中国生物医学工程学会团体标准

《儿童神经母细胞瘤临床数据集规范》团体标准

（征求意见稿）编制说明

一．工作简况

（一）任务来源及编制目的

根据中国生物医学工程学会2025年第二批团体标准制修订项目工作计划，浙江大学医学院附属儿童医院等负责起草团体标准《儿童神经母细胞瘤临床数据集规范》（项目编号：BZW/PI-174）。

儿童神经母细胞瘤（Neuroblastoma, NB）临床科研数据分布于入院评估、病理诊断、影像检查、手术记录、化疗方案、放疗参数及长期随访等多个诊疗环节中。由于国内尚无针对NB临床科研数据集的国家标准或行业标准，不同机构在数据元定义、采集格式、标注规范及质量控制方面存在显著差异，导致多中心数据异构性强、完整性与准确性不足、互操作性缺失，严重制约了大规模数据分析、科研成果转化及人工智能辅助诊疗产品的研发，且难以支撑贴合中国儿童特点的创新药械研发与精准医疗实践。本标准围绕神经母细胞瘤临床科研数据的描述、采集、预处理、标注及质量控制，统一全周期数据集元数据、数据元属性及值域代码，为跨机构数据共享与协作研究、新药研发与临床决策支持系统构建、优质医疗资源下沉与同质化诊疗提供标准化支撑，最终提升我国在儿童肿瘤领域的国际竞争力与话语权。

（二）起草及协作单位

标准草案列明的起草单位为：浙江大学医学院附属儿童医院、四川大学华西第二医院、浙江大学、杭州聚仁医疗科技有限公司、山东大学齐鲁医院、上海交通大学医学院附属儿童医院、中国人民解放军总医院（301医院）、杭州深睿医疗有限公司、浙江大学滨江研究院。

（三）主要起草人及所做工作

标准草案列明的主要起草人为：俞刚、王金湖、朱珠、李晋辉、张爱军、熊洁妮、蔡嘉斌、章立锋、贺敏、茅君卿、李林洁、顾伟忠、赵曼丽、蔡海领。各起草人自接受起草任务以来，按照统一的时间进度共同参与了框架设计、资料调研、数据元梳理

与审核、值域代码编制、文本起草及试验验证等工作。

（四）主要工作过程

2025年3月，浙江大学医学院附属儿童医院向中国生物医学工程学会提出该团体标准的立项申请，2025年4月确认正式立项。由浙江大学医学院附属儿童医院、四川大学华西第二医院、浙江大学、山东大学齐鲁医院、上海交通大学医学院附属儿童医院、中国人民解放军总医院（301医院）、杭州深睿医疗有限公司、浙江大学滨江研究院、杭州聚仁医疗科技有限公司组成的标准起草工作组接到任务后，调查收集国内外相关标准和资料，确定本标准制订的范围和原则。起草小组在前期开展工作的基础上，2025年5月至2026年11月多次进行标准研讨，于2025年11月完成了标准草稿。为了推进标准验证工作，2026年12月至2026年6月期间，多次进行线上标准技术研讨，并于2022年7月完成标准验证，形成标准征求意见稿， 2022年8月完成征求意见，2022年9月形成送审稿。

二、主要技术内容的确定依据

（一）编制原则

本标准以临床科研实际需求为导向，兼顾神经母细胞瘤疾病异质性特点与卫生健康信息标准的协调性。在内容设计上，充分参考国内外高质量文献与行业规范，结合我国临床实际构建包含管理域与数据域的标准化体系，覆盖人口学、诊断分期、治疗干预、实验室检查、影像评估、不良反应及随访预后等核心要素；已有适用国家或行业代码体系优先引用，现有标准尚不能覆盖的 NB 专病内容采用本文件内部值域代码表，确保数据元语义明确、分类层次便于识别与维护。标准明确数据准确性、一致性、隐私脱敏及审核流程等质量控制要求，支持标准化字段与原始临床记录之间的追溯。

（二）编制依据及其作用

标准草案按照 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》确定文件结构和起草方式；以 WS/T 305《卫生信息数据集元数据规范》为数据集编制依据，以 WS/T 303《卫生信息数据元标准化规则》和 WS 363.1《卫生信息数据元目录 第 1 部分：总则》组织数据元命名、定义、数据类型、表示格式和允许值等属性；数据集分类与标识符按照 WS/T 306《卫生信息数据集分类与编码规则》确定，值域代码编制遵循 WS 364.1《卫生信息数据元值域代码 第 1 部分：总则》。

