**Authors:**

**Adam Milward**

GENOMICS ENGLAND

*100,000 Genomes Project Release Notes*

*FINAL*

*Cancer Model - Release v3.1.2*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Document Record ID Key | |
| Work stream | Informatics |  | |
| Programme Director | Peter Counter | **Status** | FINAL |
| Document Owner | Amanda O’Neill | **Version** | v3.1.2\_01 |
|  |  |  |  |
| Document Author | Listed Below | **Version Date** | 15/11/2016 |

Contents

2 Document Control 2

2.1 Version History 2

2.2 Distribution list 2

2.3 Approvers 2

3 Introduction 3

3.1 Purpose 3

3.2 Audience 3

4 What’s New? 4

4.1 Additions 4

4.2 Removals 4

4.3 Changes 4

5 Package Contents 4

5.1 Cancer Clinical Data Specifications – XML Submissions and Open Clinica Forms: 4

5.2 Cancer Clinical Data XSDs and examples 5

5.3 Cancer Sample Tracking Data Specifications – CSV Submissions and examples: 6

5.4 Cancer Appendices: Reference Data and Guidance 6

6 Known Deficiencies 7

6.1 Inherent Defects 7

7 Impact of Release 7

7.1 GMC Impact 7

7.2 Biorepository Impact 7

7.3 Sequencer Impact 7

7.4 Support Impact 7

# Document Control

The controlled copy of this document is maintained in the Genomics England internal document management system. Any copies of this document held outside of that system, in whatever format (for example, paper, email attachment), are considered to have passed out of control and should be checked for currency and validity. This document is uncontrolled when printed.

## Version History

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Author | Date | Description |
| 3.1.2\_01 | A Milward | 15/11/16 | Final |
|  |  |  |  |

## Distribution list

This document must be reviewed by the following:

|  |  |
| --- | --- |
| Name | Organisation |
| Amanda O’Neill | Genomics England |
| Oliver Freeman | Oxford |
| Miles Grey | NHS England |

## 

## Approvers

This document must be approved by the following:

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Responsibility | Date | Version |
| Peter Counter | Programme Director | 16 Nov 16 | v3.1.2\_01 |
| Tom Fowler | Deputy Chief Scientist | 16 Nov 16 | v3.1.2\_01 |
|  |  |  |  |

# Introduction

## Purpose

The purpose of this document is to provide a brief summary of the 3.1.0 Cancer Data Model Package.

## 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.

# What’s New?

Summary of the main changes:

NOTE: specific changes can be found in the change tab of the Excel data specifications

## Additions

* New cancers, disease types and subtypes
* Disease types, subtypes, consent materials and presentations moved to appendices, as reference data, to provide a more flexible means of submitting data if and when types, subtypes and consent materials change. GMCs are advised to make their data acquisition systems flexible in these areas, so that they can accommodate additional subtypes or consent materials.
* Cancer specific sections including imaging, tumour markers, care plans, staging, surgery and diagnosis, for the new cancers
* Sample tracking updates to support handling of multiple samples and biopsies
* Sample tracking updates to improve quality control
* Stricter validation to ensure erroneous data is flagged earlier in the data lifecycle
* Addition of business rules to the package
* Addition of sample tracking appendices to the package
* Reason sample not sent section/schema added, element removed from sample tracking csv submission

## Removals

* Additional Data section removed
* Family History section removed

## Changes

* Referrals section and schema renamed Presentations
* Alignment with COSD and GEL Rare Disease models
* Clarifications, corrections and improvements to address issues raised during v2 and v3 consultation

# Package Contents

## Cancer Model Overview

1. **Cancer Model Conceptual Overview-3.1.2.pdf** – General overview and conceptual model for the data structures in the specification

## Cancer Clinical Data Specifications for XML Submissions and Open Clinica Forms:

1. **Cancer Model-3.1.2.docx** - Cancer Clinical Data Specification as a Word Document describing the XML message definitions for Genomic Medicine Centres (GMCs) to provide clinical data for the programme from registration, to the description of their cancer diagnosis and stage, through their treatment pathway.
2. **Cancer Model-3.1.2.xslx** - Cancer Clinical Data Specification AND CHANGE LOG as an Excel Document describing the XML message definitions for Genomic Medicine Centres (GMCs) to provide clinical data for the programme from registration, to the description of their cancer diagnosis and stage, through their treatment pathway.

CHANGES: New Additions to the model are highlighted in Green, Removals are highlighted in Red, Superficial changes (such as descriptions) are highlighted in blue.

