

T/BPMA

团 体 标 准

T/XXX.XX—2023

重大呼吸道传染病病例流调数据集

Dataset of epidemiological investigation for Severe Respiratory Infectious Disease

（征求意见稿）

目 录

前 言.....3

1 范围.....4

2 规范性引用文件.....4

3 术语和定义.....4

4 数据要素.....5

5 数据元属性.....5

6 数据格式.....6

参考文献.....13

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。

本文件由北京市疾病预防控制中心提出。

本文件由北京预防医学会归口。

文件起草单位：北京市疾病预防控制中心、北京市西城区智慧健康研究中心、京东城市（北京）数字科技有限公司。

本文件主要起草人：李刚、高燕琳、栗圆、杜婧、贾蕾、朱树宏、鲍捷、易修文。

本文件为首次发布。

重大呼吸道传染病病例流调数据集

1 范围

本标准规定了重大呼吸道传染病病例流调业务相关数据集的内容范围、数据结构、数据元属性及数据格式。

适用于指导重大呼吸道传染病病例流调的相关信息的采集、存储、共享以及信息系统的开发。

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 2260—2007 中华人民共和国行政区划代码

GB/T 2261.1—2003 个人基本信息分类与代码 第1部分：人的性别代码

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 38962—2020 个人健康信息码 数据格式

WS/T 370—2022 卫生健康信息基本数据集编制标准

WS 375.10-2012疾病控制基本数据集_第10部分：传染病报告

WS 537-2017 居民健康卡数据集

3 术语和定义

3.1

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。

[GB/T 18391.1—2009, 定义 3.3.8]

3.2

数据元标识符 data element identifier; DEI

为数据元分配的唯一标识。

[WS/T 370—2022, 内部标识7.3.1]

3.3

点位 point location; PL

点位是由经纬度定位和时间所确定的病例所在的位置，是病例活动轨迹的组成元素，多个点位的串联组成病例的活动轨迹。

4 数据要素

重大呼吸道传染病病例流调数据集主要包括人员基本信息、工作单位信息、核酸检测记录、疫苗接种状态、病例基本信息、病例活动轨迹、病例协查函、密接基本信息、流调报告等数据要素。

5 数据元属性

5.1 数据元属性设置

重大呼吸道传染病病例流调数据集的数据元由数据标识符、中文名称、数据类型、数据长度、数据值域、和数据约束等属性组成，详情如表1所示。

表 1 数据元属性

属性名称	描述
数据标识符	各数据元的唯一标识
中文名称	数据元的中文名称，应为唯一名称
英文名称	数据元的英文名称，应为唯一名称
数据类型	数据元的数据存储类型
数据长度	数据元的数据存储长度
数据值域	数据元的取值范围

5.2 数据元属性描述规则详情

5.2.1 数据元标识符

数据元标识符指数据元在某特定数据集中的唯一标识代码。采用长度16位的字母数字混合码，含3位“.”。结构见图1。

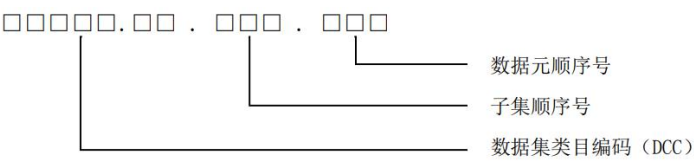


图 1 数据元标识符

a) 数据集类目编码（DCC）：用 8 位字母数字混合码表示。其中前用3位大写英文字母表示业务领域代码。卫生健康信息领域统一用 HDS 表示。第4位表示一级类目代码：用 1 位大写英文字母表示，从 A 开始顺序编码。第5位表示二级类目代码：用 1 位数字表示，数字大小无含义。如有两个以上类目（含两个），从 1 开始顺序编码。如无二级类目，编为 0。后面2位表示顺序号：用 2 位数字表示，代表二级类目下数据集的序号，数字大小无含义；从 01 开始顺序编码。顺序号与二级类目代码之间加“.”区分。

- b) 子集顺序号：用 3 位数字表示，数字大小无含义；从 001 开始顺序编码。如无子集，编为 000。子集顺序号与数据集类目编码之间加 “.” 区分。
- c) 数据元顺序号：用 3 位数字表示，代表数据元在数据集子集中的序号，数字大小无含义；从 001 开始顺序编码。数据元顺序号与子集顺序号之间加 “.” 区分。