草案对人口学资料中的性别、民族、家庭关系及国籍分别引用 GB/T 2261.1、GB/T 3304、GB/T 4761 和 GB/T 2659.1 所列代码；对神经母细胞瘤诊断与手术操作相关数据元，分别采用 ICD-10 国际疾病分类标准编码和 ICD-9-CM-3 国际疾病分类手术代码进行规范化表达。对卫生健康信息数据元目录已覆盖的数据元，在数据元专用属性表中引用其数据元标识符（DE）；对目前尚无对应数据元标识符的神经母细胞瘤专病字段，由本文件设置内部标识符并纳入专用值域代码表管理。引用版本、正式标识符及具体适用范围应在送审前逐项核对。

（三）适用范围与数据集元数据

第 1 章将适用范围确定为儿童神经母细胞瘤临床科研数据集的元数据属性、数据元公用属性、数据元专用属性、值域代码表及数据质量控制要求，适用于指导该数据集的数据元编制、字段映射、结构化采集、数据交换共享及相关信息系统建设。该范围与标准规定的元数据、数据元属性和值域代码相对应，有利于明确标准实施对象和使用边界，确保数据采集与利用的规范性。

第 5 章设置数据集元数据属性，以表格形式描述数据集名称、数据集标识符、数据集发布方-单位名称、关键词、数据集语种、数据集分类-类目名称、数据集摘要和数据集特征数据元等元数据项，用于识别数据集、说明内容范围及支持检索管理。其中数据集标识符、发布方等管理信息应按照正式发布和注册要求确定。同时，标准在第 3 章和第 4 章对神经母细胞瘤、数据质量等核心术语及 TNM 分期、数据类型等缩略语进行了统一界定，为数据集元数据的准确理解和规范应用提供基础支撑。

（四）数据模块及数据元设置依据

第5章按照儿童神经母细胞瘤患儿从入院诊治到出院随访的全周期临床科研业务流程划分十二类数据模块，以“规范性附录 A. 儿童神经母细胞瘤临床数据元专用属性”中表 A.1 至表 A.23 逐类列出数据元专用属性。数据模块的设置覆盖了人口学资料、个人家族及既往疾病情况、一般状况及体格检查、神经母细胞瘤诊断、既往治疗史、干预措施、合并用药、实验室标本采集及检验、医学检查、疗效评价、不良反应及随访预后等核心要素，确保数据集能够完整支撑多中心临床研究、科研成果转化及人工智能辅助诊疗产品研发的需求。在值域管理方面，表 A.24 规定了数据元值域代码表编码，表 A.25 至表 A.63 详细列出了具体的数据元值域代码表，以规范各数据元的允许值集合。经对草案数据元专用属性表逐行统计，共有 432 项数据元；下表为当

前草案的数据元条目统计。

数据模块	规范章节	数据表号	条目数
人口学资料	5.2.1	表 A.1~A.2	14
个人家族及既往疾病情况	5.2.2	表 A.3~A.4	17
一般状况及体格检查	5.2.3	表 A.5~A.6	25
神经母细胞瘤诊断	5.2.4	表 A.7~A.8	53
神经母细胞瘤既往治疗史	5.2.5	表 A.9～表 A.10	45
干预措施	5.2.6	表 A.11～表 A.12	66
合并用药	5.2.7	表 A.13～表 A.14	24
实验室标本采集及检验	5.2.8	表 A.15～表 A.16	89
医学检查	5.2.9	表 A.17～表 A.18	39
疗效评价	5.2.10	表 A.19～表 A.20	23
不良反应	5.2.11	表 A.21～表 A.22	15
随访及预后	5.2.12	表 A.23	22
合计	—	—	432

患者人口学资料、个人家族及既往疾病情况模块用于描述患儿身份背景、家族遗传倾向、过敏史及既往健康基线，确保入组数据真实有效，为后续研究提供基础背景信息。一般状况及体格检查模块覆盖体能状态评分（ECOG、KPS）、生命体征及浅表淋巴结情况等数据元，为患儿基础健康状态评估提供依据。神经母细胞瘤诊断模块覆盖诊断编码与依据、病理镜下所见、免疫组化与基因检测结果，以及 INSS、INRGSS 病理分期和危险度分组等数据元，为疾病的精准诊断与风险分层提供核心数据。

