..1	nodesPositiveNumberParaAortic	NODES POSITIVE NUMBER (PARA-AORTIC) (GY7080

aortic) (38955@3.1.2)	positive for the presence of tumour metastases.			from Cancer Outcomes and Services Dataset)
Extranodal spread (38958@3.1.2)	Is there evidence of extranodal spread/extension?	0..1	yesNoNotAssessable **Y** Yes **N** No **X** Not Assessable **9** Not Known	EXTRANODAL SPREAD (GY7230 from Cancer Outcomes and Services Dataset)
Nodes examined number (pelvic) (38956@3.1.2)	The number of pelvic nodes examined (Not applicable for vulval cancers). Use 0 if nodes not sent	0..1	nodesExaminedNumberPelvic	NODES EXAMINED NUMBER (PELVIC) (GY7070 from Cancer Outcomes and Services Dataset)
Nodes positive number (pelvic) (38957@3.1.2)	The number of pelvic nodes reported as being positive for the presence of tumour metastases. (Not applicable for vulval cancers)	0..1	nodesPositiveNumberPelvic	NODES POSITIVE NUMBER (PELVIC) (GY7090 from Cancer Outcomes and Services Dataset)

or in the case of,

6.3.3.15 Pathology (Kidney) [1..1]

Priority COSD data items from UROLOGY - PATHOLOGY - KIDNEY. To carry the cancer pathology details for Kidney.

Name	Description	Multiplicity	Data Type	Related To
Renal vein tumour (38922@3.1.2)	Is there evidence of tumour thrombus in the renal vein?	0..1	yesNoUnc **Y** Yes **N** No **U** Uncertain	RENAL VEIN TUMOUR (CT6650 from Cancer Outcomes and Services Dataset)

Perinephric fat invasion (38920@3.1.2)	Is there evidence of perinephric fat invasion?	0..1	yesNoUnc		PERIRENAL FAT INVASION (CT6630 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			U	Uncertain	
Adrenal invasion (38921@3.1.2)	Is there evidence of direct adrenal invasion?	0..1	yesNo		ADRENAL INVASION (UR15150 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
Gerotas fascia invasion (38923@3.1.2)	Is there evidence of invasion into Gerota's fascia?	0..1	yesNo		GEROTA'S FASCIA INVASION (UR15170 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	

or in the case of,

6.3.3.16 Pathology (Lung) [1..1]

Priority COSD data items from LUNG - PATHOLOGY. To carry Pathology details for Lung Carcinoma (Most items are only applicable where patients have surgical resection).

Name	Description	Multiplicity	Data Type		Related To
Extent of pleural invasion (38901@3.1.2)	What is the extent of pleural invasion	0..1	extentOfPleuralInvasion		EXTENT OF PLEURAL INVASION (LU10120 from Cancer Outcomes and Services Dataset)
			1	No pleural invasion	
			2	Visceral pleura only	
			3	Parietal pleura/chest wall	
			4	Mediastinal pleura	

Malignant pleural effusion (38902@3.1.2)	Is there evidence of malignant pleural effusion?	0..1	yesNoNk **Y** Yes **N** No **9** Not known	MALIGNANT PLEURAL EFFUSION (LU10170 from Cancer Outcomes and Services Dataset)
Satellite tumour nodules location (38903@3.1.2)	Record the most distant location of separate tumour nodules.	0..1	satelliteTumourNodulesLocation **1** Separate tumour nodules in same lobe **2** Separate tumour nodules in a different ipsilateral lobe **3** Separate tumour nodules in a contralateral lobe **4** No separate tumour nodules **9** Not known	SATELLITE TUMOUR NODULES LOCATION (LU10180 from Cancer Outcomes and Services Dataset)

or in the case of,

6.3.3.17 Pathology (Prostate) [1..1]

Priority COSD data items from UROLOGY - PATHOLOGY - PROSTATE. To carry the cancer pathology details for Prostate.

Name	Description	Multiplicity	Data Type	Related To
Gleason grade (tertiary) (38924@3.1.2)	Is there a different third grade in addition the primary and secondary grades and what is its value?	0..1	gleasonGradeTertiary	GLEASON GRADE (TERTIARY) (UR15230 from Cancer Outcomes and Services Dataset)
Perineural invasion (38925@3.1.2)	Is there perineural invasion (invasion into perineurium of nerve bundles - PNI)	0..1	yesNoNotAssessable **Y** Yes	PERINEURAL INVASION (SK12530 from Cancer Outcomes and Services Dataset)

			N No **X** Not Assessable **9** Not Known	
Organ confined (38926@3.1.2)	If prostatectomy was performed, is the tumour confined to the prostate?	0..1	yesNoNa **Y** Yes **N** No **X** Not applicable	ORGAN CONFINED (UR15250 from Cancer Outcomes and Services Dataset)
Seminal vesicles invasion (38927@3.1.2)	If prostatectomy was performed, is there invasion into Seminal Vesicles?	0..1	yesNoNa **Y** Yes **N** No **X** Not applicable	SEMINAL VESICLES INVASION (UR15260 from Cancer Outcomes and Services Dataset)
TURP tumour percentage (38928@3.1.2)	For Transurethral resection of prostate (TURP) only, what percentage of tumour is clinically unsuspected tumour.	0..1	turpTumourPercentage	TURP TUMOUR PERCENTAGE (UR15270 from Cancer Outcomes and Services Dataset)

or in the case of,

6.3.3.18 Pathology (Testes) [1..1]

Priority COSD data items from UROLOGY - PATHOLOGY - TESTICULAR. To carry the cancer pathology details for Testicular.

Name	Description	Multiplicity	Data Type	Related To
Rete testes invasion (38929@3.1.2)	For Seminoma only, does the tumour invade the rete testis.	1..1	yesNoNa **Y** Yes	RETE TESTES INVASION (UR15310 from Cancer Outcomes and Services Dataset)

			N No **X** Not applicable	

6.3.4 Genetic Results [0..*]

Multiple reports containing genetic results can be submitted together with each Investigations report.

Enter all abnormal genetic results and all pertinent negative results from this sample. Use one entry per gene.

Name	Description	Multiplicity	Data Type	Related To
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.3.4.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Genetic Results report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.4.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Genetic Results report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.3.4.3 Sample Details [0..1]

A maximum of one report containing Sample Details can be submitted together with each Genetic Results report. The Genetic Results report can be submitted without this information.

All sample investigations relating to germline or tumour molecular genetics should have the following:

Name	Description	Multiplicity	Data Type	Related To
Local Sample ID (12762@1.0.1)	The local identifier for the source sample	0..1	xs:string	
Sample Taken Date (12760@1.0.1)	The date upon which the sample was taken	0..1	xs:date	SAMPLE COLLECTION DATE (CR1010 from Cancer Outcomes and Services Dataset)
Sample Receipt Date (12761@1.0.1)	The date upon which the sample was received at the laboratory	0..1	xs:date	SAMPLE RECEIPT DATE (CR0770 from Cancer Outcomes and Services Dataset)

6.3.4.4 Genetic Result [1..200]

A minimum of one report containing Genetic Result must be submitted together with each Genetic Results report. Multiple reports may be submitted.

If the investigation produced genetic results, then the genetic investigation should be extended with the following items, for each genetic result

GUIDANCE: Enter all abnormal genetic results and all pertinent negative results from this sample. Use one entry per gene.

Name	Description	Multiplicity	Data Type	Related To
Genetic Test Laboratory (4563@1.0.1)	Was this test performed in a diagnostic or research laboratory?	0..1	geneticTestLaboratory **research_laboratory** Research laboratory **diagnostic_laboratory** Diagnostic laboratory	
Test Scope (6101@1.0.1)	The gene coded according to HGNC. Enter	1..1	geneScope	

	'genomewide' if genomewide, e.g. karyotype or aCGH.			
Scope Qualifiers (12764@1.0.1)	If whole locus or coding sequence of gene not covered, give details of regions covered, e.g. 'exons 3 and 8'	0..1	xs:string	
Method of Test (12765@1.0.1)	The method used to investigate the gene(s). If copy number analysis has been performed for a subset of genes, please enter separately from sequencing results	1..1	geneticTestMethod >10 enumerations, please click link above to view full list.	
Test Result (12744@1.0.1)	(for molecular results) If no defect was observed please report 'normal'; if a mutation is detected that is considered pathogenically or clinically important record 'mutation detected'; if no reliable result could be determined please report 'fail'.	1..1	molecularTestResult **normal** Normal (negative) **fail** Fail **abnormalitydetected** Pathogenic abnormality detected **vus** Variant of unknown significance detected	
Abnormal Molecular Result (14900@1.0.1)	Record the details of the abnormal genotype using Genomic Coordinates	0..1	xs:string	
Genome Build (34224@1.0.1)	Record the relevant human genome build if an abnormal genotype is specified if applicable	0..1	xs:string	
Abnormal Cytogenetic Result (34225@1.0.1)	Record the details of the cytogenetic abnormality using IGCN standards	0..1	xs:string	

6.3.5 Next Generation Sequencing [0..*]

Multiple reports containing next generation sequencing can be submitted together with each Investigations report.

Next Generation Sequencing performed outside of Genomics England.

Name	Description	Multiplicity	Data Type	Related To
Sequence Report (12766@3.1.2)	Reference to uploaded copy of test report	0..1	xs:string	
Sequence File (12767@1.0.1)	Local sequence file reference or uploaded copy of VCF	0..1	xs:string	
Comments (5182@1.0.1)	Follow-up comments	0..1	xs:string	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.3.5.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Next Generation Sequencing report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.5.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Next Generation Sequencing report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.3.6 Circulating Tumour Markers [0..*]

Multiple reports containing cancer tumour markers can be submitted together with each investigations report.

If submitted, one of the following must be submitted together with each Circulating Tumour Markers report

6.3.6.1 Circulating Tumour Markers (Prostate) [1..1]

Name	Description	Multiplicity	Data Type	Related To
psa (14929@3.1.2)	PROSTATE ONLY. Prostate Specific Antigen blood level in ng/ml, measured at time of diagnosis.	0..1	psaDiagnosis	PSA (DIAGNOSIS) (UR15070 from Cancer Outcomes and Services Dataset)
psa (pre treatment) (38916@3.1.2)	PROSTATE ONLY. Prostate Specific Antigen blood level in ng/ml, measured before treatment (including second and subsequent treatments). This is the PSA taken prior to EACH treatment (because some curative treatments may be delivered years after diagnosis.	0..1	psaPreTreatment	PSA (PRE TREATMENT) (UR15080 from Cancer Outcomes and Services Dataset)
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters	1..*	tumourID	

	i.e. RN3_A098BC			

6.3.6.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Circulating Tumour Markers (Prostate) report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.6.3 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Circulating Tumour Markers (Prostate) report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code.	0..*	diagnosisCode(snomedCt)	

(35539@3.1.2)	Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.			

or in the case of,

6.3.6.4 Circulating Tumour Markers (Ovarian) [1..1]

Name	Description	Multiplicity	Data Type	Related To
CA125 (14930@3.1.2)	Protein level	1..1	xs:double	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.3.6.5 Event Details [1..1]

One report containing Event Details must be submitted together with each Circulating Tumour Markers (Ovarian) report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.6.6 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Circulating Tumour Markers (Ovarian) report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.3.7 Investigation Report Other [0..*]

Multiple reports containing other investigations can be submitted together with each investigations report

This class is a generic mechanism to include additional investigation reports that are clinically relevant but have not been specified within the current set of investigations.

For every report please include the name of the type of report and the ID of the report type i.e.
Name: Electrolytes panel - Blood
ID: 55231-5
Reporting Standard: LOINC

In addition to the report name please provide one or more attributes that correspond to the report attribute model below. These can either be test results or metadata associated with the report.

Name	Description	Multiplicity	Data Type	Related To
Report Name (33106@3.1.2)	If the investigation is part of a standardised report / panel of tests, please populate the name that identifies the type of report i.e. Renal Biopsy	1..1	xs:string	
Report Code (33096@3.1.2)	If the investigation is part of a standardised report / panel of tests, please populate the code that identifies the type of report - if available	0..1	xs:string	
Reporting Standard (33454@3.1.2)	The standard (name and version) for the report - if available	0..1	xs:string	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention:	1..*	tumourID	

	Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC			

6.3.7.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Investigation Report Other report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.3.7.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Investigation Report Other report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to	1..*	primaryDiagnosisIcd	

	more than one diagnosis.			
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.3.7.3 Report Attribute [1..*]

One or more reports containing Report Attribute must be submitted for each Investigation Report Other report.

For every report please include one or more attributes corresponding to the model below. These can either be test results and/or metadata associated with the report i.e. a report of a urine dip could be submitted with the following attributes:

Data Standard: LOINC
ID: 2947-0
Name: Sodium [Moles/volume] in Blood
Value: 130
DateTime: 2015-09-11T21:32:52
Unit of Measure: 01

Data Standard: LOINC
ID: 6298-4
Name: Potassium [Moles/volume] in Blood
Value: 4.6
DateTime: 2015-09-11T21:32:52
Unit of Measure: 01

etc.

NOTE: Upper Range and Lower Range could be included in each attribute but they are not mandatory

Name	Description	Multiplicity	Data Type	Related To
Standard (33104@3.1.2)	The name and version of the data standard of the attribute type i.e. SNOMED CT, OPCS-4, COSD	0..1	xs:string	

Code (33097@3.1.2)	Code of the report attribute type - if available i.e. if the attribute name is "POSITIVE PROXIMAL OR DISTAL RESECTION MARGIN" and the standard is COSD then the attribute code would be CO5190	1..1	xs:string	
Name (33098@3.1.2)	Name of the test included in the investigation and/or the name of the metadata associated with the investigation i.e. POSITIVE PROXIMAL OR DISTAL RESECTION MARGIN	1..1	xs:string	
Value (33100@3.1.2)	The result of the investigation i.e. Margin involved	1..1	xs:string	
DateTime (33455@3.1.2)	Date the attribute value was recorded	1..1	xs:dateTime	
Upper Range (33101@3.1.2)	For quantitative tests, the upper range associated with the investigation within the lab - if applicable	0..1	xs:string	
Lower Range (33102@3.1.2)	For quantitative tests, the lower range associated with the investigation within the lab - if applicable	0..1	xs:string	
Report (33456@3.1.2)	File / Report associated with the investigation - if applicable	0..1	xs:string	
Unit of Measure (33099@3.1.2)	Unit of measure used to record the investigation result - if applicable	0..1	unitOfMeasurement >10 enumerations, please click link above to view full list.	

Schema 6.4 Diagnosis

We expect to receive at least one Diagnosis record per tumour for each participant.

A diagnosis event will usually correspond to when a diagnosis of cancer is agreed or confirmed. If multiple instances of cancer are diagnosed at the same time, then each should be reported as a separate diagnosis event, and each of these reports should have a different locally-allocated event reference.

The Diagnosis Event Date will be the date when cancer was confirmed or diagnosis agreed, typically the date the specimen was taken, recorded on the pathology report, which confirms the cancer.