5.2.2 中文名称

数据元的“中文名称”应当是唯一的，并且以字母、汉字、数字组成的字符串形式表示。中文名称应准确、通用、便于理解。

5.2.3 英文名称

数据元“英文名称”应当是与“数据元名称”对应且唯一的，以字母、数字式的字符串形式表示。英文名称的选用参考通用的英文术语及权威翻译。

5.2.4 数据类型

数据类型指数据元的数据存储类型，描述规则详见表2。

表 2 数据类型描述规则

数据类型	表示符	描述
字符型（string）	C	通过字符形式表达的值的类型。可包含字母字符（a~z,A~Z）、数字字符等。（默认 GB 2312）
布尔型（boolean）	B	又称逻辑型,采用 0（False）或 1（True）形式表示的逻辑值的类型。
数值型（number）	N	通过“0”到“9”数字形式表示的值的类型。
日期型（date）	D	采用 GB/T 7408 中规定的 YYYYMMDD 格式表示的值的类型。
日期时间型（datetime）	DT	采用 GB/T 7408 中规定的 YYYYMMDDThhmmss 格式表示的值的类型。（字符 T 作为时间的标志符，说明日的时间表示的开始。根据信息交换中各方的约定，在不会引起日期和时间组合的日的表示法与 GB/T7408 定义的其他表示法出现混淆的情况下，字符[T]可以省略。）
时间型（time）	T	采用 GB/T 7408 中规定的 hhmm ss 格式表示的值的类型。
二进制（binary）	BY	上述无法表示的其他数据类型，如图像、音频、视频等二进制流文件格式。

5.2.5 数据长度

数据长度指定了该数据元的最大存储长度，以数字形式表示。

5.2.6 数据值域

数据值域包括枚举值和非枚举值两类。若为可选枚举值，则需预先设置枚举值列表，如性别属性的枚举值为：男、女、未知；若为非枚举值，则需准确描述该值域的允许值。

6 数据格式

6.1 人员基本信息

人员基本信息主要包括病例的姓名、性别、出生日期、身高、体重、证件类型、证件号、联系方式、居住地址等，详细格式见表3。

表 3 人员基本信息

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.001.001	人员 ID	person_id	C	64	
HDSB1.01.001.002	姓名	name	C	64	
HDSB1.01.001.003	性别	gender	N	1	1 男；2 女
HDSB1.01.001.004	出生日期	birthday	D8		yyyyMMdd
HDSB1.01.001.005	身高	height	N	3	0~999
HDSB1.01.001.006	体重	weight	N	3	0~999
HDSB1.01.001.007	证件类型	certificate_type	N	1	1 身份证；2 护照；3 港澳居民来往内地通行证；4 台湾居民来往大陆通行证；5 外国人身份证；6 出生证明
HDSB1.01.001.008	证件号	certificate_number	C	255	
HDSB1.01.001.009	手机号码	mobile_number	C	11	
HDSB1.01.001.010	居住地址	residential_address	C	255	
HDSB1.01.001.011	居住地址所在区	residential_address_district	C	32	
HDSB1.01.001.012	居住地址所在街道	residential_address_street	C	32	
HDSB1.01.001.013	居住地址所在社区	residential_address_community	C	32	

6.2 工作单位信息

工作单位信息主要包括职业、单位名称、单位地址、单位联系人姓名、单位联系人电话等信息，详情如表4所示。

表 4 工作单位信息

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.002.001	人员编号	person_id	C	64	
HDSB1.01.002.002	职业	profession	C	32	
HDSB1.01.002.003	单位名称	company_name	C	64	
HDSB1.01.002.004	单位地址	company_address	C	255	
HDSB1.01.002.005	单位联系人姓名	company_contact_name	C	64	

HDSB1.01.002.006	单位联系人电话	company_contact_phone	C	11	
------------------	---------	-----------------------	---	----	--

6.3 核酸检测记录

核酸检测记录主要包括核酸的检测时间、检测状态、检测结果、检测机构名称、检测机构编码等，详情如表5所示。

表 5 核酸检测记录

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.003.001	核酸检测流水号	nucleic_id	C	128	
HDSB1.01.003.002	人员编号	person_id	C	64	
HDSB1.01.003.003	核酸检测结果	nucleic_detection_result	C	32	
HDSB1.01.003.004	核酸检测状态	nucleic_detection_status	C	32	
HDSB1.01.003.005	核酸检测时间	nucleic_detection_time	DT15		yyyyMMddhhffss
HDSB1.01.003.006	核酸检测机构编码	nucleic_detection_org_id	C	100	
HDSB1.01.003.007	核酸检测机构名称	nucleic_detection_org_name	C	128	