既往治疗史与干预措施模块记录既往及当前的手术细节、放疗参数、系统治疗与试验性治疗方案等数据元，用于全面还原患儿的治疗轨迹与临床干预过程。合并用药模块覆盖伴随用药详情、用药原因、合并医疗操作及方案偏移情况等数据元，用于评估治疗依从性与药物相互作用。实验室标本采集及检验、医学检查模块记录生物样本采集信息、各类检验指标定性定量结果，以及心电图、CT、MR、PET-CT 和超声等影像学检查详情，为病情监测提供客观辅助依据。

疗效评价模块覆盖评价标准、总疗效结果以及靶病灶、非靶病灶和新发病灶的具体评估数据元，用于科学量化治疗响应。不良反应模块记录不良事件详情、分级、因果关系及严重不良事件等数据元，为干预措施的安全性评价提供支撑。随访及预后模块覆盖访视记录、生存状态、后续抗肿瘤治疗、死亡或退出原因，以及肿瘤复发与远

处转移情况等数据元，用于长期追踪患儿预后结局与生存质量。

（五）数据元属性与内部标识符

第 5 章区分数据元公用属性和专用属性。数据元公用属性统一描述版本、注册机构、相关环境、分类模式、主管机构、注册状态和提交机构等管理信息；专用属性明确内部标识符、数据元标识符（DE）、数据元名称、定义、数据元值的数据类型、表示格式及数据元允许值，以支持不同机构和系统对同一字段形成一致理解。

数据元内部标识符采用“ED.XX.YY.NNN.ZZ”结构，其中 ED 为数据元标识前缀，XX 为类别代码，YY 为亚类别代码，NNN 为三位顺序代码，ZZ 为保留代码（未进一步细分时以“00”表示）。例如，CH.03.ZD.00.0003 对应“诊断依据”，CH.03.ZD.00.0005 对应“首次诊断日期”。该结构在保持唯一识别的同时反映数据元的类别归属，便于字段映射、扩展和版本维护；对卫生健康信息标准已有数据元，另以数据元标识符（DE）建立对应关系。

数据元值的数据类型按照字符型（S1、S3）、数值型（N）、日期型（D）和字节流型（BY）等类别设置，并区分自由文本、引用代码表和直接取值三种字符型表达方式。数值型数据元以表示格式约定整数位数和小数位数，日期型数据元以 D8 等格式约定表示精度，字节流型数据元用于描述波形、影像等阵列数据。数据元允许值为“—”时表示不另设固定值域，不应将破折号本身作为业务数据写入系统。

（六）值域代码表及现有代码体系的衔接

本文件规范性附录 A 对现有国家或行业标准能够覆盖的值域直接引用，对儿童神经母细胞瘤临床科研专有且现有标准尚不能覆盖的部分设置专用值域代码表。草案在“附录 A. 儿童神经母细胞瘤临床数据元专用属性”中，通过表 A.24 规定数据元值域代码表编码规则，并编制了 39 个专用值域代码表，对应表 A.25 至表 A.63。

专用值域涵盖临床研究状态、肿瘤患者体力状况 ECOG 评分与功能状态 KPS 评分、各年龄段儿童正常血压参考值、神经母细胞瘤发生位置及病理活检部位、术后病理类型、INSS 与 INRG 分期、危险度分组、远处转移部位、手术与放射治疗目的及技术、系统治疗药物类型、靶病灶评估方法、不良事件分级与转归、中止试验性治疗及受试者退出研究原因等核心内容。上述信息需在多中心临床研究、跨机构数据交换及长期随访中反复使用，通过固定代码可有效减少自由文本歧义，保障数据的一致性与可比性。

值域代码表以表 A.25 至表 A.63 编号，通过表号与数据元专用属性表中的“允许值”栏目相呼应，使“数据元—允许值—代码含义”之间可相互核对。该表号仅用于本文件内部的引用和索引，不占用 WS/T 363 系列数据元标识符（DE）的正式标识符号段。