The report of the diagnosis, in addition to Event Details, should include Morphology, Topography, and, where applicable, Integrated TNM Staging information and/or Cancer Specific Staging and Cancer Specific Diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Diagnosis (ICD) (33183@3.1.2)	The icd code for the agreed diagnosis	1..1	primaryDiagnosisIcd	PRIMARY DIAGNOSIS (ICD) (CR0370 from Cancer Outcomes and Services Dataset)
Diagnosis (SNOMED CT) (35538@3.1.2)	Optionally provide the SNOMED CT code for the diagnosis in addition to the ICD code.	0..*	diagnosisCode(snomedCt)	
Recurrence Indicator (14938@3.1.2)	An indication of whether a diagnosis of recurrence has been recorded for which a new Cancer Care Plan is required.	0..1	cancerRecurrenceCarePlanIndicator **YL** Yes, including local recurrence **YD** Yes, not including local recurrence **NN** No, not recurrence	CANCER RECURRENCE CARE PLAN INDICATOR (CR0450 from Cancer Outcomes and Services Dataset)
Metastatic Site (14937@3.1.2)	The site of the metastatic disease, if any, at diagnosis	0..1	metastaticSite **02** Brain **03** Liver	METASTATIC SITE (CR1590 from Cancer Outcomes and Services Dataset)

			04	Lung	
			06	Multiple metastatic sites	
			07	Unknown metastatic site	
			08	Skin	
			09	Distant lymph nodes	
			10	Bone (excluding Bone Marrow)	
			11	Bone marrow	
			99	Other metastatic site	
Basis Of Diagnosis (14939@3.1.2)	The basis of diagnosis of cancer records show how a cancer was identified. Please use the NHS data dictionary definition of this attribute.	1..1	basisOfDiagnosis(cancer)		BASIS OF DIAGNOSIS (CANCER) (CR0390 from Cancer Outcomes and Services Dataset)
			2	Clinical Investigation: Includes all diagnostic techniques (e.g. X-rays, endoscopy, imaging, ultrasound,	
			1	Clinical: Diagnosis made before death but without the benefit of any of the following (2-7)	
			0	Death Certificate: The only information available is from a death certificate	
			7	Histology of a primary tumour: Histological examination of tissue from the primary tumour, however obtained,	

				including all cutting and bone marrow biopsies. Also includes autopsy specimens of a primary tumour	
			6	Histology of a metastasis: Histological examination of tissues from a metastasis, including autopsy specimens	
			5	Cytology: Examination of cells whether from a primary or secondary site, including fluids aspirated using endoscopes or needles. Also including microscopic examination of peripheral blood films and trephine bone marrow aspirates	
			4	Specific tumour markers: Includes biochemical and/or immunological markers which are specific for a tumour site	
			9	Unknown: No information on how the diagnosis has been made (e.g. PAS or HISS record only)	
Tumour Laterality	Tumour laterality identifies	0..1	tumourLaterality(pathological)		

(14902@3.1.2)	the side of the body for a tumour relating to paired organs within a PATIENT based on the evidence from a pathological examination.		**B** Bilateral **R** Right **L** Left **M** Midline **9** Not Known **8** Not applicable	
Grade of Differentiation (At Diagnosis) (40402@3.1.2)	GRADE OF DIFFERENTIATION (AT DIAGNOSIS) is the definitive grade of the Tumour at the time of PATIENT DIAGNOSIS. COSD Guidance: Required for all Urological cancers except prostate and testis cancer. This data item is not applicable to CNS, Sarcoma or Haematology diagnosis.	0..1	gradeOfDifferentiationAtDiagnosis **GX** Grade of differentiation is not appropriate or cannot be assessed **G1** Well differentiated **G2** Moderately differentiated **G3** Poorly differentiated **G4** Undifferentiated / anaplastic	GRADE OF DIFFERENTIATION (AT DIAGNOSIS) (CR0410 from Cancer Outcomes and Services Dataset)
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters	1..1	tumourID	

	i.e. RN3_A098BC			

6.4.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Diagnosis report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.4.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				

CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.4.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Diagnosis report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.4.3 Morphology [1..*]

One or more reports containing Morphology must be submitted for each Diagnosis report.

Choice of ICD03 or SNOMED morphology codes.

Name	Description	Multiplicity	Data Type	Related To
Morphology (ICD) (14871@3.1.2)	The morphology code for the diagnosed cancer as defined by ICDO3. This can be recorded as well as or instead of MORPHOLOGY (SNOMED).	1..1	morphology(icd)	MORPHOLOGY (ICDO3) (CR0180 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Morphology (SNOMEDCT) (31244@3.1.2)	The morphology code for the diagnosed cancer as defined by SNOMED CT. This can be recorded as well as or instead of MORPHOLOGY (ICD).	1..1	morphology(snomedCt)	MORPHOLOGY (SNOMED CT) (CR3070 from Cancer Outcomes and Services Dataset) MORPHOLOGY (SNOMED) (CR0850 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Morphology (SNOMEDRT) (31243@3.1.2)	The morphology code for the diagnosed cancer as defined by SNOMED RT. This can be recorded as well as or instead of MORPHOLOGY (ICD).	1..1	morphology(snomed)	MORPHOLOGY (SNOMED) (CR0850 from Cancer Outcomes and Services Dataset)

Or in the case of,

6.4.3.1 Morphology (SNOMED) [1..1]

This is the morphology of the tumour as categorised by SNOMED and the version of SNOMED.

Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content.

Name	Description	Multiplicity	Data Type	Related To
Morphology (SNOMED) (42048@3.1.2)	This is the morphology of the tumour as categorised by SNOMED International / SNOMED CT	1..1	snomed	

	Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content			
SNOMED version (42049@3.1.2)	The version of SNOMED used to encode MORPHOLOGY (SNOMED) and TOPOGRAPHY (SNOMED) Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content	1..1	snomedVersion **01** SNOMED II **02** SNOMED 3 **03** SNOMED 3.5 **04** SNOMED RT **05** SNOMED CT **99** Not Known	

6.4.4 Topography [0..1]

A maximum of one report containing Topography can be submitted together with each Diagnosis report. The Diagnosis report can be submitted without this information.

Choice of ICD03 or SNOMED topography codes.

If submitted, one of the following must be submitted together with each Topography report.

Name	Description	Multiplicity	Data Type	Related To
Topography (ICD) (31228@3.1.2)	This is the topographical site of the tumour as categorised by ICD03	1..1	topographyIcdo3	Morphology (ICD) (Cancer Model) MORPHOLOGY (ICDO3) (CR0180 from Cancer Outcomes and Services Dataset) TOPOGRAPHY (ICDO3) (CR0480 from Cancer Outcomes and Services Dataset)

Or in the case of,				
Topography (SNOMEDCT) (14876@3.1.2)	This is the topographical site of the tumour as categorised by SNOMED CT.	1..1	topographySnomedCt	TOPOGRAPHY (SNOMED CT) (CR3060 from Cancer Outcomes and Services Dataset)
Or in the case of,				
Topography (SNOMEDRT) (31227@3.1.2)	This is the topographical site of the tumour as categorised by SNOMED RT	1..1	topographySnomed	TOPOGRAPHY (SNOMED) (CR0530 from Cancer Outcomes and Services Dataset)

Or in the case of,

6.4.4.1 Topography (SNOMED) [1..1]

This is the topographical site of the tumour as categorised by SNOMED International / SNOMED CT.

Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017.

Name	Description	Multiplicity	Data Type	Related To
Topography (SNOMED) (42052@3.1.2)	This is the topographical site of the tumour as categorised by SNOMED International / SNOMED CT Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017.	1..1	snomed	
SNOMED version (42049@3.1.2)	The version of SNOMED used to encode MORPHOLOGY (SNOMED) and TOPOGRAPHY (SNOMED)	1..1	snomedVersion **01** SNOMED II	

	Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content		**02** SNOMED 3 **03** SNOMED 3.5 **04** SNOMED RT **05** SNOMED CT **99** Not Known	

6.4.5 Integrated TNM [0..1]

A maximum of one report containing Integrated TNM can be submitted together with each Diagnosis report. The Diagnosis report can be submitted without this information.

Record Integrated TNM stage of the tumour with TNM version where applicable.

Name	Description	Multiplicity	Data Type	Related To
Integrated TNM Version (14943@3.1.2)	The UICC edition number used for Tumour, Node and Metastasis (TNM) staging for cancer diagnosis.	1..1	tnmEditionNumber	TNM EDITION NUMBER (CR2070 from Cancer Outcomes and Services Dataset)

6.4.5.1 TNM Details [1..2]

A minimum of one report containing TNM Details must be submitted together with each Integrated TNM report. Multiple reports may be submitted.

Record overall TNM stage grouping of the tumour OR component T, N and M stage.

If submitted, one of the following must be submitted together with each TNM Details report

Name	Description	Multiplicity	Data Type	Related To
TNM Stage	Record the overall TNM stage	1..1	tnmStageGroupingIntegrat	TNM STAGE GROUPING

Grouping (Integrated) (14942@3.1.2)	grouping of the tumour, derived from each T, N and M component after treatment. This classification is based on all the evidence available to the clinician(s) with responsibility for assessing the patient. Such evidence arises from physical examination, imaging, endoscopy, biopsy, surgical exploration and other relevant examinations. The overall integrated TNM stage grouping indicates the tumour stage after treatment and/or after all available evidence has been collected. Note: Use UICC coding.		ed	(INTEGRATED) (CR0610 from Cancer Outcomes and Services Dataset)

Or in the case of,

6.4.5.2 Component TNM [1..1]

Component T, N, M scores reported individually.

Name	Description	Multiplicity	Data Type	Related To
Integrated T (14944@3.1.2)	Tumour stage, if integrated TNM not supplied. This is the UICC code which classifies the size and extent of the primary tumour after treatment and/or after all available evidence has been collected.	1..1	tCategoryIntegratedStage	T CATEGORY (INTEGRATED STAGE) (CR0620 from Cancer Outcomes and Services Dataset)
Integrated N (14945@3.1.2)	Nodes stage, if integrated TNM not supplied. This is the UICC code which classifies the absence or presence and extent of regional lymph node metastases after treatment and/or after all available evidence has been collected	1..1	nCategoryIntegratedStage	N CATEGORY (INTEGRATED STAGE) (CR0630 from Cancer Outcomes and Services Dataset)
Integrated M (14946@3.1.2)	Metastasis stage, if integrated TNM not supplied. This is the UICC code which classifies the	0..1	mCategoryIntegratedStage	M CATEGORY (INTEGRATED STAGE) (CR0640 from Cancer Outcomes and Services

	absence or presence of distant metastases after treatment and/or after all available evidence has been collected.			Dataset)

6.4.6 Cancer Specific Staging [0..1]

A maximum of one report containing Cancer Specific Staging can be submitted together with each Diagnosis report. The Diagnosis report can be submitted without this information.

Cancer specific staging data.

Where applicable, and if available, the most appropriate cancer specific staging details should be submitted as part of the diagnosis submission.

See COSD guidance for the recording of cancer site specific staging.

Name	Description	Multiplicity	Data Type		Related To
International Neuroblastoma Risk Group (INRG) Staging System (38874@3.1.2)	Related to COSD v7 CT7050. The International Neuroblastoma Risk Group Staging System (INRGSS) was designed for the International Neuroblastoma Risk Group (INRG) pre-treatment classification system. Unlike the INSS (above), the INRGSS uses only the results of imaging tests taken before surgery. It does not include surgical results or spread to lymph nodes to determine the stage. Knowledge regarding the presence or absence of image defined risk factors (IDRF) is required for this staging system. (See COSD v7 User Guide for more information)	0..1	inrgNeuroblastoma		
			L1	Stage L1: The Tumour is located only in the area where it started; no IDRFs are found on imaging scans, such as CT or MR	
			L2	Stage L2: The tumour has not spread beyond the area where it started and the nearby tissue; IDRFs are found on imaging scans, such as CT or MR	
			M	Stage M: The tumour has spread to other parts of the body (except stage MS, see below)	

			MS	Stage MS: The tumour has spread to only the skin, liver, and/or bone marrow (less than 10% marrow involvement) in patients less than 18 months	
Modified Dukes Stage (33023@3.1.2)	Dukes' stage of disease at diagnosis (based on pathological evidence but upgraded to Dukes D if clinical evidence of metastasis) Dukes D should be recorded if metastatic spread is identified either in the preoperative staging process, e.g. on CT scanning, MRI, USS, chest x-ray or at the time of operation. It is accepted that a small number of D cases are cured by further treatment such as liver resection, but for COSD metastatic spread distant from the primary should always be recorded as D.	0..1	modifiedDukes		MODIFIED DUKES (CO5170 from Cancer Outcomes and Services Dataset)
			A	Dukes A Tumour confined to wall of bowel, nodes negative	
			B	Dukes B Tumour penetrates through the muscularis propria to involve extramural tissues, nodes negative	
			C1	Dukes C1 Metastases confined to regional lymph nodes (node/s positive but apical node negative)	
			C2	Dukes C2 Metastases present in nodes at mesenteric artery ligature (apical node positive)	
			D	Dukes D Metastatic spread outside the operative field	
			99	Not Known	

6.4.6.1 AJCC Stage [0..1]

A maximum of one report containing AJCC Stage can be submitted together with each Cancer Specific Staging report. The Cancer Specific Staging report can be submitted without this information.

AJCC STAGE GROUP and version.

Name	Description	Multiplicity	Data Type	Related To
AJCC Stage Group (38871@3.1.2)	AJCC STAGE GROUP, not the UICC TNM Stage Grouping, should be collected for stageable skin cancers. American Joint Committee on Cancer staging of tumour at diagnosis. This is the final integrated stage as agreed by MDT. See COSD User Guide for site specific options.	1..1	ajccStageGroup >10 enumerations, please click link above to view full list.	AJCC STAGE GROUP (SK12510 from Cancer Outcomes and Services Dataset)
AJCC Stage Group Version (38872@3.1.2)	AJCC Stage Group Version	1..1	xs:string	

6.4.6.2 Final Figo Stage [0..1]

A maximum of one report containing Final Figo Stage can be submitted together with each Cancer Specific Staging report. The Cancer Specific Staging report can be submitted without this information.

Final Figo Stage and Version

Name	Description	Multiplicity	Data Type	Related To
Final Figo Stage (33029@3.1.2)	The FIGO stage is generally confirmed at pathology review in MDT meetings following surgery for uterine and vulval malignancies and for ovarian malignancies undergoing primary surgery. For ovarian malignancies planned to undergo neoadjuvant chemotherapy and for cases of cervical cancer (which is staged clinically), the final FIGO stage is determined at the time of review of clinical findings, imaging, cytology and biopsy histology at the MDT meeting.	1..1	finalFigoStage >10 enumerations, please click link above to view full list.	
Final Figo Stage Version	Version of final figo used for staging	1..1	xs:string	

(33088@3.1.2)				

6.4.6.3 Staging (Upper GI) [0..1]

A maximum of one report containing Staging (Upper GI) can be submitted together with each Cancer Specific Staging report. The Cancer Specific Staging report can be submitted without this information.

Priority COSD data items from UPPER GI - STAGING - LIVER HCC and PANCREAS.

Name	Description	Multiplicity	Data Type	Related To
Barcelona clinic liver cancer (bclc) stage (38904@3.1.2)	The Barcelona Clinic Liver Cancer (BCLC) stage includes both anatomic and non-anatomic factors and is widely used within the UK to predict prognosis and determine treatment.	0..1	barcelonaClinicLiverCancerBclcStage **0** Very early **A** Early **B** Intermediate **C** Advanced **D** Terminal	BARCELONA CLINIC LIVER CANCER (BCLC) STAGE (UG14520 from Cancer Outcomes and Services Dataset)
Child-pugh score (38905@3.1.2)	Record the overall Child-Pugh score. This is the level of disease of the liver.	0..1	childPughScore **A** Child-Pugh A **B** Child-Pugh B **C** Child-Pugh C	CHILD-PUGH SCORE (UG14530 from Cancer Outcomes and Services Dataset)
Portal invasion (38907@3.1.2)	Record whether there is involvement of the portal vein. (From UPPER GI - STAGING - LIVER HCC)	0..1	portalInvasion **Y** Present **N** Not present **9** Not known	PORTAL INVASION (UG14550 from Cancer Outcomes and Services Dataset)
Number of lesions (radiological) (38931@3.1.2)	Radiologically determined number of lesions. (From UPPER GI - STAGING - LIVER HCC)	0..1	numberOfLesionsRadiological	NUMBER OF LESIONS (RADIOLOGICAL) (UG14540 from Cancer Outcomes and Services Dataset)
Clinical stage (pancreatic cancer) (38908@3.1.2)	COSD UG14560, UPPER GI - STAGING - PANCREAS. Description: 'Clinically agreed stage based on	0..1	clinicalStagePancreaticCancer **10** Localised and	CLINICAL STAGE (PANCREATIC CANCER) (UG14560 from Cancer Outcomes and Services

	radiological findings of tumour extent in order to offer treatment recommendations. The category selected depends on tumour location within the pancreas and the arterial or venous involvement.			resectable	Dataset)
			20	Borderline resectable	
			30	Unresectable (locally advanced or metastatic)	
			31	Unresectable (locally advanced)	
			32	Unresectable (metastatic)	
Trans arterial chemoembolisation (38910@3.1.2)	Was Trans Arterial Chemoembolisation (TACE) carried out?	0..1	yesNoNk		TRANS ARTERIAL CHEMOEMBOLISATION (UG13580 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			9	Not known	

6.4.6.4 Staging (Urology - Testicular) [0..1]

A maximum of one report containing Staging (Urology - Testicular) can be submitted together with each Cancer Specific Staging report. The Cancer Specific Staging report can be submitted without this information.

Priority COSD data items from UROLOGY - STAGING - TESTICULAR. To carry staging details for Urology (Testicular).

Name	Description	Multiplicity	Data Type		Related To
Stage grouping (testicular) (38912@3.1.2)	TESTICULAR ONLY. Nationally agreed anatomical stage groupings as defined by The Royal Marsden Hospital (RMH).	0..1	stageGroupingTesticular >10 enumerations, please click link above to view full list.		STAGE GROUPING (TESTICULAR) (UR15300 from Cancer Outcomes and Services Dataset)
Extranodal metastases (38913@3.1.2)	For testicular Stage 4 patients only Indicate the extent of metastatic spread (multiple items can be selected)	0..1	extranodalMetastases		EXTRANODAL METASTASES (UR15320 from Cancer Outcomes and Services Dataset)
			H	Liver involvement	
			B	Brain involvement	
			M	Mediastinal involvement	

			N	Neck nodes	
			L	Lung involvement	
Lung metastases sub-stage grouping (38914@3.1.2)	For testicular cancer only Where lung metastases are identified, specify the RMH grouping.	0..1	lungMetastasesSubStageGrouping		LUNG METASTASES SUB-STAGE GROUPING (UR15330 from Cancer Outcomes and Services Dataset)
			L1	less than or equal to 3 metastases	
			L2	Greater than 3 metastases	
			L3	Greater than 3 metastases, one or more greater than or equal to 2cm diameter	

6.4.7 Cancer Specific Diagnosis [0..1]

A maximum of one report containing Cancer Specific Diagnosis can be submitted together with each Diagnosis report. The Diagnosis report can be submitted without this information.