## Cancer Clinical Data XSDs and examples

1. **XSDs** - Contains the XML Schemas used to validate the XML messages submitted by GMCs.

The following XSDs are included:

* RegistrationAndConsentsCancer-v3.1.2.xsd
* RiskFactorAssessmentsCancer-v3.1.2.xsd
* DiagnosesCancer-v3.1.2.xsd
* DiseaseInformationUpdatesCancer-v3.1.2.xsd
* CarePlansCancer-v3.1.2.xsd
* ConsentUpdatesCancer-v3.1.2.xsd
* InterventionsCancer-v3.1.2.xsd
* InvestigationsCancer-v3.1.2.xsd
* PresentationsCancer-v3.1.2.xsd
* ReasonsSampleNotSent-v3.1.2.xsd
* Withdrawals-v3.1.2.xsd
* DeathsCancer-v3.1.2.xsd

1. **Example XML** Files

The following XML files are included:

* example\_3.1.2\_reason\_sample\_not\_sent\_2.xml
* example\_3.1.2\_presentations\_1.xml
* example\_3.1.2\_disease\_information\_update\_2.xml
* example\_3.1.2\_death\_1.xml
* example\_3.1.2\_withdrawals\_2.xml
* example\_3.1.2\_care\_plans\_1.xml
* example\_3.1.2\_care\_plan\_2.xml
* example\_3.1.2\_riskfactors\_1.xml
* example\_3.1.2\_consent\_updates\_2.xml
* example\_3.1.2\_diagnoses\_2.xml
* example\_3.1.2\_diagnoses\_1.xml
* example\_3.1.2\_interventions\_1.xml
* example\_3.1.2\_registration\_and\_consents\_1.xml
* example\_3.1.2\_investigations\_1.xml

## Cancer Sample Tracking Data Specifications – CSV Submissions and examples:

1. **Cancer Sample Tracking-3.1.2.docx** - Cancer Sample Tracking Specification as a Word Document describing the csv message definitions for Genomic Medicine Centres (GMCs) to provide sample tracking data for the programme.
2. **Cancer Sample Tracking-3.1.2.xslx** - Cancer Sample Tracking Specification AND CHANGE LOG as an Excel Document describing the csv message definitions for Genomic Medicine Centres (GMCs) to provide sample tracking data for the programme. New CHANGES: Additions to the model are highlighted in Green, Removals are highlighted in Red, Superficial changes (such as descriptions) are highlighted in blue
3. **Example CSV** Files

The following csv files are included:

* example\_3.1.2\_metadata\_1

## Cancer Appendices: Reference Data and Guidance

1. **Cancer Model v3.1.2\_Appendix B\_Tissue\_Source\_Disease\_Type\_Matrix\_v1** - Disease Type / Subtype Matrix
2. **Cancer Model v3.1.2\_Appendix B\_Tissue\_Source\_Disease\_Type\_Matrix\_v2** - Tissue Source / Disease Type Matrix
3. **Cancer Model v3.1.2\_Appendix C\_Sample\_Type\_Test\_Result\_Type\_Matrix\_v1** - Sample Type / Test Result Type Matrix
4. **Cancer Model v3.1.2\_Appendix D\_Sample\_Metafile\_Completion\_Guidance\_v1** - Sample Metadata File Completion Guidance for each Sample Type
5. **Cancer Model v3.1.2\_Appendix E\_Test\_Result\_Type\_Test\_Result\_Value\_v1** - Test Result Type / Test Result Value
6. **Cancer Model v3.1.2\_Appendix F\_Consent\_Forms\_Participant\_Info\_Additional\_v1** - Consent Forms, Participant Information Sheets, Additional Materials and Consent Update Enumerations
7. **Cancer Model v3.1.2\_Appendix G\_Presentation\_v2** – Disease Type / Presentations

**NB**: More example xml files and csv files will be issued in due course.

# Known Deficiencies

## Inherent Defects

none

# Impact of Release

## GMC Impact

Valid processing of CSV files for Gear to against the published model.

Valid processing of XML submissions for Gear 2 against the published model.

## Biorepository Impact

none

## Sequencer Impact

none

## Support Impact

None anticipated beyond existing considerations.

The contact details for the Genomics England Contacts Desk are:

**Telephone: 0300 3 035 035**

**E-mail:** [**ssd.genomics@hscic.gov.uk**](mailto:ssd.genomics@hscic.gov.uk)

END OF RELEASE NOTE