ISO/TS 23357:2023-07 (E)

Genomics informatics - Clinical genomics data sharing specification for next-generation sequencing

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Terms and definitions 2

4 Abbreviated terms 4

5 Summary of the clinical genomic information model 4
5.1 General 4
5.2 Patient 5
5.2.1 General 5
5.2.2 Identifiers 5
5.2.3 Name 5
5.2.4 Sex 5
5.2.5 Birth data 5
5.2.6 Ethnicity 5
5.2.7 List of diagnosis 5
5.2.8 Treatment 6
5.3 Specimen 6
5.3.1 General 6
5.3.2 Tissue or organ of origin 6
5.3.3 Collection date 6
5.3.4 Type of specimen 7
5.4 Experimental equipment 7
5.4.1 General 7
5.4.2 Quality control 7
5.4.3 Base calling information 7
5.5 Analysis equipment 9
5.5.1 General 9
5.5.2 Read alignment 10
5.5.3 Alignment post processing 11
5.5.4 Variant calling 11
5.5.5 Variant annotation 12
5.6 Derived data 12
5.6.1 General 12
5.6.2 FASTAQ FASTQ 13
5.6.3 Sequence alignment map (SAM) 13
5.6.4 Binary alignment map (BAM) 13
5.6.5 Compressed reference-oriented alignment map (CRAM) 13
5.6.6 Variant cell format (VCF) 13
5.6.7 Mutation annotation format (MAF) 13

Bibliography 14