New section. COSD Cancer specific diagnosis data can be submitted as a separate Event or with Diagnosis Event. Submission is optional however, if submitted, one of the following sections must be provided

6.4.7.1 Diagnosis (Colorectal) [1..1]

One report containing Diagnosis (Colorectal) must be submitted together with each Cancer Specific Diagnosis report.

Priority COSD data items from COLORECTAL - DIAGNOSIS. To carry diagnosis details for colorectal cancer.

Name	Description	Multiplicity	Data Type		Related To
Synchronous Tumour Indicator (42056@3.1.2)	Related to COSD v7, CO5400. Record any synchronous tumours in the Colon as identified by the clinician at presentation. Synchronous tumours are defined as discrete	1..*	synchronousTumourIndicator		
			1	CAECUM	
			2	APPENDIX	

	tumours apparently not in continuity with other primary cancers originating in the same site or tissue.		**3**	ASCENDING COLON	
			4	HEPATIC FLEXURE	
			5	TRANSVERSE COLON	
			6	SPLENIC FLEXURE	
			7	DESCENDING COLON	
			8	SIGMOID COLON	
			9	RECTOSIGMOID	
			10	RECTUM	

Schema 6.5 Cancer Care Plan

This section includes details applicable to care planning, which are normally discussed at the MDT meeting. Multiple cancer care plans can be submitted.

The Cancer Care Plan Event Date will be the date that the patient agrees to a Care Plan following the recommendation by the MDT.

Name	Description	Multiplicity	Data Type	Related To
Start Date (14961@1.0.1)	Start date for the proposed treatment. This may or may not be known at the time of care planning, and therefore is optional.	0..1	xs:date	
Treatment Intent (14962@3.1.2)	Intent of the proposed treatment. The intention of a Cancer Care Plan developed within a Cancer Care Spell.	1..1	cancerCarePlanIntent **C** Curative **Z** Non Curative **X** No active treatment **9** Not Known	CANCER CARE PLAN INTENT (CR0460 from Cancer Outcomes and Services Dataset)
No Cancer Treatment Reason (14965@3.1.2)	Code for decision not to treat. The main reason why no active cancer treatment is specified within a Cancer Care Plan.	0..1	noCancerTreatmentReason **01** Patient declined treatment **02** Unfit: poor performance status **03** Unfit: significant co-morbidity **04** Unfit: advanced stage cancer	NO CANCER TREATMENT REASON (CR0490 from Cancer Outcomes and Services Dataset)

			05 Unknown primary site **06** Died before treatment **07** No active treatment available **08** Other **10** Monitoring only **99** Not Known	
Performance Status (14963@3.1.2)	Performance status of the participant. A World Health Organisation classification indicating a PERSON's status relating to activity / disability.	1..1	performanceStatusAdult **0** Able to carry out all normal activity without restriction **1** Restricted in physically strenuous activity, but able to walk and do light work **2** Able to walk and capable of all self care, but unable to carry out any work. Up and about more than 50% of	PERFORMANCE STATUS (ADULT) (CR0510 from Cancer Outcomes and Services Dataset)

			waking hours **3** Capable of only limited self care, confined to bed or chair more than 50% of waking hours **4** Completely disabled. Cannot carry on any self care. Totally confined to bed or chair **9** Not recorded	
Outcome of MDT (33513@3.1.2)	Freetext report on the outcome of MDT discussions	0..1	xs:string	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_"	1..*	tumourID	

	proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC			

6.5.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Cancer Care Plan report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.5.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.5.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Cancer Care Plan report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.5.3 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Cancer Care Plan report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.5.4 Cancer Specific Care Plan [0..1]

A maximum of one report containing Cancer Specific Care Plan can be submitted together with each Cancer Care Plan report. The Cancer Care Plan report can be submitted without this information.

Priority COSD data items from cancer specific care plans. Submission is optional however, if submitted, one of the following sections must be provided.

If submitted, one of the following must be submitted together with each Cancer Specific Care Plan report.

6.5.4.1 Cancer Care Plan (Urology) [1..1]

Priority COSD data items from UROLOGY - CANCER CARE PLAN. To carry the cancer care plan details for Urology cancer.

Name	Description	Multiplicity	Data Type	Related To
Normal LDH (38911@3.1.2)	TESTICULAR ONLY. This is the upper limit of normal for the LDH (Lactate Dehydrogenase Level) assay which is used to calculate S Category.	1..1	normalLdh	NORMAL LDH (UR15020 from Cancer Outcomes and Services Dataset)

or in the case of,

6.5.4.2 Cancer Care Plan (CNS) [1..1]

Priority COSD data items from CNS - CANCER CARE PLAN. To carry cancer care plan details for CNS cancer.

Name	Description	Multiplicity	Data Type	Related To
MDT provisional diagnosis (ICD) (38936@3.1.2)	Working diagnosis as defined at MDT where the first definitive treatment is agreed. This is the clinical opinion which may also be informed by biopsy, radiological and/or other investigations.	0..1	mdtProvisionalDiagnosisIcd	MDT PROVISIONAL DIAGNOSIS (ICD) (BA3080 from Cancer Outcomes and Services Dataset)
Primary diagnosis (ICD radiological) (38935@3.1.2)	The preliminary primary diagnosis based on radiological examination recorded pre treatment. In many cases this will be the definitive clinical diagnosis, but needs to be distinguished from the subsequent pathological diagnosis - if it becomes available.	0..1	primaryDiagnosisIcdRadiological	PRIMARY DIAGNOSIS (ICD RADIOLOGICAL) (BA3060 from Cancer Outcomes and Services Dataset)

or in the case of,

6.5.4.3 Cancer Care Plan (Lung) [1..1]

Priority COSD data items from LUNG - CANCER CARE PLAN. To carry care plan details for Lung Carcinoma.

Name	Description	Multiplicity	Data Type		Related To
Mediastinal sampling indicator (38899@3.1.2)	Record if the patient had a mediastinoscopy, mediastinotomy, open mediastinal sampling or other type of mediastinal biopsy (e.g. Endobronchial ultrasound or transbronchial needle aspiration biopsy)'	1..1	yesNoNk		MEDIASTINAL SAMPLING INDICATOR (LU10060 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			9	Not known	

Schema 6.6 Intervention

Interventions are treatment events. For each intervention there is a set of core essential data items.

In addition to core essential data items for each intervention, there are specific data items for Surgery, Systemic Anti-Cancer Therapy, Radiotherapy, and Other Treatment, as applicable. Multiple interventions can be submitted.

6.6.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Intervention report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.6.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.6.2 Surgery And Other Procedures [0..*]

Multiple reports containing surgery and other procedures can be submitted together with each interventions report.

For surgical and other procedures, in addition to the core intervention data items, the following are required:

Name	Description	Multiplicity	Data Type	Related To
Main Specialty Code (29100@3.1.2)	The main specialty code of the consultant performing the intervention	0..1	specialty	
Primary Procedure (12806@3.1.2)	OPCS code for the primary procedure. Primary procedure is the main procedure carried out.	1..1	opcsProcedureCodes	PRIMARY PROCEDURE (OPCS) (CR0720 from Cancer Outcomes and Services Dataset)

Other Procedures (14968@3.1.2)	OPCS codes for other procedures. This is a procedure other than the PRIMARY PROCEDURE (OPCS), carried out and recorded for CDS or Hospital Episode Statistics purposes. (This may occur more than once).	0..*	opcsProcedureCodes	PROCEDURE (OPCS) (CR0730 from Cancer Outcomes and Services Dataset)
ASA score (38937@3.1.2)	Related to COSD v7, CR6010. The ASA physical status classification system is a system for assessing the fitness of patients before surgery.	0..1	asaScore	
			1 A normal healthy patient.	
			2 A patient with mild systemic disease.	
			3 A patient with severe systemic disease	
			4 A patient with severe systemic disease that is a constant threat to life.	
			5 A moribund patient who is not expected to survive without the operation.	

			6 A declared brain-dead patient whose organs are being removed for donor purposes	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.6.2.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Surgery And Other Procedures report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values

previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.6.2.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Surgery And Other Procedures report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.6.2.3 Cancer Specific Surgery [0..1]

A maximum of one report containing Cancer Specific Surgery can be submitted together with each Surgery And Other Procedures report. The Surgery And Other Procedures report can be submitted without this information.

Priority COSD data items from dance specific surgery. Submission is optional however, if submitted, one of the following sections must be provided

6.6.2.3.1 Surgery (CNS) [1..1]

One report containing Surgery (CNS) must be submitted together with each Cancer Specific Surgery report.

Priority COSD data items from CNS - SURGERY & OTHER PROCEDURES. To carry surgery and other procedure details for CNS cancer.

Name	Description	Multiplicity	Data Type	Related To
Tumour location (surgical) (38938@3.1.2)	Surgically determined anatomical location of lesion(s) or where centered.	0..1	locationSurgical >10 enumerations, please click link above to view full list.	TUMOUR LOCATION (SURGICAL) (BA3100 from Cancer Outcomes and Services Dataset)
Excision or Procedure Type (38939@3.1.2)	Identify type of excision or procedure (where performed)	0..1	excisionOrProcedureType **1** Limited (<50%) **2** Partial (50-69%) **3** Subtotal (70-95%) **4** Total Macroscopic **5** Extent Uncertain **6** CSF Division Procedure **9** Not Known	

6.6.3 Systemic Anti-Cancer Therapy [0..*]

Multiple reports containing systemic anti-cancer therapy can be submitted together with each Interventions report.

For each course of therapy, in addition to the core intervention data items, the following are required:

Name	Description	Multiplicity	Data Type	Related To
Drug Treatment Intent (14970@3.1.2)	Treatment intent.	1..1	drugTreatmentIntent **D** Disease Modification **P** Palliative **A** Adjuvant **C** Curative **N** Neoadjuvant	DRUG TREATMENT INTENT (CR1070 from Cancer Outcomes and Services Dataset)
Drug Regimen (14971@3.1.2)	The drug regimen prescribed. To be consistent with the National Regimen List.	1..1	drugRegimenAcronym	DRUG REGIMEN ACRONYM (CR1080 from Cancer Outcomes and Services Dataset)
Main Specialty Code (29100@3.1.2)	The main specialty code of the consultant performing the intervention	0..1	specialty	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric	1..*	tumourID	

	characters i.e. RN3_A098BC			

6.6.3.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Systemic Anti-Cancer Therapy report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.6.3.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Systemic Anti-Cancer Therapy report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis	Optionally, provide the related cancer diagnosis as SNOMED CT	0..*	diagnosisCode(snomedCt)	

(SNOMEDCT) (35539@3.1.2)	code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.			

6.6.4 Radiotherapy [0..*]

Multiple reports containing radiotherapy can be submitted together with each interventions report.

For each course of radiotherapy, in addition to the core intervention data items, the following are required:

Name	Description	Multiplicity	Data Type	Related To
Radiotherapy Site (14973@3.1.2)	OPCS code for site. A classification of part of the body to which the RADIOTHERAPY ACTUAL DOSE is administered.	1..1	anatomicalTreatmentSiteRadiotherapy	ANATOMICAL TREATMENT SITE (RADIOTHERAPY) (CR1140 from Cancer Outcomes and Services Dataset)
Dose (14974@3.1.2)	The total prescribed absorbed radiation dose, measured in Grays, given to the ICRU Reference Point for the whole prescription. http://www.icru.org/home/reports/prescribing-recording-and-reporting-photon-beam-therapy-report-62	1..1	radiotherapyTotalDose	RADIOTHERAPY TOTAL DOSE (CR2080 from Cancer Outcomes and Services Dataset)
Prescription (14976@3.1.2)	Reference to uploaded copy of prescription document	0..1	xs:string	
Plan (14975@3.1.2)	Reference to uploaded copy of radiotherapy plan	0..1	xs:string	
Radiotherapy Intent (29103@3.1.2)	Intent of radiotherapy	0..1	radiotherapyIntent **01** Palliative	RADIOTHERAPY INTENT (CR1570 from Cancer Outcomes and Services Dataset)

			02 Anti-cancer **03** Other	
Main Specialty Code (29100@3.1.2)	The main specialty code of the consultant performing the intervention	0..1	specialty	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.6.4.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Radiotherapy report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.6.4.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Radiotherapy report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.6.4.3 Radiotherapy Details [1..1]

One report containing Radiotherapy Details must be submitted together with each Radiotherapy report.

One of the following must be submitted together with each Radiotherapy Details report

6.6.4.3.1 Brachytherapy [1..1]

Name	Description	Multiplicity	Data Type	Related To
Brachytherapy Type (14977@3.1.2)	The type of Brachytherapy Treatment Course being used, if applicable.	1..1	brachytherapyType **BI** Interstitial **BC** Intra-cavity	BRACHYTHERAPY TYPE (CR1200 from Cancer Outcomes and Services Dataset)

			BT Not otherwise specified **US** Unsealed Source	

or in the case of,

6.6.4.3.2 External Beam [1..1]

Name	Description	Multiplicity	Data Type	Related To
External Beam Type (14978@3.1.2)	Type of external beam, if applicable. The prescribed type of beam for a Teletherapy Treatment/ Exposure	1..1	externalBeamType **imrt** IMRT **stereotactic** Stereotactic **2dxrt** 2DXRT **3dxrt** 3DXRT **4dxrt** 4DXRT **electrons** Electrons **protons** Protons	
Fractions (14981@3.1.2)	Dose fractions, if external beam therapy. The total number of Fractions or hyperfraction delivered as part of a RADIOTHERAPY PRESCRIPTION.	1..1	radiotherapyTotalFractions	RADIOTHERAPY TOTAL FRACTIONS (CR2090 from Cancer Outcomes and Services Dataset)

6.6.5 Cancer Specific Treatments [0..*]

Multiple reports containing cancer specific treatments can be submitted together with each Interventions report.

Name	Description	Multiplicity	Data Type	Related To

Main Specialty Code (29100@3.1.2)	The main specialty code of the consultant performing the intervention	0..1	specialty	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.6.5.1 Event Details [1..1]

One report containing Event Details must be submitted together with each Cancer Specific Treatments report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	

Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.6.5.2 Related Cancer Diagnoses [1..1]

One report containing Related Cancer Diagnoses must be submitted together with each Cancer Specific Treatments report.

Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.

Name	Description	Multiplicity	Data Type	Related To
Related Cancer Diagnosis (ICD) (14892@1.0.1)	Cancer diagnoses that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	1..*	primaryDiagnosisIcd	
Related Cancer Diagnosis (SNOMEDCT) (35539@3.1.2)	Optionally, provide the related cancer diagnosis as SNOMED CT code as well as the ICD code. Related Cancer Diagnosis is the diagnosis that led to the reported clinical event. More than one diagnosis can be provided for the same event, e.g. where an event pertains to more than one diagnosis.	0..*	diagnosisCode(snomedCt)	

6.6.5.3 Cancer Specific Treatment [1..1]

One report containing Cancer Specific Treatment must be submitted together with each Cancer Specific Treatments report.

One *of the following must be submitted together with each Cancer Specific Treatment report*

6.6.5.3.1 Other Treatment (Bladder) [1..1]

Priority COSD Data items from UROLOGY - TREATMENT - BLADDER. To carry the cancer treatment details for Bladder.

Name	Description	Multiplicity	Data Type		Related To
Intravesical chemotherapy received indicator (38915@3.1.2)	Related to UR15100, updated description in COSD v7.0: BLADDER ONLY (Only required for patients having chemotherapy) Record as YES for patients having intravesical chemotherapy to distinguish from intravenous	0..1	yesNoNk		INTRAVESICAL CHEMOTHERAPY RECEIVED INDICATOR (UR15100 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			9	Not known	
Intravesical Immunotherapy Received Indicator (39085@3.1.2)	Related to UR15110, updated description in COSD v7.0: BLADDER ONLY (Only required for patients having Immunotherapy) Record as YES for patients having intravesical Immunotherapy to distinguish from intravenous	0..1	yesNoNk		INTRAVESICAL IMMUNOTHERAPY RECEIVED INDICATOR (UR15110 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			9	Not known	

or in the case of,

6.6.5.3.2 Other Treatment (Upper GI) [1..1]

Priority COSD data items from UPPER GI - LIVER METS and LIVER HCC. To carry other procedure details for LIVER METS and Liver HCC.

