

团 体 标 准

T/SZAS 33—2021

严重急性呼吸系统综合征冠状病毒 2 (SARS-CoV-2) 核酸检测数据集

The data set of nucleic acid detection for severe acute respiratory syndrome
coronavirus 2 (SARS-CoV-2)

2021 - 06 - 04 发布

2021 - 06 - 19 实施

深圳市标准化协会 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 数据元目录	2
6 数据格式标准	2
7 数据归档目录	3
附录 A（资料性） 数据元目录.....	5
附录 B（资料性） 数据元值域代码表.....	9
附录 C（资料性） 检测结果数据样例.....	10

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由深圳华大医学检验实验室提出。

本文件由深圳市标准化协会归口。

本文件起草单位：深圳华大医学检验实验室、深圳市华大智造软件技术有限公司、深圳华大基因股份有限公司、北京大学深圳医院、惠州市第一妇幼保健院、惠州市第二妇幼保健院。

本文件主要起草人：刘小燕、钟宏彬、吕春杰、唐美芳、徐平、杜酉寅、吴锐锋、单日强、吴昊、李倩一、姜华艳。

严重急性呼吸系统综合征冠状病毒2（SARS-CoV-2） 核酸检测数据集

1 范围

本文件规定了严重急性呼吸系统综合征冠状病毒2（SARS-CoV-2）核酸检测数据的范围以及数据元的定义，数据集包括严重急性呼吸系统综合征冠状病毒2（SARS-CoV-2）核酸检测相关数据元 and 值域。

本文件适用于人的严重急性呼吸系统综合征冠状病毒2（SARS-CoV-2）核酸检测数据信息的存储、治理、交换与共享。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 2261.1 个人基本信息分类与代码 第1部分：人的性别代码

GB/T 2659 世界各国和地区名称代码

GB/T 18391.1-2009 信息技术 元数据注册系统(MDR) 第1部分：框架

GB/T 19495.5-2018 转基因产品检测 实时荧光定量聚合酶链式反应(PCR)检测方法

3 术语和定义

下列术语和定义适用于本文件。

3.1

实时荧光 PCR *real-time polymerase chain reaction; RT-PCR*

在聚合酶链式反应体系中加入荧光基团，利用荧光信号积累实时监控整个PCR进程，并通过标准曲线对未知模板进行定量分析的方法。

[来源：GB/T 19495.5-2018，3.1.3]

3.2

实时荧光核酸恒温扩增检测 *simultaneous amplification and testing; SAT*

将核酸恒温扩增技术和实时荧光检测技术相结合的一种核酸检测技术。

3.3

元数据 *metadata*

定义和描述其他数据的数据。

[来源：GB/T 18391.1-2009，3.2.16]

3.4

数据元 *data element; DE*

由一组属性规定其定义、标识、表示和允许值的数据单元。

[来源：GB/T 18391.1-2009，3.3.8]

4 缩略语

下列缩略语适用于本文件。

S: 字符串型 (string)

L: 布尔型 (boolean)

N: 数值型 (number)

DT: 日期时间型 (datetime)

5 数据元目录

5.1 数据元目录公用属性

数据元目录公用属性如表1所示。

表1 数据元目录公用属性

属性名称	描述
版本	V1.0
注册机构	注册机构名称
相关环境	生物信息、生物大数据
分类模式	分类法
主管机构	主管机构名称
注册状态	标准状态
提交机构	提交机构名称

5.2 数据元目录专用属性

5.2.1 严重急性呼吸系统综合征冠状病毒 2 (SARS-CoV-2) 核酸检测数据元目录专用属性包括个人信息元数据目录、表型信息元数据目录、送检信息元数据目录、检测过程信息元数据目录、检查结果信息元数据目录五部分。