6.4 疫苗接种状态

疫苗接种记录主要包括接种的剂次、接种日期、疫苗种类、疫苗厂家、接种点等，详情如表6所示。

表 6 人员疫苗接种记录

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.004.001	疫苗接种流水号	vaccine_id	C	128	
HDSB1.01.004.002	人员编号	person_id	C	64	
HDSB1.01.004.003	接种剂次	vaccine_dose	N	1	0~9
HDSB1.01.004.004	接种日期	vaccine_date	D8		yyyyMMdd
HDSB1.01.004.005	疫苗种类	vaccine_type	C	64	
HDSB1.01.004.006	疫苗厂家	vaccine_manufacturer	C	64	
HDSB1.01.004.007	接种点	vaccine_area	C	128	

6.5 病例基本信息

病例基本信息主要包括病例类型、病例来源、毒株类型、是否感染、发现方式、病例最终状态、是否有症状，是否存在基础病等，详细格式见表7。

表7 病例基本信息

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.005.001	病例用户 ID	person_id	C	64	
HDSB1.01.005.002	病例类型	case_type	N	1	1 病例；2 提级管控
HDSB1.01.005.003	病例来源	case_source	N	1	1 本区录入；2 非本区录入
HDSB1.01.005.004	发现方式	find_mode	N	2	1 境外输入病例续发；2 本土病例续发；3 主动就诊
HDSB1.01.005.005	毒株是否 XBB	strain_category	N	1	1 是；2 否
HDSB1.01.005.007	是否本地	local_type	N	1	1 本土；2 境外
HDSB1.01.005.008	是否为政府机关、医务人员等	is_special_people	N	1	1 是；2 否
HDSB1.01.005.009	是否为学生	is_student	N	1	1 是；2 否
HDSB1.01.005.010	是否有症状	is_symptom	N	1	1 是；2 否
HDSB1.01.005.011	主要症状描述	symptom_detail	C	255	
HDSB1.01.005.012	是否存在基础病	is_basic_diseases	N	1	1 是；2 否
HDSB1.01.005.013	基础病详情	basic_diseases_detail	C	64	
HDSB1.01.005.014	是否感染	infect_status	N	1	1 是；2 否

6.6 病例活动轨迹

病例活动轨迹主要包括病例的活动轨迹涉及到的点位名称、点位地址、点位归属区、街道、社区、地点类型、地点类型备注、点位等级、点位类型、以及到访的开始时间和结束时间等，详情如表8所示。

表8 病例活动轨迹

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.006.001	病例用户 ID	person_id	C	64	
HDSB1.01.006.002	点位名称	point_name	C	128	
HDSB1.01.006.003	点位地址	point_address	C	255	
HDSB1.01.006.004	点位归属区	point_district	C	32	
HDSB1.01.006.005	点位归属街道	point_street	C	32	

HDSB1.01.006.006	点位归属社区	point_community	C	32	
HDSB1.01.006.007	经度	lng	C	32	
HDSB1.01.006.008	纬度	lat	C	32	
HDSB1.01.006.009	地点类型	location_type	C	16	
HDSB1.01.006.010	地点类型备注	location_remark	C	255	
HDSB1.01.006.011	点位等级	point_level	N	1	1 原始点位；2 被合并后的点位；3 无效的中间点位
HDSB1.01.006.012	点位类型	origin_type	N	1	1 原始点位；2 复合点位
HDSB1.01.006.013	开始时间	start_time	DT15		yyyyMMddhhffss
HDSB1.01.006.014	结束时间	end_time	DT15		yyyyMMddhhffss

6.7 密接基本信息

密接人员基础信息主要包括人员分类、关联的病例ID、接触方式、涉及场所、首次暴露日期、末次暴露日期等，详情如表9所示。

表 9 密接基本信息

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.007.001	风险人员 ID	person_id	C	64	
HDSB1.01.007.002	关联的病例 ID	related_case_id	C	64	
HDSB1.01.007.003	接触方式	contact_mode	C	64	
HDSB1.01.007.004	涉及场所	involve_place	C	64	
HDSB1.01.007.005	首次暴露日期	found_firt_day	D8		yyyyMMdd
HDSB1.01.007.006	末次暴露日期	found_last_day	D8		yyyyMMdd
HDSB1.01.007.007	管控口径	control_caliber	C	1024	
HDSB1.01.007.008	健康码类型	health_code_type	N	1	1 正常；2 红码；3 黄码