Name	Description	Multiplicity	Data Type		Related To
Ablative therapy type (38909@3.1.2)	Describe type of ablative (i.e. locally destructive treatment) therapy used if any	0..1	ablativeTherapyType		ABLATIVE THERAPY TYPE (UG13560 from Cancer Outcomes and Services Dataset)
			N	None	
			R	Radiofrequency ablation	
			O	Other ablative treatment	
			9	Not known	

Trans arterial chemoembolisation (38910@3.1.2)	Was Trans Arterial Chemoembolisation (TACE) carried out?	0..1	yesNoNk		TRANS ARTERIAL CHEMOEMBOLISATION (UG13580 from Cancer Outcomes and Services Dataset)
			Y	Yes	
			N	No	
			9	Not known	

Schema 6.8 Death

This section is used to submit details of the date and cause of death. Participants remain in the programme after their death but this information is crucial both for project implementation (to ensure appropriate future contact with the family) and for research.

All fetal participants should also have a death details form completed at the time of recruitment; this is because samples from ongoing pregnancies are not eligible for the programme, as genome sequencing is not yet fast enough to be used in the context of ongoing pregnancy. The date of death should refer to the date the intra-uterine death was discovered, or feticide was carried out.
A report of death should include (see NHS data dictionary):
http://www.datadictionary.nhs.uk/data_dictionary/classes/p/person_death_details_at.asp?shownav=1

Name	Description	Multiplicity	Data Type	Related To
Death Location (12777@1.0.1)	Location of death	0..1	deathLocation **3** Voluntary hospice / Specialist Palliative Care unit **2** NHS hospice / Specialist Palliative Care unit **1** Hospital **6** Other **5** Care Home **4** PATIENT's own home	
Significant (12781@1.0.1)	Significant condition not leading to death. Coded according to the International Classification of Diseases (ICD) code of the condition leading to death as	0..*	deathCauseCode	

	recorded on the death certificate.			
Immediate Cause (12778@1.0.1)	Immediate cause of death. Coded according to the International Classification of Diseases (ICD) code of the condition leading to death as recorded on the death certificate.	1..*	deathCauseCode	
Condition (12780@1.0.1)	Condition leading to death. Coded according to the International Classification of Diseases (ICD) code of the condition leading to death as recorded on the death certificate.	0..*	deathCauseCode	
Underlying Cause (12779@1.0.1)	Underlying cause of death. Coded according to the International Classification of Diseases (ICD) code of the condition leading to death as recorded on the death certificate.	0..*	deathCauseCode	

6.8.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Death report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)

Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.8.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.8.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Death report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

Schema 6.9 Consent Update

This section reports changes in a cancer participant's consent status after they have joined the project, including changes in preference relating to the participant's options regarding additional findings (Consent Details (29742.1)).
If a patient withdraws from the project, please complete a Withdrawal form, not a Consent update form.

Name	Description	Multiplicity	Data Type	Related To
Name and Version of Consent Form Update (34558@3.1.2)	Name and Version of cancer update form used. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	1..1	xs:string	
Consent Given (12545@1.0.1)	Yes no answer to consent given	1..1	yesNo **yes** Yes **no** No	
Consent Form (12546@1.0.1)	File name of uploaded PDF copy of consent form - requested format [ParticipantId]_consent_[TimeStamp].pdf	0..1	xs:string	
Person Taking Consent (12547@1.0.1)	The full name of the person taking consent	1..1	xs:string	
Name and Version of Participant Information Sheet Update (34564@3.1.2)	Name and Version of information sheet presented. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	1..1	xs:string	
Name and Version of Assent Form (34552@3.1.2)	Name and Version of Cancer Assent form used. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	0..1	xs:string	
Assent Form (34543@1.0.1)	File name of the uploaded PDF copy of the assent form. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials	0..1	xs:string	

	and enumerations.			
Additional optional consent materials (34556@1.0.1)	Names and versions of additional cancer consent materials used. Please see appendix F for latest list of consent forms, participant information sheets, additional optional consent materials and enumerations.	0..1	xs:string	

6.9.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Consent Update report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.9.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.9.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Consent Update report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

6.9.3 Consent Details [1..1]

One report containing Consent Details must be submitted together with each Consent Update report.

Details corresponding to the questions and responses on the consent form.

Name	Description	Multiplicity	Data Type	Related To
Health Related Additional Findings (34544@1.0.1)	Health-related additional findings: Does the participant want these looked for and fed back to their clinical team?	1..1	yesNo **yes** Yes **no** No	
Reproductive Additional Findings (34546@1.0.1)	Reproductive additional findings: Does the participant want these looked for and fed back to their clinical team?	0..1	yesNoNotRelevant **yes** yes **no** no **not_relevant** not relevant	

Schema 6.10 Reason Sample Not Sent

This section reports if sample has been taken from a participant but not sent.

Name	Description	Multiplicity	Data Type		Related To
Reason Sample Not Sent (33116@3.1.2)	Reason sample not sent from GMC to Biorepository, if applicable. Note the only instances a blood sample shouldn't be sent, is if the tumour sample was not sent, or if there has been a successfully sequenced germline previously. As they have to send the both samples together.	1..1	tumourSampleNotSentReason		
			tumour_sample_not_taken	Tumour sample not taken	
			tumour_type_not_eligible	Tumour type not eligible	
			poorly_cellular_tumour	Poorly cellular tumour (Less than 40 percent neoplastic cells)	
			insufficient_tumour_post_neoadjuvant_chemotherapy	Insufficient tumour post neoadjuvant chemotherapy	
			insufficient_dna	Insufficient DNA	
			no_cancer_diagnosed	No Cancer Diagnosed	
			ffpe_not_optimally_fixed	FFPE not optimally fixed	
			ffpe_not_optimally_processed	FFPE not optimally processed	
			high_necrosis	High necrosis (over 20	

				percent)	
			other	Other	

6.10.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Reason Sample Not Sent report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.10.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number (14821@1.0.1)	The COMMUNITY HEALTH INDEX NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	chiNumber	
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.10.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Reason Sample Not Sent report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

Schema 6.11 Presentation

Presentation details associated with the Genomics England tumour sample.

Name	Description	Multiplicity	Data Type	Related To
Presentation (12838@3.1.2)	Symptoms presented, recorded against supplied enumerations (for example, breast cancer presentation may be: breast mass, altered breast appearance, axillary mass, other mass, nipple discharge, or screening). The list of disease types will be validated against the types contained in Appendix G. These may be subject to change and GMCs are requested to ensure that data capture systems are flexible enough to accommodate changes to the list of disease contained in Appendix G	1..*	xs:string	
Tumour ID (42230@3.1.2)	A locally allocated identifier for the participant's tumour. This should be unique for each tumour submitted from a patient. Two tumours resected at the same time would have unique Tumour IDs. All sample reports and event reports that relate to a Genomics England tumour sample must have a locally allocated Tumour ID. Tumour IDs must be unique within the context of a GMC Clinic and should conform to the following convention: Clinic ID proceeded by "_" proceeded by the local tumour identifier used to refer to a tumour, which must be between 1 and 16 alphanumeric characters i.e. RN3_A098BC	1..*	tumourID	

6.11.1 Participant Identifiers [1..1]

One report containing Participant Identifiers must be submitted together with each Presentation report.

The following information is used to identify the participant and must be included with all data submissions.

Name	Description	Multiplicity	Data Type	Related To
Participant ID (12502@1.0.1)	Participant Identifier (supplied by Genomics England)	1..1	participantId	
Date of Birth (12505@1.0.1)	The date on which a PERSON was born or is officially deemed to have been born.	1..1	xs:date	PERSON BIRTH DATE (NHS Data Dictionary GEL Subset)
Surname (12507@1.0.1)	The participant's surname	1..1	personFamilyName	PERSON FAMILY NAME (CR0050 from Cancer Outcomes and Services Dataset)
Forenames (12508@1.0.1)	The participant's forenames	1..1	personGivenName	PERSON GIVEN NAME (CR0060 from Cancer Outcomes and Services Dataset)

6.11.1.1 Person Identifier [1..1]

One report containing Person Identifier must be submitted together with each Participant Identifiers report.

Choice of one of either NHS Number (Wales & England) OR CHI Number (Scotland) OR Health and Care Number (Northern Ireland).

One of the following must be submitted together with each Person Identifier report.

Name	Description	Multiplicity	Data Type	Related To
NHS Number (12506@1.0.1)	Validated NHS number for participant	1..1	nhsNumber	
Or in the case of,				
CHI Number	The COMMUNITY HEALTH INDEX	1..1	chiNumber	

(14821@1.0.1)	NUMBER (CHI NUMBER) uniquely identifies a PATIENT on the Community Health Index (Scotland) within the NHS in Scotland. It is the equivalent of the NHS NUMBER in England and Wales.			
Or in the case of,				
Health and Care Number (42126@1.0.1)	Validated HEALTH AND CARE NUMBER (H&C NUMBER). Uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.	1..1	healthAndCareNumber	

6.11.2 Event Details [1..1]

One report containing Event Details must be submitted together with each Presentation report.

The local 'Event Reference' combined with the Event Date, is used to identify the actual participant event involved. GMCs must ensure that local event references are not re-used to refer to different events, therefore two different events on the same date can be easily identified. GMCs must also ensure no two participants are given the same event reference.

If a second submission is received with the same local event reference, then Genomics England will assume that this is an update pertaining to the same participant event, and will be used in place of the original submission; any data values previously supplied that are still applicable must therefore be included.

Name	Description	Multiplicity	Data Type	Related To
Event Date (12727@1.0.1)	Date of the clinical event or observation being reported e.g. date biopsy was taken	1..1	xs:dateTime	
Event Reference (14858@1.0.1)	Unique identifier for local record of clinical event or observation	1..1	xs:string	

7 Data Types

Age

(Cancer Model)

Age in years

Unit of Measure	Year
Regular Expression	\d{1,3}
Rule based on xs:nonNegativeInteger (XMLSchema)	minInclusive(0)
Rule based on xs:integer (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseInteger(string(x)) in BigInteger

Usages

Risk Factors for Breast Cancer

Risk Factors for Endometrial Cancer

Risk Factors for Malignant Melanoma

Risk Factors for Ovarian Cancer

Risk Factors for Testicular Cancer

Postcode

(PDS)

The UK format Postcode, 8 character string, as per BS7666. The 8 characters field allows a space to be inserted to differentiate between the inward and outward segments of the code, enabling full use to be made of Royal Mail postcode functionality.
N.B. Must be capitalized

Regular Expression	^(GIR ?0AA|[A-PR-UWYZ]([0-9]{1,2}|([A-HK-Y][0-9]([0-9ABEHMNPRV-Y])?)|[0-9][A-HJKPS-UW])?[0-9][ABD-HJLNP-UW-Z]{2})$

Usages

Participant Contact Details

SCATEGORY

(Cancer Outcomes and Services Dataset)

TESTICULAR ONLY. Based on serum tumour markers AFP, HCG and LDH. For Testicular Cancer S category is an additional prognostic factor.

Rule string(2)

Code	Description
SX	Marker studies not available or not performed
S0	Normal
S1	LDH (UNITS/LITRE) - Less than 1.5 x normal, HCG (MILLIUNITS/MILLILITRE) - Less than 5,000, AFP (NANOGRAMS/MILLILITRE) - Less than 1,000
S2	LDH (UNITS/LITRE) - 1.5-10 x normal, HCG (MILLIUNITS/MILLILITRE) - 5,000-50,000, AFP (NANOGRAMS/MILLILITRE) - 1,000-10,000
S3	LDH (UNITS/LITRE) - Greater than 10 x normal, HCG (MILLIUNITS/MILLILITRE) - Greater than 50,000, AFP (NANOGRAMS/MILLILITRE) - Greater than 10,000

Usages

Cancer Specific Grading

Skintype

(Cancer Model)

Skin type

Code	Description
i	I Always burns, never tans
ii	II Usually burns, tans minimally
iii	III Sometimes mild burn, tans uniformly
iv	IV Burns minimally, always tans well
v	V Very rarely burns, tans very easily
vi	VI Never burns, never tans
unknown	unknown

Usages

Risk Factors for Malignant Melanoma

ablativeTherapyType

(Cancer Outcomes and Services Dataset)

Describe type of ablative (i.e. locally destructive treatment) therapy used if any

Code	Description
N	None
R	Radiofrequency ablation
O	Other ablative treatment
9	Not known

Usages

Other Treatment (Upper GI)

addressLine

(PDS)

Includes main, temporary and correspondence addresses
5 lines excludes postcode, may be vernacular or PAF derived. The following address lines should normally be present although there may be some exceptions:
-1 or 2,
-and 4

Regular Expression .{2,175}

Usages

Participant Contact Details

ajccStageGroup

(Cancer Model)

AJCC STAGE GROUP, not the UICC TNM Stage Grouping, should be collected for stageable skin cancers. American Joint Committee on Cancer staging of tumour at diagnosis. This is the final integrated stage as agreed by MDT. See COSD User Guide for site specific options.

Rule based on ajccStageGroup (Cancer Outcomes and Services Dataset)

x ==~/[a-zA-Z0-9]{1,2}/

Code	Description
1	Stage I
1a	Stage IA
1b	Stage IB
2	Stage II
2a	Stage IIA
2b	Stage IIB
2c	Stage IIC
3	Stage III
3a	Stage IIIA
3b	Stage IIIB
3c	Stage IIIC
4	Stage 4

Usages

AJCC Stage

alkBlastomaMarker

(Cancer Model)

Blastoma marker test result

Code	Description
not_tested	not tested
wild_type	wild-type
mutant	mutant

Usages

Childhood Tumour Markers

allredScore

(Cancer Model)

ALLRED Score

Regular Expression	[0-8]
Rule based on xs:integer (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseInteger(string(x)) in BigInteger

Usages

Pathology (Breast)

anatomicalSideImaging

(Cancer Outcomes and Services Dataset)

*IMAGING CODE (NICIP)
and/or
*IMAGING CODE (SNOMED CT)
and/or
*CANCER IMAGING MODALITY and IMAGING ANATOMICAL SITE and ANATOMICAL SIDE (IMAGING)
is required.

The side of the body that is the subject of an Imaging or Radiodiagnostic Event.

Code	Description
L	Left
R	Right
M	Midline
B	Bilateral
8	Not applicable
9	Not Known

Usages

Imaging

anatomicalTreatmentSiteRadiotherapy

(Cancer Outcomes and Services Dataset)

The OPCS4 anatomical site code of the site subjected to radiotherapy

Rule	x==~/[a-zA-Z0-9.]{3,5}/
Based On	OPCS-4 (Cancer Outcomes and Services Dataset)
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Radiotherapy

asaScore

(Cancer Outcomes and Services Dataset)

The ASA physical status classification system is a system for assessing the fitness of patients before surgery.

Code	Description
1	A normal healthy patient.
2	A patient with mild systemic disease.
3	A patient with severe systemic disease
4	A patient with severe systemic disease that is a constant threat to life.

5	A moribund patient who is not expected to survive without the operation.
6	A declared brain-dead patient whose organs are being removed for donor purposes

Usages

Surgery And Other Procedures

barcelonaClinicLiverCancerBclcStage

(Cancer Outcomes and Services Dataset)

The Barcelona Clinic Liver Cancer (BCLC) stage includes both anatomic and non-anatomic factors and is widely used within the UK to predict prognosis and determine treatment.

Code	Description
0	Very early
A	Early
B	Intermediate
C	Advanced
D	Terminal

Usages

Staging (Upper GI)

basisOfDiagnosis(cancer)

This is the method used to confirm the cancer.

Code	Description
2	Clinical Investigation: Includes all diagnostic techniques (e.g. X-rays, endoscopy, imaging, ultrasound,
1	Clinical: Diagnosis made before death but without the benefit of any of the following (2-7)
0	Death Certificate: The only information available is from a death certificate
7	Histology of a primary tumour: Histological examination of tissue from the primary tumour, however obtained, including all cutting and bone marrow biopsies. Also includes autopsy specimens of a primary tumour
6	Histology of a metastasis: Histological examination of tissues from a metastasis, including autopsy specimens
5	Cytology: Examination of cells whether from a primary or secondary site, including fluids aspirated using endoscopes or needles. Also including microscopic examination of peripheral blood films and trephine bone marrow aspirates

4	Specific tumour markers: Includes biochemical and/or immunological markers which are specific for a tumour site
9	Unknown: No information on how the diagnosis has been made (e.g. PAS or HISS record only)

Usages

Diagnosis

biomarkerPresent

(Cancer Model)

Indication of biomarkers presence

Based On MarkerPresent (Cancer Model)

Code	Description
present	present
not_tested	no tested
equivocal	equivocal
absent	absent

Usages

Colorectal Tumour Markers

Ovarian Tumour Markers

blastomaMarker

(Cancer Model)

Code	Description
not_tested	not tested
amplified	amplified
non_amplified	non amplified

Usages

Childhood Tumour Markers

brachytherapyType

(Cancer Outcomes and Services Dataset)

The type of Brachytherapy Treatment Course being used.

Rule string(2)

Code	Description
BI	Interstitial
BC	Intra-cavity
BT	Not otherwise specified

US	Unsealed Source

Usages

Brachytherapy

breastDensity

(Cancer Model)

Breast density at most recent available pre-surgical mammogram: based on percentage of fibroglandular tissue relative to total area on the two view mammogram.