5.2.2 个人信息元数据目录描述个人关键信息相关的数据元，例如姓名、性别代码、年龄(年)、年龄(月)、注册证件类型代码、注册证件号码等。

5.2.3 表型信息元数据目录描述受检者健康及表型的数据元，例如发病日期、就诊日期、就诊当日体温、呼吸道症状标志等。

5.2.4 送检信息元数据目录描述送检信息数据元，例如送检单位、送检科室、样本类型等。

5.2.5 检测过程信息元数据目录描述检测过程的数据元，例如检测技术名称、核酸提取试剂盒厂家名称、检测试剂盒厂家名称、检测者、审核者等。

5.2.6 检测结果信息元数据目录描述对数据进行解读形成最终检测结果的数据元，例如检测结果。

5.2.7 具体每个数据元的标识符、名称、定义、信息保护、单位、数据类型见附录 A。数据元允许值见附录 B。

6 数据格式标准

6.1 原始数据格式代码

原始数据格式代码如表 2 所示。

表2 原始数据格式

数据格式说明标识符	数据格式名称	规格说明
SF01.01.001	下机原始文件	具体格式按仪器自动生成的下机文件为准。

6.2 结果数据格式代码

结果数据格式代码如表3所示。

表3 结果数据格式

数据格式说明标识符	数据格式名称	规格说明
SF01.02.001	SARS-CoV-2 检测结果	见附录 C

7 数据归档目录

7.1 归档目录专用属性

7.1.1 归档目录结构

归档目录结构如表4所示。

表4 归档目录结构

第一级	第二级	第三级
Item_id	Raw_data	Rawdata_id
Item_id	Result_data	

注：每级目录情况说明见7.2。第一级目录和第二级目录为强制性目录，其他级目录为推荐性目录。

7.2 概述

本章节是对7.1.1归档目录结构具体每级目录的情况说明。共3级目录，每一级目录均包涵目录标识符、目录名称、目录定义及父目录。

7.2.1 数据归档第一级目录

目录标识符：DI01.01.01

目录名称：Item_id

定义：该目录用于存档SARS-CoV-2核酸检测相关数据。

7.2.2 数据归档第二级目录

目录标识符：DI01.02.01

目录名称：Raw_data

定义：该目录用于存档SARS-CoV-2核酸检测过程中检测仪器生成的原始数据。

父目录：DI01.01.01

目录标识符：DI01.02.02

目录名称：Result_data

定义：该目录用于存档SARS-CoV-2核酸检测的结果。

父目录：DI01.01.01

7.2.3 数据归档第三级目录

目录标识符：DI01.03.01

目录名称：Rawdata_id

定义：存放原始数据的指定文件。

父目录：DI01.02.01

附录 A
(资料性)
数据元目录

A.1 概述

本附录说明了推荐性数据元的标识符，名称，定义，信息保护，单位，数据类型和数据元允许值。且有新的数据元加入可以顺延排入。

A.2 个人信息元数据目录

个人信息如表A.1所示。

表A.1 个人信息

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE02.01.001.00	姓名	受检者在公安管理部门正式登记注册的姓氏和名称	保护		S	
DE02.01.002.00	性别代码	受检者生理性别在特定编码体系中的代码	不保护		S	GB/T 2261.1 个人基本信息分类与代码 第1部分:人的性别代码
DE02.01.003.00	年龄(年)	个体从出生当日公元纪年日起到计算当日止生存的时间长度,计量单位为岁。	不保护	岁	N	
DE02.01.004.00	年龄(月)	年龄不足1周岁的实足年龄的月龄,以分数形式表示:分数的整数部分代表实足月龄,分数部分分母为30,分子为不足1个月的天数	不保护		S	
DE02.01.005.00	注册证件类型代码	注册证件类型代码	保护		S	B.3 注册证件类型代码表
DE02.01.006.00	注册证件号码	注册证件号码	保护		S	
DE02.01.007.00	监护人注册证件类型代码	监护人注册证件类型代码	保护		S	B.3 注册证件类型代码表
DE02.01.008.00	监护人注册证件号码	监护人注册证件号码	保护		S	
DE02.01.009.00	现住社区名称	本人现住地址所属社区组织机构名称	不保护		S	
DE02.01.010.00	现住医院名称	本人现住医疗机构的组织机构名称	不保护		S	
DE02.01.011.00	现住地址-省(自治区、直辖市)	本人现住地址中的省、自治区或直辖市名称	不保护		S	
DE02.01.012.00	现住地址-市(地区、州)	本人现住地址中的市、地区或州的名称	不保护		S	

表A.1 个人信息 (续)

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE02.01.012.00	电话号码	联系电话的号码,包括国际、国内区号和分机号	保护		S	
DE02.01.013.00	发病前 14 天来自国家	发病前 2 周疫情发生地停留情况:来自国家的名称。	不保护		S	GB/T 2659 世界各国和地区名称代码表
DE02.01.014.00	发病前 14 天来自地区-省(自治区、直辖市)	发病前 2 周疫情发生地停留情况:来自地区的省、自治区或直辖市名称。	不保护		S	
DE02.01.015.00	发病前 14 天来自地区-市(地区、州)	发病前 2 周疫情发生地停留情况:来自地区的市、地区或州的名称。	不保护		S	
DE02.01.016.00	交通联合检疫(J)人员标志	是否是交通联合检疫(J)人员	不保护		L	
DE02.01.017.00	密切接触者(M)标志	是否属于确诊病例密切接触者(M)	不保护		L	
DE02.01.018.00	重点疫区人员(S)标志	是否是来自重点疫区人员(S)	不保护		L	
DE02.01.019.00	发热门诊患者(F)标志	是否是发热门诊患者(F)	不保护		L	
DE02.01.020.00	非发热门诊患者(N)标志	是否是非发热门诊患者(N)	不保护		L	
DE02.01.021.00	其他人员情况	不属于以上五类人员的其他人员情况	不保护		S	

A.3 表型信息元数据目录

表型信息如表A.2所示。

表A.2 表型信息

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE03.01.001.00	发病日期	疾病发病症状的首次出现当日的公元纪年日期和时间的完整描述	不保护		DT	
DE03.01.002.00	就诊日期	患者就诊当日的公元纪年日期和时间的完整描述	不保护		DT	
DE03.01.003.00	就诊当日体温	就诊当日体温的测量值,计量单位为℃	不保护	℃	N	
DE03.01.004.00	呼吸道症状标志	有无呼吸道症状	不保护		L	
DE03.01.005.00	消化道症状标志	有无消化道症状	不保护		L	