6.8 病例协查函

病例协查函主要包括函件名称、类型、日期、状态、内容、标注、发出时间、反馈时间、发送状态、接收人信息、以及关联的病例信息、风险行为、风险人员类型、协查方式等等，详情如表10所示。

表 10 病例协查函

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.008.001	函件编号	letter_number	C	64	
HDSB1.01.008.002	函件名称	letter_name	C	100	
HDSB1.01.008.003	函件类型	letter_type	N	1	1 协查函；2 协管函
HDSB1.01.008.004	函件标识	letter_flag	N	1	1 点位函件；2 病例函件；3 密接函件
HDSB1.01.008.005	函件标注	letter_tagging	C	64	
HDSB1.01.008.006	函件内容	letter_content	C	1024	
HDSB1.01.008.007	函件日期	letter_date	D8		yyyyMMdd
HDSB1.01.008.008	函件状态	letter_status	N	1	1 待提交；2 已提交；3 区审核通过；4 区驳回；5 市审核通过；6 市驳回；7 已接收；8 公安驳回
HDSB1.01.008.009	协查方式	assistant_investigation	N	1	1 健康宝；2 支付宝；3 微信支付；4 其他
HDSB1.01.008.010	接收人信息	accept_person	C	64	
HDSB1.01.008.011	关联病例姓名	case_name	C	64	
HDSB1.01.008.012	关联病例证件号码	case_certificate	C	255	
HDSB1.01.008.013	风险行为	risk_behavior_type	N	1	1 公交车；2 网约车；3 外卖；4 轨迹
HDSB1.01.008.015	发出时间	issue_time	DT15		yyyyMMddhhffss
HDSB1.01.008.016	反馈时间	feedback_time	DT15		yyyyMMddhhffss
HDSB1.01.008.017	备注	remark	C	1024	

6.9 流调报告

流调报告主要包括病例报告种类、标题、文件类型、文件说明、文件状态、存储桶名称、文件后缀格式、文件名称、大小、提交状态、区审批人、市审批人、提交时间、审核时间、通过时间等，详情如表11所示。

表 11 流调报告

数据标识符	中文名称	英文名称	数据类型	数据长度	数据值域
HDSB1.01.009.001	资源 ID	resource_id	C	64	
HDSB1.01.009.002	标题	report_title	C	255	
HDSB1.01.009.003	病例报告种类	report_type	N	1	1、4、8 小时报告

HDSB1.01.009.004	文件类型	document_type	C	10	
HDSB1.01.009.005	文件说明	document_description	C	255	
HDSB1.01.009.006	文件状态	file_status	N	1	0 未下载, 1 已下载, 2 已撤回
HDSB1.01.009.007	存储桶名称	file_bucket_name	C	255	
HDSB1.01.009.008	文件后缀格式	file_suffix	C	255	
HDSB1.01.009.009	文件名称	file_name	C	255	
HDSB1.01.009.010	文件大小	file_length	N	6	0~999999
HDSB1.01.009.011	提交状态	resource_status	N	1	1 待提交; 2 待区审核; 3 待市审核; 4 区返修; 5 市返修; 6 已通过
HDSB1.01.009.012	区审批人	region_approval_user	C	128	
HDSB1.01.009.013	市审批人	city_approval_user	C	128	
HDSB1.01.009.014	提交时间	upload_time	DT15		yyyyMMddhhffss
HDSB1.01.009.015	审核时间	examine_time	DT15		yyyyMMddhhffss
HDSB1.01.009.016	通过时间	pass_time	DT15		yyyyMMddhhffss

参考文献

- [1] GB/T 2260—2007 中华人民共和国行政区划代码
- [2] GB/T 2261.1—2003 个人基本信息分类与代码 第 1 部分:人的性别代码
- [3] GB/T 38962—2020 个人健康信息码 数据格式
- [4] GB/T 18391.1—2009 信息技术元数据注册系统(MDR) 第1部分:框架
- [5] WS537—2017 居民健康卡数据集
- [6] WS375.10—2012疾病控制基本数据集_第10部分: 传染病报告
- [7][8] YQFK A001—2021 疫情防控管理平台 数据格式