Code	Description
birads_0	additional imaging evaluation and/or comparison to prior mammogram is needed
birads_1	glandular tissue is less than 25%
birads_2	scattered fibroglandular densities (25-50%)
birads_3	heterogeneously dense (50-75%)
birads_4	extremely dense breast (75-100%)

Usages

Risk Factors for Breast Cancer

breastInvasiveTumourGrade

(Cancer Model)

Code	Description
1	1
2	2
3	3
x	Not assessable

Usages

Cancer Specific Grading

cancerCarePlanIntent

(Cancer Outcomes and Services Dataset)

The intention of a Cancer Care Plan developed within a Cancer Care Spell.

Code	Description
C	Curative
Z	Non Curative
X	No active treatment
9	Not Known

Usages

Cancer Care Plan

cancerImagingModality

(Cancer Outcomes and Services Dataset)

*IMAGING CODE (NICIP)
and/or
*IMAGING CODE (SNOMED CT)
and/or
*CANCER IMAGING MODALITY and IMAGING ANATOMICAL SITE and ANATOMICAL SIDE (IMAGING)
is required.

The type of imaging procedure used during an Imaging or Radiodiagnostic Event for a Cancer Care Spell.
NB: PET Scan also includes PET-CT Scan.

Rule x==~/[a-zA-Z0-9]{4}/

Code	Description
C01X	Standard Radiography
C01M	Mammogram
C02X	CT Scan
C02C	Virtual colonoscopy
C03X	MRI Scan
C04X	PET Scan
C05X	Ultrasound Scan
C06X	Nuclear Medicine imaging

C08A	Angiography
C08B	Barium
C08U	Urography (IV and retrograde)
C09X	Intervention radiography.
CXXX	Other

Usages

Imaging

cancerRecurrenceCarePlanIndicator

(Cancer Outcomes and Services Dataset)

An indication of whether a diagnosis of recurrence has been recorded for which a new Cancer Care Plan is required. A new record should be completed for a recurrence.

Rule string(2)

Code	Description
YL	Yes, including local recurrence
YD	Yes, not including local recurrence
NN	No, not recurrence

Usages

Diagnosis

cancerVascularOrLymphaticInvasion

(Cancer Outcomes and Services Dataset)

An indication of the presence or absence of unequivocal tumour in lymphatic and/or vascular spaces.

Rule string(2)

Code	Description
NU	No - vascular/lymphatic invasion not present
YU	Yes - vascular/lymphatic invasion present
YV	Vascular invasion only present
YL	Lymphatic invasion only present
YB	Both lymphatic and vascular invasion present"
UU	Uncertain whether vascular/lymphatic invasion is present or not
XX	Cannot be assessed
99	Not Known

Usages

Sample Pathology

chiNumber

(NHS Data Dictionary GEL Subset)

The Community Health Index (CHI) is a population register, which is used in Scotland for health care purposes. The CHI number uniquely identifies a person on the index.

Regular Expression	[a-zA-Z0-9]{10}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Person Identifier

childPughScore

(Cancer Outcomes and Services Dataset)

Record the overall Child-Pugh score. This is the level of disease of the liver.

Code	Description
A	Child-Pugh A
B	Child-Pugh B
C	Child-Pugh C

Usages

Staging (Upper GI)

childhoodChronicExposure

(Cancer Model)

Number of years spent living in a country with high UV light between 0 and15 years.

Regular Expression	\d|1[0-5]

Usages

Risk Factors for Malignant Melanoma

chromosomalabnormalityNeuroblastoma

(Cancer Model)

chromosomal abnormality (neuroblastoma)

Code	Description
not_tested	not tested
segmental	segmental
non_segmental	non-segmental

Usages

Childhood Tumour Markers

clinicalStagePancreaticCancer

(Cancer Outcomes and Services Dataset)

Clinically agreed stage based on radiological findings of tumour extent in order to offer treatment recommendations. The category selected depends on tumour location within the pancreas and the arterial or venous involvement

Rule string(2)

Code	Description
10	Localised and resectable
20	Borderline resectable
30	Unresectable (locally advanced or metastatic)
31	Unresectable (locally advanced)
32	Unresectable (metastatic)

Usages

Staging (Upper GI)

consentWithdrawalOptions

(Genomics England Shared)

Genomics England Consent Withdrawal Options

Code	Description
full_withdrawal	OPTION 2: FULL WITHDRAWAL: No further use
partial_withdrawal	OPTION 1: PARTIAL WITHDRAWAL: No further contact

Usages

Withdrawal

consultantCode

(Cancer Outcomes and Services Dataset)

The GMC code of the consultant

Rule	x==~/[a-zA-Z0-9]{1,8}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Consultant Details

coreBiopsyBreast

(Cancer Outcomes and Services Dataset)

Needle core biopsy opinion.

Rule	x==~/[a-zA-Z0-9]{1,3}/

Code	Description
B1	Normal
B2	Benign
B3	Uncertain malignant potential
B4	Suspicious

B5a	Malignant (In situ)
B5b	Malignant (Invasive)
B5c	Malignant (Not assessable)

Usages

Pathology (Breast)

coreBiopsyNode

(Cancer Outcomes and Services Dataset)

Needle biopsy opinion on axillary lymph node.

Rule string(2)

Code	Description
B1	Normal
B2	Benign
B3	Uncertain malignant potential
B4	Suspicious
B5	Malignant

Usages

Pathology (Breast)

cytologyNode

(Cancer Outcomes and Services Dataset)

Cytology opinion on axillary lymph node.

Rule string(2)

Code	Description
C1	Inadequate/unsatisfactory specimen
C2	Benign
C3	Uncertain
C4	Suspicious of malignancy
C5	Malignant

Usages

Pathology (Breast)

dcisTumourGrade

(Cancer Model)

Code	Description
I	Low

i	Intermediate
h	High
x	Not assessable

Usages

Cancer Specific Grading

deathCauseCode

(Genomics England Shared)

DEATH CAUSE ICD CODE is the International Classification of Diseases (ICD) code of the condition leading to death as recorded on the death certificate.

Regular Expression based on deathCauseIcdCode (Cancer Outcomes and Services Dataset)	[a-zA-Z0-9.]{3,6}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Death

deathLocation

(Genomics England Shared)

The type of LOCATION at which a PERSON died.

For the purposes of the Community Information Data Set this is either the LOCATION where the PATIENT expressed a preference to die, or where they actually died.

Based On deathLocationType (NHS Data Dictionary GEL Subset)

Code	Description
3	Voluntary hospice / Specialist Palliative Care unit
2	NHS hospice / Specialist Palliative Care unit
1	Hospital
6	Other
5	Care Home
4	PATIENT's own home

Usages

Death

detrusorMusclePresenceIndicator

(Cancer Outcomes and Services Dataset)

BLADDER ONLY
Presence or absence of detrusor muscle in the specimen

Code	Description

1	Present
2	Absent
9	Not known

Usages

Pathology (Bladder)

diagnosisCode(snomedCt)

(Cancer Model)

SNOMED CT CODE

Regular Expression based on snomedCt (Genomics England Shared)	\d{6,18}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Diagnosis

Related Cancer Diagnoses

diameterInMm

(Genomics England Shared)

diameter in mm

Unit of Measure	millimeter (meter*10^-3
Rule	x==~/\d*\.?\d*/
Rule based on xs:decimal (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDecimal(string(x)) in BigDecimal

Usages

Sample Pathology

distanceToMargin

(Cancer Model)

Distance to closest relevant margin (mm). Distance to nearest margin whether invasive or non invasive. (For COSD measurement to the nearest mm is sufficient but may be recorded to nearest tenth of mm)

Unit of Measure	millimeter (meter*10^-3
Regular Expression based on distanceToMargin (Cancer Outcomes and Services Dataset)	\d{1,2}.\d{1}
Rule based on xs:decimal (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDecimal(string(x)) in BigDecimal

Usages

Pathology (Breast)

distanceToSerosa

(Cancer Outcomes and Services Dataset)

Specify the tumour free distance to the serosa in mm

Rule	x==~/[0-9]{1,2}/
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Pathology (Endometrial)

drugRegimenAcronym

(Cancer Outcomes and Services Dataset)

DRUG REGIMEN ACRONYM

Rule	x==~/[a-zA-Z0-9.\/()]{1,35}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Systemic Anti-Cancer Therapy

drugTreatmentIntent

(Cancer Outcomes and Services Dataset)

DRUG TREATMENT INTENT

Code	Description

D	Disease Modification
P	Palliative
A	Adjuvant
C	Curative
N	Neoadjuvant

Usages

Systemic Anti-Cancer Therapy

durationInYears

(Cancer Model)

Number of years

Unit of Measure	Year (A year is the orbital period of the Earth moving in its orbit around the Sun
Regular Expression	\d{1,3}
Rule based on xs:nonNegativeInteger (XMLSchema)	minInclusive(0)
Rule based on xs:integer (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseInteger(string(x)) in BigInteger

Usages

Risk Factors for Breast Cancer

Risk Factors for Endometrial Cancer

Risk Factors for HPB Cancer

Risk Factors for Ovarian Cancer

emailAddress

(Genomics England Shared)

A Valid Email Address i.e. someone@somedomain.com

Rule	x ==~/(?:[a-z0-9!#$%&'*+\/=?^_`{\|}~-]+(?:\.[a-z0-9!#$%&'*+\/=?^_`{|}~-]+)*|"(?:[\x01-\x08\x0b\x0c\x0e-\x1f\x21\x23-\x5b\x5d-\x7f]|\\[\x01-\x09\x0b\x0c\x0e-\x7f])*")@(?:(?:[a-z0-9](?:[a-z0-9-]*[a-z0-9])?\.)+[a-z0-9](?:[a-z0-9-]*[a-z0-9])?|\[(?:(?:25[0-5]|2[0-4][0-9]|[01]?[0-9][0-9]?)\.){3}(?:25[0-5]|2[0-4][0-9]|[01]?[0-9][0-9]?|[a-z0-9-]*[a-z0-9]:(?:[\x01-\x08\x0b\x0c\x0e-\x1f\x21-\x5a\x53-\x7f]|\\[\x01-\x09\x0b\x0c\x0e-\x7f])+)\]/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Participant Contact Details

epidermalGrowthFactorReceptorMutationalStatus

(Cancer Model)

Code	Description
1	*Wild type*
2	*Mutation*
3	Failed analysis

4	Not assessed
5	Wild type/non-sensitising mutation
6	Sensitising/activating mutation

Usages

Lung Tumour Markers

ethnicCategory

(Cancer Outcomes and Services Dataset)

The ethnicity of a PERSON, as specified by the PERSON.. The 16+1 ethnic data categories defined in the 2001 census is the national mandatory standard for the collection and analysis of ethnicity.
(The Office for National Statistics has developed a further breakdown of the group from that given, which may be used locally.)

Rule x==~/[a-zA-Z0-9]{1,2}/

Code	Description
D	Mixed: White and Black Caribbean
E	Mixed: White and Black African
F	Mixed: White and Asian
G	Mixed: Any other mixed background
A	White: British

B	White: Irish
C	White: Any other White background
L	Asian or Asian British: Any other Asian background
M	Black or Black British: Caribbean
N	Black or Black British: African
H	Asian or Asian British: Indian
J	Asian or Asian British: Pakistani
K	Asian or Asian British: Bangladeshi
P	Black or Black British: Any other Black background
S	Other Ethnic Groups: Any other ethnic group
R	Other Ethnic Groups: Chinese
Z	Not stated
99	Not known

Usages

Registration

excisionMargin

(Cancer Outcomes and Services Dataset)

An indication of whether the excision margin was clear of the tumour and if so, by how much.

Where there is more than one measurement, record the closest or closest relevant margin.
Where actual measurements are not taken use options 01, 05 or 06.
Note that not some values are applicable to specific tumour types

Rule string(2)

Code	Description
01	Excision margins are clear (distance from margin not stated)
02	Excision margins are clear (tumour >5mm from the margin)
03	Excision margins are clear (tumour >1mm but less than or equal to 5mm from the margin
04	Tumour is less than or equal to 1mm from excision margin, but does not reach margin
05	Tumour reaches excision margin
06	Uncertain
07	Margin not involved =>1mm
08	Margin not involved <1mm
09	Margin not involved 1-5mm
98	Not applicable

99	Not Known

Usages

Sample Pathology

excisionOrProcedureType

(Cancer Model)

Identify type of excision or procedure (where performed)

Code	Description
1	Limited (<50%)
2	Partial (50-69%)
3	Subtotal (70-95%)
4	Total Macroscopic
5	Extent Uncertain
6	CSF Division Procedure
9	Not Known

Usages

Surgery (CNS)

extentOfPleuralInvasion

(Cancer Outcomes and Services Dataset)

What is the extent of pleural invasion

Code	Description
1	No pleural invasion
2	Visceral pleura only
3	Parietal pleura/chest wall
4	Mediastinal pleura

Usages

Pathology (Lung)

externalBeamType

(Cancer Model)

External Beam Type

Code	Description
imrt	IMRT
stereotactic	Stereotactic
2dxrt	2DXRT

3dxrt	3DXRT
4dxrt	4DXRT
electrons	Electrons
protons	Protons

Usages

External Beam

extranodalMetastases

(Cancer Outcomes and Services Dataset)

For testicular Stage 4 patients only)
Indicate the extent of metastatic spread (multiple items can be selected)

Code	Description
H	Liver involvement
B	Brain involvement
M	Mediastinal involvement
N	Neck nodes
L	Lung involvement

Usages

Staging (Urology - Testicular)

featuresOfLargestLesionRadiological

(Cancer Outcomes and Services Dataset)

Radiologically identified features of the largest lesion such as density, necrosis recorded pre treatment. This may involve selection of more than one value.

Rule string(2)

Code	Description
01	Contrast-enhancement
02	Calcification
03	Mass effect
04	Hydrocephalus
05	Haemorrhage
06	Cystic/multi-cystic
07	Dural tail
08	Brain oedema
09	Cord signal change
10	Cord compression

Usages

Imaging (CNS)

figoGrade

(Cancer Model)

As per RCPath minimum dataset

Code	Description
1	1
2	2
3	3

Usages

Cancer Specific Grading

finalFigoStage

(Cancer Model)

The FIGO stage is generally confirmed at pathology review in MDT meetings following surgery for uterine and vulval malignancies and for ovarian malignancies undergoing primary surgery. For ovarian malignancies planned to undergo neoadjuvant chemotherapy and for cases of cervical cancer (which is staged clinically), the final FIGO stage is determined at the time of review of clinical findings, imaging, cytology and biopsy histology at the MDT meeting.

Code	Description
ia	IA
ib	IB

ic1	IC1
ic2	IC2
ic3	IC3
iia	IIA
iib	IIB
iiia1_i	IIIA1(i)
iiia1_ii	IIIA1(ii)
iiia2	IIIA2
iiib	IIIB
iiic	IIIC
iva	IVA
ivb	IVB
i	I
ii	II
iii	III
iv	IV

Usages

Final Figo Stage

frenchGradingSystem

(Cancer Model)

Code	Description
g1	G1-Well differentiated (Low grade)
g2	G2-Moderately differentiated (Intermediate grade)
g3	G3-Poorly differentiated (High grade)

Usages

Cancer Specific Grading

fuhrmanGradingSystem

(Cancer Model)

Fuhrman grade according to RCP guidance.
Please see: Fuhrman SA, Lasky LC, Limas C. Prognostic significance of morphologic parameters in renal cell carcinoma. Am J Surg Pathol. 1982 Oct. 6(7):655-63.

Code	Description
1	1
2	2

3	3
4	4

Usages

Cancer Specific Grading

geneScope

(Genomics England Shared)

The gene or genes considered

Based On	hgncSymbol (Genomics England Shared)
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Genetic Result

geneticTestLaboratory

(Genomics England Shared)

Was genetic testing performed in a diagnostic or research laboratory?

Based On genetic_test_laboratory (Genomics England Shared)

Code	Description

research_laboratory	Research laboratory
diagnostic_laboratory	Diagnostic laboratory

Usages

Genetic Result

geneticTestMethod

(Genomics England Shared)

Genetic Test Method

Code	Description
sequencing	Sequencing
sequencing_and_targeted_copy_number_analysis	Sequencing and targeted copy number analysis
copy_number_analysis	Copy number analysis
other_snv_analysis	Other SNV analysis e.g. OLA
targeted_copy_number_analysis	Targeted copy number analysis e.g. MLPA/qPCR
acgh	aCGH
snp_array	SNP array
karyotype	Karyotype

fish	FISH
methylation_testing	Methylation testing
microsatellite_analysis	Microsatellite analysis
fanconi_breakage_testing	Fanconi (MMC/DEB) breakage testing
radiation_hypersensitivity	Radiation hypersensitivity (AT)
uv_hypersensitivity	UV hypersensitivity
unscheduled_dna_synthesis	Unscheduled DNA synthesis
single_gene_sequencing	Single Gene Sequencing
gene_panel	Gene Panel
ihc	IHC
translocation	Translocation eg qPCR/sequencing/FISH/IHC
other	Other

Usages

Genetic Result

genomicsEnglandConsentWithdrawalForms

(Genomics England Shared)

List of consent withdrawal forms used by Genomics England

Code	Description
6a	Withdrawal information and form – for adult or child participants (6a)
6b	Consultee declaration of advice regarding adult participant withdrawal information – for consultees (withdrawal) (6b)

Usages

Withdrawal

gleasonGrade

(Cancer Model)

Please see:
Epstein JI et al Am J Surg Path 2005: 29: 1228-42
Pierorazio PM et al. BJU Int 2013: 111: 753-60
For further detail

Code	Description
1	1
2	2
3	3
4	4
5	5

Usages

Gleason Grade

gleasonGradeTertiary

(Cancer Outcomes and Services Dataset)

Is there a different third grade in addition the primary and secondary grades and what is its value?