表A.2 表型信息（续）

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE03.01.006.00	第N次检测	第几次严重急性呼吸系统综合征冠状病毒2核酸检测。每次检测填写一条，N代表第几次检测的数值；当次检测进行多次复查的为一次检测。	保护		N	
DE03.01.007.00	第N次检测结果	每次检测填写一条，N代表第几次检测的值。第N次核酸检测结果在特定分类中的代码。	保护		S	B.1 定性结果代码表
DE03.01.008.00	第N次检测时间	每次检测填写一条，N代表第几次检测的值。第N次检测当日的公元纪年日期和时间的完整描述。	不保护		DT	
DE03.01.009.00	第N次检测地区	每次检测填写一条，N代表第几次检测的值。第N次检测地区完整地址。	保护		S	

A.4 送检信息元数据目录

送检信息如表A.3所示。

表A.3 送检信息

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE09.01.001.00	送检单位	样品的送检单位名称	不保护		S	
DE09.01.002.00	送检科室	医院送检的科室名称	不保护		S	
DE09.01.003.00	送检医师	送检医师在公安管理部门正式登记注册的姓氏和名称	保护		S	
DE09.01.004.00	采样日期	采集标本时的公元纪年日期和时间的完整描述	不保护		DT	
DE09.01.005.00	样本编号	样本唯一标识的编号	保护		N	
DE09.01.006.00	样本类型	采集样本的类型	不保护		S	

A.5 检测过程信息元数据目录

检测过程信息如表A.4所示。

表A.4 检测过程信息

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE01.01.001.00	检测技术名称	检测技术名称在特定分类中的代码	不保护		S	B.2 检测技术名称代码表
DE01.01.002.00	核酸提取试剂盒厂家名称	核酸提取试剂盒生产厂家的名称	不保护		S	
DE01.01.003.00	检测试剂盒厂家名称	检测试剂盒生产厂家的名称	不保护		S	
DE01.01.004.00	检测者	检测者在公安管理部门正式登记注册的姓氏和名称	保护		S	

表A.4 检测过程信息（续）

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE01.01.005.00	审核者	审核人在公安管理部门正式登记注册的姓氏和名称	保护		S	
DE01.01.006.00	报告日期	报告当日的公元纪年日期的完整描述	不保护		DT	
DE01.01.006.00	原始数据标识	该样本检测时对应的原始数据 ID	保护		S	

A.6 检查结果信息元数据目录

检查结果信息如表A.5所示。

表A.5 检测结果信息

标识符	名称	定义	信息保护	单位	数据类型	数据元允许值
DE05.06.001.00	检测结果	定性检测结果在特定分类中的代码	不保护		S	B.1 定性结果代码表

附录 B
(资料性)
数据元值域代码表

B.1 定性结果代码

定性结果代码规定了定性结果的代码。

采用1位数字顺序代码，从“1”开始编码，按升序排列，见表B.1。

表B.1 定性结果代码表

代码	定性结果
1	阴性
2	阳性
3	不确定

B.2 检测技术名称代码

检测技术名称代码规定了检测技术名称的代码。

采用1位数字顺序代码，从“1”开始编码，按升序排列，见表B.2。

表B.2 检测技术名称代码表

代码	检测技术名称
1	荧光定量 PCR
2	实时荧光核酸恒温扩增检测
9	其他

B.3 注册证件类型代码表

注册证件类型代码规定了注册证件类型代码。

采用1位数字顺序代码，从“1”开始编码，按升序排列，见表B.3。

表B.3 注册证件类型代码表

代码	注册证件类型
1	居民身份证
2	居民户口本
3	护照
4	军官证
5	驾驶证
6	港澳居民来往内地通行证
7	台湾居民来往内地通行证
99	其他法定有效证件

附录 C
(资料性)
检测结果数据样例

检测结果数据格式为文本格式，涵盖了5.2提及的SARS-CoV-2核酸检测数据的五部分元数据目录，详细信息可见附录A。样例可见表C.1。

表C.1 检测结果数据样例

个人信息	姓名
	性别代码
	年龄
	注册证件类型代码
	注册证件号码
	现住社区名称
	现住医院名称
	现住地址-省(自治区、直辖市)
	现住地址-市(地区、州)
	电话号码
	发病前 14 天来自国家
	发病前 14 天来自地区-省(自治区、直辖市)
	发病前 14 天来自地区-市(地区、州)
	交通联合检疫(J)人员标志
	密切接触者(M)标志
	重点疫区人员(S)标志
	发热门诊患者(F)标志
	非发热门诊患者(N)标志
	其他人员情况
表型信息	发病日期
	就诊日期
	就诊当日体温
	临床表现
	第 N 次检测
	第 N 次检测结果
	第 N 次检测时间
	第 N 次检测地区
送检信息	送检单位
	送检科室
	送检医师
	采样日期
	样本编号
	样本类型

表C.1 检测结果数据样例（续）

检测过程信息	检测技术名称
	核酸提取试剂盒厂家名称
	检测试剂盒厂家名称
	检测者
	审核者
	报告日期
	原始数据标识
检测结果信息	检测结果