Rule x==~/[1-5]|8/

Usages

Pathology (Prostate)

gliomagrading

(Cancer Model)

Glioma (WHO 2007)

Code	Description
i	I
ii	II
iii	III
iv	IV

Usages

Cancer Specific Grading

gradeOfDifferentiation(pathological)

(Cancer Outcomes and Services Dataset)

GRADE OF DIFFERENTIATION (PATHOLOGICAL) is the definitive grade of the Tumour based on the evidence from a pathological examination.

Rule string(2)

Code	Description
G4	Undifferentiated / anaplastic
G3	Poorly differentiated
G2	Moderately differentiated
G1	Well differentiated
GX	Grade of differentiation is not appropriate or cannot be assessed

Usages

Sample Pathology

gradeOfDifferentiationAtDiagnosis

(Cancer Outcomes and Services Dataset)

GRADE OF DIFFERENTIATION (AT DIAGNOSIS) is the definitive grade of the Tumour at the time of PATIENT DIAGNOSIS.

Rule string(2)

Code	Description
GX	Grade of differentiation is not appropriate or cannot be assessed
G1	Well differentiated
G2	Moderately differentiated
G3	Poorly differentiated
G4	Undifferentiated / anaplastic

Usages

Diagnosis

healthAndCareNumber

(Genomics England Shared)

The HEALTH AND CARE NUMBER (H&C NUMBER) uniquely identifies a PATIENT within the NHS in Northern Ireland. It is the equivalent of the NHS NUMBER in England and Wales.

The HEALTH AND CARE NUMBER is ten numeric digits in length, and is in the same format as the NHS NUMBER in England (3 3 4 format with the tenth digit being a modulus 11 check digit). HEALTH AND CARE NUMBERS are however unique from NHS NUMBERS allocated in England as they are taken from a range of numbers reserved for Northern Ireland (320 000 001 to 399 999 999 plus check digit).

Regular Expression [a-zA-Z0-9]{10}

Rule based on import static javax.xml.bind.DatatypeConverter.*

xs:string (XMLSchema)	true && (x = parseString(string(x)))

Usages

Person Identifier

her2IshStatus

(Cancer Outcomes and Services Dataset)

Record the result of the ISH (in-situ hybridization) test.

Code	Description
P	Positive
N	Negative

Usages

Pathology (Breast)

her2Status

(Cancer Outcomes and Services Dataset)

Code	Description
P	Positive
N	Negative
B	Borderline

X	Not performed

Usages

Pathology (Breast)

imagingAnatomicalSite

(Cancer Outcomes and Services Dataset)

*IMAGING CODE (NICIP)
and/or
*IMAGING CODE (SNOMED CT)
and/or
*CANCER IMAGING MODALITY and IMAGING ANATOMICAL SITE and ANATOMICAL SIDE (IMAGING)
is required.

A classification of the part of the body that is the subject of an Imaging Or Radiodiagnostic Event.

Rule	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Imaging

imagingCode(NICIP)

(Cancer Outcomes and Services Dataset)

IMAGING CODE (NICIP) is the National Interim Clinical
Imaging Procedure Code Set code which is used to identify both the
test modality and body site of the test.

Rule	x==~/[a-zA-Z0-9]{1,6}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Imaging Code

infectionHistory

(Cancer Model)

Code	Description
none	none
previous	previous
current	current
unknown	unknown

Usages

Risk Factors for HPB Cancer

inrgNeuroblastoma

(Cancer Model)

INSS (neuroblastoma)

Code	Description
L1	Stage L1: The Tumour is located only in the area where it started; no IDRFs are found on imaging scans, such as CT or MR
L2	Stage L2: The tumour has not spread beyond the area where it started and the nearby tissue; IDRFs are found on imaging scans, such as CT or MR
M	Stage M: The tumour has spread to other parts of the body (except stage MS, see below)
MS	Stage MS: The tumour has spread to only the skin, liver, and/or bone marrow (less than 10% marrow involvement) in patients less than 18 months

Usages

Cancer Specific Staging

isrctNumber

(Cancer Model)

A simple numeric system for the unique identification of randomised controlled trials worldwide

Regular Expression	ISRCTN\d{8}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Registration

leibovichScore

(Cancer Model)

The Leibovich score is a scoring algorithm to predict survival for patients with metastatic renal cell carcinoma. Please provide, if applicable. 0-11, not more than 12

Regular Expression	[0-9]|[1][0-2]
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Cancer Specific Grading

lesionSizeRadiological

(Cancer Outcomes and Services Dataset)

Radiological estimate in millimetres of the maximum diameter of the tumour measured prior to treatment (largest lesion if more than one). Record as "0" to indicate not assessable for diffuse tumours (e.g. gliomatosis cerebri).

Unit of Measure	Millimetres
Regular Expression	^\d{1,3}(?:\.\d{1,2})?$
Rule based on xs:decimal (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDecimal(string(x)) in BigDecimal

Usages

Imaging (CNS)

locationSurgical

(Cancer Outcomes and Services Dataset)

Surgically determined anatomical location

Rule string(2)

Code	Description
01	Frontal lobe
02	Temporal lobe
03	Parietal lobe
04	Occipital lobe
05	Pineal region
06	Hypothalamic
07	Basal ganglia/thalamic
08	Cerebellar
09	Midbrain
10	Pons
11	Medulla

12	Fourth ventricle
13	Third ventricle
14	Lateral ventricle
15	Parasagittal/parafalcine dura
16	Posterior fossa convexity dura
17	Convexity dura
18	Petrous temporal bone
19	Orbital roof
20	Skull vault
21	Scalp
22	Anterior cranial fossa
23	Middle cranial fossa
25	Infratemporal fossa
26	Pterygopalatine fossa
27	Anterior clinoid dura
28	Sphenoid wing dura

29	Subfrontal dura
30	Suprasellar dura
31	Clival dura
32	Cavernous sinus
33	Cerebellopontine angle
34	Jugular bulb
35	Venous angle dura
36	Foramen magnum
37	Cervical intramedullary
38	Cervical intradural
39	Cervical extradural
40	Cervical bony
41	Thoracic intramedullary
42	Thoracic intradural
43	Thoracic extradural
44	Thoracic bony

45	Lumbar intramedullary
46	Lumbar intradural
47	Lumbar extradural
48	Lumbar bony
98	Other

Usages

Imaging (CNS)

Surgery (CNS)

lungMetastasesSubStageGrouping

(Cancer Outcomes and Services Dataset)

(For testicular cancer only)
Where lung metastases are identified, specify the RMH grouping.

Rule string(2)

Code	Description
L1	less than or equal to 3 metastases
L2	Greater than 3 metastases
L3	Greater than 3 metastases, one or more greater than or equal to 2cm diameter

Usages

Staging (Urology - Testicular)

mCategoryIntegratedStage

(Cancer Outcomes and Services Dataset)

This is the UICC code which classifies the absence or presence of distant metastases after treatment and/or after all available evidence has been collected.

Rule based on mCategory (Cancer Outcomes and Services Dataset)	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Component TNM

mCategoryPathological

(Cancer Outcomes and Services Dataset)

M CATEGORY (PATHOLOGICAL) is the Union for International Cancer Control (UICC) code which classifies the absence or presence of distant metastases based on the evidence from a pathological examination.

Rule based on mCategory (Cancer Outcomes and Services Dataset)	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

pTNM

mammogramResult

(Cancer Outcomes and Services Dataset)

Result of the mammogram. This will normally be the result of the mammogram taken at the first outpatient appointment at the breast clinic. If the patient attends more than one breast clinic, the result of each mammogram should be recorded.

Rule string(2)

Code	Description
R1	Normal
R2	Benign
R3	Uncertain
R4	Suspicious
R5	Malignant

Usages

Imaging (Breast)

mdtProvisionalDiagnosisIcd

(Cancer Outcomes and Services Dataset)

Working diagnosis as defined at MDT where the first definitive treatment is agreed. This is the clinical opinion which may also be informed by biopsy, radiological and/or other investigations.

Regular Expression	[a-zA-Z0-9.]{4,6}
Regular Expression based on ICD-10 (Cancer Outcomes and Services Dataset)	[a-zA-Z0-9.]{4,6}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Cancer Care Plan (CNS)

metastaticSite

(Cancer Outcomes and Services Dataset)

The site of the metastatic disease, if any, at diagnosis.

Rule string(2)

Code	Description
02	Brain
03	Liver
04	Lung
06	Multiple metastatic sites
07	Unknown metastatic site

08	Skin
09	Distant lymph nodes
10	Bone (excluding Bone Marrow)
11	Bone marrow
99	Other metastatic site

Usages

Diagnosis

modifiedDukes

(Cancer Outcomes and Services Dataset)

Dukes' stage of disease at diagnosis (based on pathological evidence but upgraded to Dukes D if clinical evidence of metastasis
Dukes D should be recorded if metastatic spread is identified either in the preoperative staging process, e.g. on CT scanning, MRI, USS, chest x-ray or at the time of operation.
It is accepted that a small number of D cases are cured by further treatment such as liver resection, but for COSD metastic spread distant from the primary should always be recorded as D.

Rule x==~/[a-zA-Z0-9]{1,2}/

Code	Description
A	Dukes A Tumour confined to wall of bowel, nodes negative
B	Dukes B Tumour penetrates through the muscularis propria to involve extramural tissues, nodes negative

C1	Dukes C1 Metastases confined to regional lymph nodes (node/s positive but apical node negative)
C2	Dukes C2 Metastases present in nodes at mesenteric artery ligature (apical node positive)
D	Dukes D Metastatic spread outside the operative field
99	Not Known

Usages

Cancer Specific Staging

molecularDiagnosticsCode

(Cancer Model)

Chromosomal or genetic markers associated with the brain tumour.

This may involve selection of more than one values for each tumour.

Code	Description
01	*Evidence of IDH1 or IDH2 mutation*
02	*Evidence of methylation of the MGMT gene CpG island*
03	*Evidence of total loss of 1p and 19q*
04	*Evidence of KIAA 1549-BRAF fusion gene*
05	*Other*

06	Evidence of ALK rearrangement
07	Evidence of native ALK
08	Evidence of ATRX mutation
09	Evidence of wt ATRX
10	Evidence of BRAF V600E mutation
11	Evidence of wt BRAF
12	Evidence of KIAA1549-BRAF fusion
13	Evidence of BRAF/RAF1 mutations, or fusions involving genes other than KIAA1549
14	Evidence of C11orf95-RELA fusion
15	Evidence of native C11orf95 and RELA
16	Evidence of amplification or fusion of C19MC locus (chr.19q13.42)
17	Evidence of unaltered C19MC locus (chr.19q13.42)
18	Evidence of CDK4/6 amplification
19	Evidence of CDK4/6 normal copy number
20	Evidence of CDKN2A locus homozygous deletion

21	Evidence of CDKN2A locus normal copy number
22	Evidence of CCND1/2/3 amplification
23	Evidence of CCND1/2/3 normal copy number
24	Evidence of CTNNB1 mutation
25	Evidence of wt CTNNB1
26	Evidence of amplification of EGFR
27	Evidence of mutation / rearrangement of EGFR
28	Evidence of unaltered EGFR
29	Evidence of EWSR1-FLI1 fusion
30	Evidence of native EWSR1 and FLI1
31	Evidence of FGFR1 mutation / rearrangement / fusion
32	Evidence of unaltered FGFR1
33	Evidence of H3F3A/H3F3B (H3.3) K27M mutation
34	Evidence of H3F3A/H3F3B (H3.3) wt K27
35	Evidence of H3F3A/H3F3B (H3.3) G34R/V mutation
36	Evidence of H3F3A/H3F3B (H3.3) wt G34

37	Evidence of HIST1H3B K27M mutation
38	Evidence of HIST1H3B wt K27
39	Evidence of HIST1H3C K27M mutation
40	Evidence of HIST1H3C wt K27
41	Evidence of ID2 amplification
42	Evidence of ID2 normal copy number
43	IDH1 (codon 132) or IDH2 (codon 172) mutation identified
44	IDH1 (codon 132) and IDH2 (codon 172) wt confirmed
45	Evidence of KLF4 K409Q and TRAF7 mutations
46	Evidence of wt KLF4 and TRAF7
47	Evidence of MAP2K1 mutation
48	Evidence of wt MAP2K1
49	Evidence of MET amplification
50	Evidence of MET normal copy number
51	Evidence of significant MGMT promoter methylation
52	Evidence of unmethylated MGMT promoter

53	Evidence of MYC/MYCN amplification
54	Evidence of MYC/MYCN normal copy number
55	Evidence of NF1 biallelic loss / mutation
56	Evidence of unaltered NF1
57	Evidence of NF2 biallelic loss / mutation
58	Evidence of unaltered NF2
59	Evidence of NKTR fusions
60	Evidence of native NKTR
61	Evidence of PTEN biallelic loss / mutation
62	Evidence of unaltered PTEN
63	Evidence of SDHB or SDHD mutation
64	Evidence of wt SDHB and SDHD
65	Evidence of SHH pathway activation
66	Evidence of normal SHH pathway
67	Evidence of inactivation of SMARCB1 (INI1)
68	Evidence of wt SMARCB1 (INI1)

69	Evidence of inactivation of SMARCA4
70	Evidence of wt SMARCA4
71	Evidence of TERT promotor mutation
72	Evidence of wt TERT promotor
73	Evidence of TP53 mutation
74	Evidence of wt TP53
75	Evidence of TSC1 or TSC2 mutation
76	Evidence of wt TSC1 and TSC2
77	Evidence of VHL mutation
78	Evidence of wt VHL gene
79	Evidence of WNT pathway activation
80	Evidence of normal WNT pathway
81	Evidence of WWTR1-CAMTA1 fusion
82	Evidence of native WWTR1 and CAMTA1
83	Evidence of codeletion of chr.1p and chr.19q
84	Evidence of total chr.1p loss but normal copy number of chr.19q

85	Evidence of normal copy number of both chr.1p and chr.19q
86	Evidence of monosomy chr.6
87	Evidence of chr.6 normal copy number
88	Evidence of polysomy chr.7
89	Evidence of chr.7 normal copy number
90	Evidence of loss of chr.10 or chr.10q
91	Evidence of chr.10 normal copy number
92	Evidence of loss of chr.22 or chr.22q
93	Evidence of chr.22 or chr.22q normal copy number
98	Other
99	Not Known (Not Recorded)

Usages

Pathology (CNS)

molecularSubgroupMedulloblastoma

(Cancer Model)

molecular subgroup (medulloblastoma)

Code	Description
not_tested	Not tested
shh	SHH
wnt	WNT
non_shh	Non-SHH
non_wnt	Non-WNT

Usages

Childhood Tumour Markers

molecularTestResult

(Genomics England Shared)

If no defect was observed please report 'normal'; if a mutation is detected that is considered pathogenically or clinically important record 'mutation detected'; if no reliable result could be determined please report 'fail'.

Code	Description
normal	Normal (negative)
fail	Fail
abnormalitydetected	Pathogenic abnormality detected
vus	Variant of unknown significance detected

Usages

Genetic Result

morphology(icd)

(Cancer Outcomes and Services Dataset)

Morphology ICD03 code

Regular Expression	[a-zA-Z0-9.\-\/]{5,7}
Regular Expression based on ICD-O-3 (Cancer Outcomes and Services Dataset)	[a-zA-Z0-9.\-\/]{5,7}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Morphology

morphology(snomed)

(Cancer Outcomes and Services Dataset)

This is the morphology of the tumour as categorised by SNOMED RT

Regular Expression	[a-zA-Z0-9]{6,8}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Morphology

morphology(snomedCt)

(Cancer Outcomes and Services Dataset)

Regular Expression	\d{6,18}
Regular Expression based on snomedCt (Cancer Outcomes and Services Dataset)	\d{6,18}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Morphology

myometrialInvasion

(Cancer Outcomes and Services Dataset)

Is there microscopic evidence of myometrial invasion?

Code	Description
1	None
2	Less than 50%
3	Greater than or equal to 50%

Usages

Pathology (Endometrial)

nCategoryIntegratedStage

(Cancer Outcomes and Services Dataset)

This is the UICC code which classifies the absence or presence and extent of regional lymph node metastases after treatment and/or after all available evidence has been collected.

Rule based on nCategory (Cancer Outcomes and Services Dataset)	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Component TNM

nCategoryPathological

(Cancer Outcomes and Services Dataset)

N CATEGORY (PATHOLOGICAL) is the Union for International Cancer Control (UICC) code which classifies the absence or presence and extent of regional lymph node metastases based on the evidence from a pathological examination.

Rule based on nCategory (Cancer Outcomes and Services Dataset)	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.*

```
true && (x = parseString(string(x)))
```

Usages

pTNM

nhsNumber

(NHS Data Dictionary GEL Subset)

Http://www.datadictionary.nhs.uk/data_dictionary/attributes/n/nhs/nhs_number_de.asp?query=nhs%20number&rank=100&shownav=1
The NHS NUMBER, the primary identifier of a PERSON, is a unique identifier for a PATIENT within the NHS in England and Wales. This will not vary by any ORGANISATION of which a PERSON is a PATIENT. It is mandatory to record the NHS NUMBER. There are exceptions, such as Accident and Emergency care, sexual health and major incidents, as defined in existing national policies. The NHS NUMBER is 10 numeric digits in length. The tenth digit is a check digit used to confirm its validity. The check digit is validated using the Modulus 11 algorithm and the use of this algorithm is mandatory. There are 5 steps in the validation of the check digit. Further guidance is available from the Health and Social Care Information Centre website.

Rule

```
def isValid = false
if (x.size() == 10) {
Integer total = 0
Integer i = 0
for (i = 0; i <= 8; i++) {
def digit = x.substring(i, (i+1))
def factor = 10 - i
total = total + (digit.toInteger() * factor) }
def checkDigit = (11 - (total.mod(11)))
if (checkDigit == 11) { checkDigit = 0 }
def check = x.substring(9,10)
if (check.toInteger() == checkDigit && checkDigit!=10) { isValid = true }
}
return isValid
```

Rule based on

```
import static javax.xml.bind.DatatypeConverter.*
```

xs:string (XMLSchema)	true && (x = parseString(string(x)))

Usages

Person Identifier

noCancerTreatmentReason

(Cancer Outcomes and Services Dataset)

The main reason why no active cancer treatment is specified within a Cancer Care Plan.

Rule	string(2)

Code	Description
01	Patient declined treatment
02	Unfit: poor performance status
03	Unfit: significant co-morbidity
04	Unfit: advanced stage cancer
05	Unknown primary site
06	Died before treatment
07	No active treatment available
08	Other

10	Monitoring only
99	Not Known

Usages

Cancer Care Plan

nodesExaminedNumberParaAortic

(Cancer Outcomes and Services Dataset)

The number of para-aortic nodes examined.
(Not applicable for vulval cancers) Use 0 if nodes not sent.

Regular Expression	\d{1,2}
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Pathology (Gynaecology)

nodesExaminedNumberPelvic

(Cancer Outcomes and Services Dataset)

The number of pelvic nodes examined (Not applicable for vulval cancers). Use 0 if nodes not sent

Rule	x ==~/\d{0,3}/
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Pathology (Gynaecology)

nodesPositiveNumberParaAortic

(Cancer Outcomes and Services Dataset)

The number of para-aortic nodes reported as being positive for the presence of tumour metastases.
(Not applicable for vulval cancers)

Regular Expression	\d{0,3}
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Pathology (Gynaecology)

nodesPositiveNumberPelvic

(Cancer Outcomes and Services Dataset)

The number of pelvic nodes reported as being positive for the presence of tumour metastases.
(Not applicable for vulval cancers)

Regular Expression	\d{1,2}
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Pathology (Gynaecology)

normalLdh

(Cancer Outcomes and Services Dataset)

TESTICULAR ONLY. This is the upper limit of normal for the LDH (Lactate Dehydrogenase Level) assay which is used to calculate S Category.

Regular Expression	\d{0,6}

Usages

Cancer Care Plan (Urology)

numberOfLesionsRadiological

(Cancer Outcomes and Services Dataset)

Radiologically determined number of lesions.

Regular Expression	\d{1,2}
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Imaging (CNS)

Staging (Upper GI)

numberOfNodesExamined

(Cancer Outcomes and Services Dataset)

The number of local and regional nodes examined.

Regular Expression	\d{1,3}

Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Sample Pathology

numberOfNodesPositive

(Cancer Outcomes and Services Dataset)

The number of local and regional nodes reported as being positive for the presence of Tumour metastases.

Regular Expression	\d{1,3}
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Sample Pathology

omentalInvolvement

(Cancer Outcomes and Services Dataset)

For endometrium, ovary, fallopian tube and primary peritoneum cancers, is there involvement of the omentum

Code	Description
1	Involved - deposit size not specified
2	Involved - deposit(s) 20mm or less

3	Involved - deposit(s) greater than 20mm
4	Not involved
X	Not assessable/Not sent

Usages

Pathology (Gynaecology)

opcsProcedureCodes

(Cancer Model)

OPCS Procedure Code. Allows for multiple codes delimited by "+" where the procedure cannot be described using a single code or information such as laterality is recorded.

Rule	x ==~/([a-zA-Z0-9.\-\/]{3,5}\+)*([a-zA-Z0-9.\-\/]{3,5})/
Based On	OPCS-4 (Cancer Outcomes and Services Dataset)
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Surgery And Other Procedures

organisationSiteCode

(NHS Data Dictionary GEL Subset)

An Organisation site code or ODS code identifies an NHS Organisation uniquely

Regular Expression	[a-zA-Z0-9]{3,9}
Rule based on organisationSiteCode (Cancer Outcomes and Services Dataset)	x ==~/[a-zA-Z0-9]{3,9}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Consultant Details

Registration

participantId

(Genomics England Shared)

Genomics England participant identifier (supplied by Genomics England)

Regular Expression	\d{9}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Participant Identifiers

pathologyInvestigationType

(Cancer Model)

The type of pathology investigation carried out. Although this item is based on COSD CR0760, an additional value 'BM' for Bone Marrow Aspirate has been added for the purposes of this project in order to collect haematological bone marrow aspirate samples.

Code	Description
CY	Cytology
BU	Biopsy NOS
EX	Excision
PE	Partial Excision
RE	Radical Excision
FE	Further Excision
CU	Curettage
SB	Shave Biopsy
PB	Punch Biopsy
IB	Incisional Biopsy
BM	Bone Marrow Aspirate
99	Uncertain/other

Usages

Sample Pathology

pathologyReportText

(Cancer Outcomes and Services Dataset)

The full text from the pathology report which may be required by Registries to calculate diagnosis and staging details

Rule	maxLength(270000)
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Sample Pathology

performanceStatusAdult

(Cancer Outcomes and Services Dataset)

A World Health Organisation classification indicating a PERSON's status relating to activity / disability.

Code	Description
0	Able to carry out all normal activity without restriction
1	Restricted in physically strenuous activity, but able to walk and do light work
2	Able to walk and capable of all self care, but unable to carry out any work. Up and about more than 50% of waking hours
3	Capable of only limited self care, confined to bed or chair

	more than 50% of waking hours
4	Completely disabled. Cannot carry on any self care. Totally confined to bed or chair
9	Not recorded

Usages

Cancer Care Plan

peritonealWashings

(Cancer Outcomes and Services Dataset)

Were peritoneal washings submitted and if so were malignant cells seen?

Code	Description
1	Positive
2	Negative
X	Not sent/Not assessable

Usages

Pathology (Endometrial)

personFamilyName

(Cancer Outcomes and Services Dataset)

That part of a PERSON's name which is used to describe family, clan, tribal group, or marital association.

Regular Expression	[a-zA-Z0-9\s-]{3,35}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Participant Identifiers

personFamilyNameAtBirth

(Cancer Outcomes and Services Dataset)

The PATIENT's surname at birth.

Regular Expression	[a-zA-Z0-9\s-]{3,35}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Registration

personGivenName

(Cancer Outcomes and Services Dataset)

The forename(s) or given name(s) of a PERSON.

Regular Expression	[a-zA-Z0-9\s-]{3,35}
Rule based on	import static javax.xml.bind.DatatypeConverter.*

xs:string (XMLSchema)	true && (x = parseString(string(x)))

Usages

Participant Identifiers

personHeightInMetres

(Genomics England Shared)

Height of the patient, in metres, to 2 decimal places (n.nn).

Unit of Measure	meter (The 1889 definition of the metre, based on the international prototype of platinum-iridium, was replaced by the 11th CGPM (1960) using a definition based on the wavelength of krypton 86 radiation. This change was adopted in order to improve the accuracy with which the definition of the metre could be realized, the realization being achieved using an interferometer with a travelling microscope to measure the optical path difference as the fringes were counted. In turn, this was replaced in 1983 by the 17th CGPM (1983, Resolution 1) that specified the current definition, as follows: The metre is the length of the path travelled by light in vacuum during a time interval of 1/299 792 458 of a second. It follows that the speed of light in vacuum is exactly 299 792 458 metres per second, c0 = 299 792 458 m/s. The original international prototype of the metre, which was sanctioned by the 1st CGPM in 1889, is still kept at the BIPM under conditions specified in 1889.
Regular Expression	\d{1}\.\d{1,2}
Rule based on xs:double (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDouble(string(x)) in Double

Usages

personObservationWeight

(Cancer Outcomes and Services Dataset)

Weight of the patient, in kilograms with up to three decimal places (nnn.nnn).

Unit of Measure	kilogram (The international prototype of the kilogram, an artefact made of platinum-iridium, is kept at the BIPM under the conditions specified by the 1st CGPM in 1889 when it sanctioned the prototype and declared: This prototype shall henceforth be considered to be the unit of mass. The 3rd CGPM (1901), in a declaration intended to end the ambiguity in popular usage concerning the use of the word "weight", confirmed that: The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram. The complete declaration appears here. It follows that the mass of the international prototype of the kilogram is always 1 kilogram exactly, m(grand K) = 1 kg. However, due to the inevitable accumulation of contaminants on surfaces, the international prototype is subject to reversible surface contamination that approaches 1 µg per year in mass. For this reason, the CIPM declared that, pending further research, the reference mass of the international prototype is that immediately after cleaning and washing by a specified method (PV, 1989, 57, 104-105 and PV, 1990, 58, 95-97). The reference mass thus defined is used to calibrate national standards of platinum-iridium alloy (Metrologia, 1994, 31, 317-336).
Regular Expression	\d{1,3}\.\d{1,3}
Rule based on xs:double (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDouble(string(x)) in Double

Usages

General Risk Factors

personPhenotypicSexClassification

(NHS Data Dictionary GEL Subset)

A classification of PERSON PHENOTYPIC SEX
http://www.datadictionary.nhs.uk/data_dictionary/attributes/p/person/person_phenotypic_sex_classification_de.asp?shownav=1

Based On personPhenotypicSex (Genomics England Shared)

Code	Description
2	Female
1	Male
9	Indeterminate

Usages

Registration

personStatedGenderCode

(Genomics England Shared)

The participant's current gender. COSD v7 update

Code	Description
1	Male
2	Female

9	Indeterminate (Unable to be classified as either male or female)
X	Not Known (PERSON STATED GENDER CODE not recorded)

Usages

Registration

portalInvasion

(Cancer Outcomes and Services Dataset)

Record whether there is involvement of the portal vein

Code	Description
Y	Present
N	Not present
9	Not known

Usages

Staging (Upper GI)

positiveNegativeUnknown

(Genomics England Shared)

Positive negative or unknown result

Based On posNegUnk (Genomics England Shared)

Code	Description
unknown	unknown
negative	negative
positive	positive

Usages

Risk Factors for Ovarian Cancer

primaryDiagnosisIcd

(Cancer Outcomes and Services Dataset)

See DIAGNOSTIC CODING for details on coding and PRIMARY DIAGNOSES for the standardised definition of primary diagnosis

Regular Expression	[a-zA-Z0-9.]{3,6}
Regular Expression based on ICD-10 (Cancer Outcomes and Services Dataset)	[a-zA-Z0-9.]{4,6}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Diagnosis

Related Cancer Diagnoses

Sample Pathology

primaryDiagnosisIcdRadiological

(Cancer Outcomes and Services Dataset)

The preliminary primary diagnosis based on radiological examination recorded pre treatment. In many cases this will be the definitive clinical diagnosis, but needs to be distinguished from the subsequent pathological diagnosis - if it becomes available.

Rule	x ==~/[a-zA-Z0-9.]{4,6}/
Regular Expression based on primaryDiagnosisIcd (Cancer Outcomes and Services Dataset)	[a-zA-Z0-9.]{3,6}
Regular Expression based on ICD-10 (Cancer Outcomes and Services Dataset)	[a-zA-Z0-9.]{4,6}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Cancer Care Plan (CNS)

principalDiagnosticImagingType

(Cancer Outcomes and Services Dataset)

Indicate the principal imaging procedure undertaken to diagnose the tumour.

NB: PET Scan also includes PET-CT Scan

Code	Description

1	CT Scan
2	MRI Scan
3	PET Scan

Usages

Imaging (CNS)

psaDiagnosis

(Cancer Outcomes and Services Dataset)

PROSTATE ONLY. Prostate Specific Antigen blood level in ng/ml, measured at time of diagnosis.

Unit of Measure	nanogrammes per microliter (Nanogrammes per microliter
Regular Expression	\d{1,5}\.\d{1}
Rule based on xs:decimal (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDecimal(string(x)) in BigDecimal

Usages

Circulating Tumour Markers (Prostate)

psaPreTreatment

(Cancer Outcomes and Services Dataset)

PROSTATE ONLY. Prostate Specific Antigen blood level in ng/ml, measured before treatment (including second and subsequent treatments).

This is the PSA taken prior to EACH treatment (because some curative treatments may be delivered years after diagnosis.

Regular Expression	\d{1,5}\.\d{1}
Rule based on xs:decimal (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseDecimal(string(x)) in BigDecimal

Usages

Circulating Tumour Markers (Prostate)

radiotherapyInChildhood

(Cancer Model)

Data type for Radiotherapy in Childhood

Code	Description
cns	CNS
non_cns	non_CNS
none	none
unknown	unknown

Usages

Risk Factors for Glioma Cancer

radiotherapyIntent

(Cancer Outcomes and Services Dataset)

RADIOTHERAPY INTENT

Code	Description
01	Palliative
02	Anti-cancer
03	Other

Usages

Radiotherapy

radiotherapyTotalDose

(Cancer Outcomes and Services Dataset)

Max n3.n2

Regular Expression	^\d{1,3}(?:\.\d{1,2})?$

Usages

Radiotherapy

radiotherapyTotalFractions

(Cancer Outcomes and Services Dataset)

Max n2

Regular Expression	\d{1,2}
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

External Beam

sarcomatoidGrading

(Cancer Model)

Code	Description
present	present
absent	absent

Usages

Cancer Specific Grading

satelliteTumourNodulesLocation

(Cancer Outcomes and Services Dataset)

Record the most distant location of separate tumour nodules

Code	Description
1	Separate tumour nodules in same lobe

2	Separate tumour nodules in a different ipsilateral lobe
3	Separate tumour nodules in a contralateral lobe
4	No separate tumour nodules
9	Not known

Usages

Pathology (Lung)

serviceReportIdentifier

(Cancer Outcomes and Services Dataset)

A unique identifier of a SERVICE REPORT.

Rule	x==~/[a-zA-Z0-9]{1,18}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Sample Pathology

smokingStatus

(Cancer Outcomes and Services Dataset)

Specify the current smoking status of the patient.

Code	Description
1	Current smoker
2	Ex smoker
3	Non-smoker - history unknown
4	Never smoked
Z	Not Stated (PERSON asked but declined to provide a response)
9	Unknown

Usages

General Risk Factors

snomed

(Cancer Model)

Snomed ct or rt codes

Regular Expression	[a-zA-Z0-9]{6,18}

Usages

Morphology (SNOMED)

Topography (SNOMED)

snomedCt

(Genomics England Shared)

SNOMED CT CODE

Regular Expression	\d{6,18}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Sample Pathology

Imaging Code

snomedVersion

(Cancer Model)

The version of SNOMED used to encode MORPHOLOGY (SNOMED) PATHOLOGY and TOPOGRAPHY (SNOMED) PATHOLOGY

Versions of SNOMED prior to SNOMED CT cease to be licenced by The International Health Terminology Standards Development Organisation (IHTSDO) after April 2017 other than for historical content

Code	Description
01	SNOMED II
02	SNOMED 3
03	SNOMED 3.5
04	SNOMED RT

05	SNOMED CT
99	Not Known

Usages

Morphology (SNOMED)

Topography (SNOMED)

specialty

(Cancer Model)

The specialty code of the person performing the event

Regular Expression	[a-zA-Z0-9]{3}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Cancer Specific Treatments

Radiotherapy

Surgery And Other Procedures

Systemic Anti-Cancer Therapy

stageGroupingTesticular

(Cancer Outcomes and Services Dataset)

TESTICULAR ONLY. Nationally agreed anatomical stage groupings as defined by The Royal Marsden Hospital (RMH).

Code	Description
1	Stage 1
1S	Stage 1S
1M	Stage 1M
2A	Stage 2A
2B	Stage 2B
2C	Stage 2C
3A	Stage 3A
3B	Stage 3B
3C	Stage 3C
4A	Stage 4A
4B	Stage 4B
4C	Stage 4C

Usages

Staging (Urology - Testicular)

synchronousTumourIndicator

(Cancer Model)

Record any synchronous tumours in the Colon as identified by the clinician at presentation. Synchronous tumours are defined as discrete tumours apparently not in continuity with other primary cancers originating in the same site or tissue.

Code	Description
1	CAECUM
2	APPENDIX
3	ASCENDING COLON
4	HEPATIC FLEXURE
5	TRANSVERSE COLON
6	SPLENIC FLEXURE
7	DESCENDING COLON
8	SIGMOID COLON
9	RECTOSIGMOID
10	RECTUM

Usages

Diagnosis (Colorectal)

tCategoryIntegratedStage

(Cancer Outcomes and Services Dataset)

This is the UICC code which classifies the size and extent of the primary tumour after treatment and/or after all available evidence has been collected.

Rule based on tCategory (Cancer Outcomes and Services Dataset)	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Component TNM

tCategoryPathological

(Cancer Outcomes and Services Dataset)

T CATEGORY (PATHOLOGICAL) is the Union for International Cancer Control (UICC) code which classifies the size and extent of the primary Tumour based on the evidence from a pathological examination.

Rule based on tCategory (Cancer Outcomes and Services Dataset)	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

pTNM

tnmEditionNumber

(Cancer Outcomes and Services Dataset)

The UICC edition number used for Tumour, Node and Metastasis (TNM) staging for cancer diagnosis.

Rule	x==~/[a-zA-Z0-9]{1,2}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Integrated TNM

pTNM

tnmStageGroupingIntegrated

(Cancer Outcomes and Services Dataset)

Record the overall TNM stage grouping of the tumour, derived from each T, N and M component after treatment. This classification is based on all the evidence available to the clinician(s) with responsibility for assessing the patient. Such evidence arises from physical examination, imaging, endoscopy, biopsy, surgical exploration and other relevant examinations.

The overall integrated TNM stage grouping indicates the tumour stage after treatment and/or after all available evidence has been collected.

Note: Use UICC coding.

Rule	x==~/[a-zA-Z0-9]{1,5}/
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

TNM Details

topographyIcdo3

(Cancer Outcomes and Services Dataset)

The topographical site code for the tumour as defined by ICDO3. This will normally be derived by Registries.

Regular Expression	[a-zA-Z0-9.\-\/]{3,7}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Topography

topographySnomed

(Cancer Outcomes and Services Dataset)

This is the topographical site of the tumour as categorised by SNOMED RT

Regular Expression	[a-zA-Z0-9]{6,8}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Topography

topographySnomedCt

(Cancer Outcomes and Services Dataset)

For use in pilot project only at present. Please contact cosd@ncin.org.uk for further details.

This is the topographical site of the tumour as categorised by SNOMED CT.

Regular Expression based on snomedCt (Cancer Outcomes and Services Dataset)	\d{6,18}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Topography

tubeInvolvement

(Cancer Outcomes and Services Dataset)

For endometrial and fallopian cancers, is there microscopic involvement

Code	Description
1	Not involved
2	Right involved
3	Left involved
4	Both involved

X	Not assessable

Usages

Pathology (Gynaecology)

tumourGradeOvarianSerous

(Cancer Model)

Code	Description
l	Low
h	High

Usages

Cancer Specific Grading

tumourGradeUrology

(Cancer Outcomes and Services Dataset)

BLADDER ONLY.
Specify whether LOW, HIGH Grade or PUNLMP (Papillary Urothelial Neoplasm of Low Maligant Potential).

Code	Description
L	Low
H	High

P	PunImp
X	Not applicable

Usages

Cancer Specific Grading

tumourID

(Cancer Model)

Genomics England Tumour Identifier

Regular Expression	[a-zA-Z0-9]{3,9}_[a-zA-Z0-9]{1,16}
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Cancer Care Plan

Cancer Specific Treatments

Circulating Tumour Markers (Ovarian)

Circulating Tumour Markers (Prostate)

Diagnosis

Disease Information Update (Tumour Sample)

Genetic Results

Imaging

Investigation Report Other

Next Generation Sequencing

Presentation

Radiotherapy

Sample Pathology

Surgery And Other Procedures

tumourLaterality(pathological)

(Cancer Outcomes and Services Dataset)

Tumour laterality identifies the side of the body for a tumour relating to paired organs within a PATIENT based on the evidence from a pathological examination.

Code	Description
B	Bilateral
R	Right
L	Left
M	Midline
9	Not Known
8	Not applicable

Usages

Diagnosis

tumourSampleNotSentReason

(Cancer Model)

Reason tumour sample not sent from GMC to Biorepository

Code	Description
tumour_sample_not_taken	Tumour sample not taken
tumour_type_not_eligible	Tumour type not eligible
poorly_cellular_tumour	Poorly cellular tumour (Less than 40 percent neoplastic cells)
insufficient_tumour_post_neoadjuvant_chemotherapy	Insufficient tumour post neoadjuvant chemotherapy
insufficient_dna	Insufficient DNA
no_cancer_diagnosed	No Cancer Diagnosed
ffpe_not_optimally_fixed	FFPE not optimally fixed
ffpe_not_optimally_processed	FFPE not optimally processed
high_necrosis	High necrosis (over 20 percent)
other	Other

Usages

Reason Sample Not Sent

tumourType

(Cancer Model)

Code	Description

primary	Primary; source of cancer tumour sample
recurrence_of_primary_tumour	Recurrence; a tumour has returned at the site of the original cancer
metastatic_recurrence	Metastatic (different cancer site) which developed and was sampled after presentation
metastases	Metastatic (different cancer site) which was present and sampled at diagnosis instead of the primary tumour

Usages

Sample Pathology

turpTumourPercentage

(Cancer Outcomes and Services Dataset)

For TURP only, what percentage of tumour if clinically unsuspected tumour.

Regular Expression	[0-9]{1,2}|100
Rule based on Integer (Cancer Outcomes and Services Dataset)	is Integer

Usages

Pathology (Prostate)

ukTelephoneNumber

(Cancer Model)

uk phone number

Regular Expression	^(((\+44\s?\d{4}|\(?0\d{4}\)?)\s?\d{3}\s?\d{3})|((\+44\s?\d{3}|\(?0\d{3}\)?)\s?\d{3}\s?\d{4})|((\+44\s?\d{2}|\(?0\d{2}\)?)\s?\d{4}\s?\d{4}))(\s?\#(\d{3,5}))?$
Rule based on xs:string (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* true && (x = parseString(string(x)))

Usages

Consultant Details

Participant Contact Details

ultrasoundExaminationResult

(Cancer Outcomes and Services Dataset)

Result of the ultrasound examination.

Rule	string(2)

Code	Description
U1	Normal
U2	Benign
U3	Indeterminate/probably benign
U4	Suspicious of malignancy

U5	Highly suspicious of malignancy

Usages

Imaging

unitOfMeasurement

(NHS Data Dictionary GEL Subset)

Unit of measurement
http://www.datadictionary.nhs.uk/data_dictionary/attributes/u/unit_of_measurement_de.asp?shownav=1

Code	Description
35	Kilocalories (kcal)
36	Millimoles (mmol)
33	International Units per kilogram (IU/kg)
34	Grams (g)
39	Milligrams per millimole (mg/mmol)
37	Millimoles per mole (mmol/mol)
38	Picomoles per litre (pmol/L)
43	Cubic Millimetres (mm3)
42	Millimetres of water (mmH2O)

41	Micrograms per millilitre (µg/ml)
40	Nanograms per litre (ng/l)
22	Celsius (ºC)
23	Millimetres (mm)
24	Grams per decilitre (g/dl)
25	Grams per litre (g/l)
26	Milligrams per litre (mg/l)
27	Nanograms per millilitre (ng/ml)
28	International Units per litre (IU/L)
29	Decilitres (d/l)
30	Square Millimetres (mm2)
32	Grays (Gy)
31	Millilitres (ml) (Retired September 2013)
19	Milligrams (mg)
17	Beats per minute (bpm)
18	Centimetres (cm)

15	Millimetres of mercury (mmHg)
16	Litres (l)
13	Square Metres (m2)
14	Millilitres per Minute (ml/min)
11	Metres (m)
12	Picograms (pg)
21	Minutes
20	Millilitres (ml)
49	Kilopascals (KPa)
48	Grams per kilogram per day (g/kg/day)
08	Number (Retired September 2013)
45	Millilitres per Minute divided by 1.73 Square Metres (ml/min/1.73m2)
09	Percentage (%)
44	Litres per week per 1.73 metres squared (l/week/1.73^2)
47	5 Millimetres Squared
46	number times ten raised to the power of nine per litre

	(x109/l)
04	Micrograms per millimole (µg/mmol)
05	Microgram albumin per hour (µg/ml/hr)
06	Microgram albumin per minute (µg/min)
07	Microgram albumin per 24 hours (µg/24hr)
01	Millimoles per litre (mmol/L)
02	Micromoles per litre (µmol/L)
03	Micrograms per litre (µg/L)
10	Kilograms (kg)
51	Megavolts
52	5 Millimetres Squared
50	Femtolitres (fl)

Usages

Report Attribute

whoTumourGradeCns

(Cancer Outcomes and Services Dataset)

The grade of the tumour using WHO classification for tumours of the central nervous system. FOR INTRA AXIAL AND EXTRA AXIAL ONLY.

Code	Description
1	I
2	II
3	III
4	IV

Usages

Cancer Specific Grading

xs:base64Binary

(XMLSchema)

Base64-encoded arbitrary binary data

Regular Expression	[a-zA-Z0-9=]*

Usages

Report Attribute

xs:date

(XMLSchema)

Calendar date.Format YYYY-MM-DD. Example, May the 31st, 1999 is: 1999-05-31.

Rule	import static javax.xml.bind.DatatypeConverter.*
	parseDateTime(string(x)) in Calendar

Usages

Cancer Care Plan

Participant Identifiers

Sample Details

Sample Pathology

xs:dateTime

(XMLSchema)

Specific instant of time. ISO 8601 extended format YYYY-MM-DDThh:mm:ss. Example, to indicate 1:20 pm on May the 31st, 1999 for Eastern Standard Time which is 5 hours behind Coordinated Universal Time (UTC): 1999-05-31T13:20:00-05:00.

Rule	import static javax.xml.bind.DatatypeConverter.*
	parseDateTime(string(x)) in Calendar

Usages

Event Details

Report Attribute

xs:double

(XMLSchema)

Double-precision 64-bit floating point type legal literals {0, -0, INF, -INF and NaN} Example, -1E4, 12.78e-2, 12 and INF

Rule	import static javax.xml.bind.DatatypeConverter.*

	parseDouble(string(x)) in Double

Usages

Circulating Tumour Markers (Ovarian)

xs:nonNegativeInteger

(XMLSchema)

Infinite set {0, 1, 2,...}. Sign omitted, “+” assumed. Example: 1, 0, 12678967543233, +100000.

Rule	minInclusive(0)
Rule based on xs:integer (XMLSchema)	import static javax.xml.bind.DatatypeConverter.* parseInteger(string(x)) in BigInteger

Usages

General Risk Factors

Risk Factors for Breast Cancer

Risk Factors for Endometrial Cancer

Risk Factors for Ovarian Cancer

Risk Factors for Renal Cancer

xs:string

(XMLSchema)

Character strings in XML.

Rule	import static javax.xml.bind.DatatypeConverter.*
	true && (x = parseString(string(x)))

Usages

AJCC Stage

Cancer Care Plan

Consent

Consent Update

Consultant Details

Event Details

Final Figo Stage

Genetic Result

Imaging

Investigation Report Other

Next Generation Sequencing

Radiotherapy

Report Attribute

Sample Details

Sample Pathology

Withdrawal

yesNo

(Genomics England Shared)

Boolean, yes no response

Code	Description
yes	Yes
no	No

Usages

Consent

Consent Details

Consent Update

Pathology (Kidney)

yesNoNa

(Cancer Outcomes and Services Dataset)

yes, no, not applicable

Code	Description
Y	Yes
N	No
X	Not applicable

Usages

Pathology (Prostate)

Pathology (Testes)

yesNoNk

(Cancer Outcomes and Services Dataset)

yes, no, not known

Code	Description
Y	Yes
N	No
9	Not known

Usages

Cancer Care Plan (Lung)

Other Treatment (Bladder)

Other Treatment (Upper GI)

Pathology (Lung)

Staging (Upper GI)

yesNoNotAssessable

(Cancer Outcomes and Services Dataset)

yes, no, not assessable

Code	Description
Y	Yes
N	No
X	Not Assessable
9	Not Known

Usages

Pathology (Endometrial)

Pathology (Gynaecology)

Pathology (Prostate)

yesNoNotRelevant

(Genomics England Shared)

yes, no, not relevant

Code	Description
yes	yes
no	no
not_relevant	not relevant

Usages

Consent Details

yesNoUnc

(Cancer Outcomes and Services Dataset)

Code	Description
Y	Yes
N	No
U	Uncertain

Usages

Pathology (Kidney)

yesNoUnk

(Genomics England Shared)

Code	Description
yes	Yes
no	No
unknown	Unknown

Usages

Risk Factors for Ovarian Cancer

Risk Factors for Testicular Cancer

Sample Pathology

8 Business Rules

8.1.1 Additional Findings reporting during consent update (40309)

Data Elements	Name and Version of Consent Form Update Health Related Additional Findings Reproductive Additional Findings
Component	**Consent update, open Clinica and Mercury**
Rule Focus	**Additional findings questions available based on consent form chosen**
Trigger	**Consent xml submitted through mercury or consent question selected in open clinica**
Description	**This rule prompts users to offer the right additional finding questions when updating a consent.**
Error Condition	**WARNING**
Issue Record	**Email, BuRST**
Notification	**Submission processed, warning logged**
Notification Target	**GMC submitter, GEL service management**
Last Updated	**2016-10-12**
Version Created	**2016-09-28**
Status	**FINAL**

8.1.2 Disease Type and Subtype Consistency (42262)

Data Elements	Disease Type Disease Subtype Disease Type
Component	**Sample Tracking**
Rule Focus	**GMC to GEL Submissions**
Trigger	**GMC GEL Sample Metadata CSV Received, Registration, Disease Information Received**
Description	**Combination of disease type and subtype submitted must be consistent with the combinations marked within Appendix A.**
Error Condition	**WARNING**
Issue Record	**Email, BuRST**
Notification	**Submission processed, warning logged**
Notification Target	**GMC submitter, GEL service management**
Last Updated	**2016-10-12**
Version Created	**2016-09-28**
Status	**FINAL**

8.1.3 Consent for additional findings (40300)

Data Elements	Name and Version of Consent Form Health Related Additional Findings Reproductive Additional Findings
Component	**Consent, Open Clinica and Mercury**
Rule Focus	**Additional findings questions available based on consent form chosen**
Trigger	**Consent xml submitted through mercury or consent question selected in open clinica**
Description	**This rule allows the right choice of additional findings question to be offered based on the consent form issued.**
Error Condition	**NA for open clinica. Mercury should raise a warning log**
Issue Record	**Logs**
Notification	**Logs**
Notification Target	**System administrator**
Last Updated	**2016-08-01**
Version Created	**2016-07-01**
Status	**FINAL**

8.1.4 Consent form corresponding to patient information sheet (40301)

Data Elements	Name and Version of Consent Form Name and Version of Participation Information Sheet Name and Version of Consent Form Update Name and Version of Participant Information Sheet Update
Component	**Consent, Open Clinica and Mercury**
Rule Focus	**Patient Information sheet available based on consent form chosen**
Trigger	**Consent xml submitted through mercury or consent question selected in open clinica**
Description	**This rule allows the right choice of patient information sheet to correspond with consent form chosen**
Error Condition	**WARNING**
Issue Record	**Email, BuRST**
Notification	**Submission processed, warning logged**
Notification Target	**GMC submitter, GEL service management**
Last Updated	**2016-10-12**
Version Created	**2016-09-28**
Status	**FINAL**

8.1.5 Consent options must be consistent with Appendix F (42277)

Data Elements	Name and Version of Consent Form Name and Version of Participation Information Sheet Name and Version of Consent Form Update Name and Version of Participant Information Sheet Update Name and Version of Assent Form Additional optional consent Materials
Component	**Consent, Open Clinica and Mercury**
Rule Focus	**Patient Information sheet available based on consent form chosen**
Trigger	**Consent xml submitted through mercury or consent question selected in open clinica**
Description	**Consent Forms Must be consistent with the enumerations in Appendix F. NOTE: Appendix F will be periodically updated.**
Error Condition	**VALIDATION ERROR**
Issue Record	**Email, BuRST**
Notification	**Submission Failing Validation, CSV rejected**
Notification Target	**GMC submitter, GEL service management**
Last Updated	**2016-10-12**

Version Created	**2016-09-28**
Status	**FINAL**