同一代码表原则上采用固定位数的数字代码，需要设置“其他”类目时按草案规则设置收容代码，并结合说明字段保留具体内容。对性别、民族、医保类型等现有代码体系能够覆盖的字段，保留正式代码及其语义，避免因本地重新编号破坏交换一致性；对神经母细胞瘤特有的分期、危险度分组、病理类型等专业内容，则依据国际权威指南与临床共识制定专用代码，确保数据集既符合通用数据交换规范，又满足专科科研的精细化需求。

（七）技术指标与新旧标准对比说明

本标准主要技术要求体现在儿童神经母细胞瘤临床科研数据元的语义定义、数据类型、表示格式、值域编码及模块间引用关系，属于数据集结构规范，不涉及诊疗技术或医疗器械性能指标。诊断分期、危险度分组、疗效评价及不良反应等字段用于结构化记录临床科研全流程产生的真实世界数据；其类型与格式应能够准确容纳多中心研究中的原始采集结果。涉及 INSS/INRGSS 分期、ECOG/KPS 评分、不良事件分级等专业内容的定义与取值范围，均以国际权威指南、临床共识及现行国家卫生信息标准为明确依据，确保数据元在不同研究机构间的语义一致性与互操作性，具体实施要求应通过配套数据采集方案或电子病例报告表（eCRF）设计予以明确。

现草案未列明被代替标准，暂按制定项目编写本说明。若正式立项属于修订项目，应补充旧版标准编号、技术条款变化及新旧水平对比；在立项性质确认前，不作“全面优于现有标准”等结论。

三、主要验证分析、技术经济论证及预期效果

（一）验证单位、内容与方法

2025 年 12 月至 2026 年 1 月，由主要起草单位牵头，联合杭州聚仁医疗科技有限公司、杭州深睿医疗科技有限公司、深圳扬奇医芯智能科技有限公司等多家机构共同开展验证。录 A 儿童神经母细胞瘤临床数据元专用属性中的“数据元属性”和“数据元值域代码表”，确认相关条款的可理解性、可操作性和可实施性。验证对象为标准草案文本及其数据元属性表、值域代码表，不涉及实物样品。

验证支撑材料包括标准草案文本、数据元属性表、值域代码表，以及来自多中心的电子病历、医嘱、LIS 检验、影像报告及随访记录等典型字段样例。验证工作依托各参与单位的实际运行环境，主要使用脱敏字段、标准草案文本和专家复核记录；电子表格、文档编辑工具和字段映射工具用于整理验证证据，不作为标准符合性判定对象。验证以临床数据与标准数据元之间的映射核查为核心，对数据元名称、定义、数据类型、表示格式、计量单位及允许值域进行逐项比对。杭州深睿医疗科技有限公司、深圳扬奇医芯智能科技有限公司协助开发符合标准的 eCRF 模板与校验规则。结果分为“符合”和“不符合”：能够找到对应证据且支持标准化采集、映射和复用的判定为“符合”；缺失必要内容或存在实质性冲突的判定为“不符合”。

（二）验证结果及结论

验证报告共列出 15 项验证项目，覆盖人口学资料、个人家族及既往疾病情况、一般状况及体格检查、神经母细胞瘤诊断、既往治疗史、干预措施、合并用药、实验室标本采集及检验、医学检查、疗效评价、不良反应、随访及预后十二类数据元专用属性，以及值域代码表编码规则与具体代码表，单项结论均为“符合”。报告附录按对应标准条款列明验证资料、材料来源、验证过程记录和评价结论，为各项判定提供支撑。

根据验证报告的综合评估，标准草案总体结构合理，数据元体系完整覆盖儿童神经母细胞瘤从初诊、治疗到长期随访的全过程关键节点，能够满足流行病学研究、危险度分层、个体化治疗及新药临床试验等核心科研需求。通过多中心真实世界数据的映射核查，验证了标准在解决“同义异名”“分期表述不一致”等现实问题上的有效性，确认其能够支撑跨机构数据整合与国际协作项目（如 INRG、COG）的数据交换要求。经检验，《儿童神经母细胞瘤临床数据集规范》草案符合要求。

本次验证属于数据集规范的文本验证和字段映射验证，不对具体临床诊疗流程、数据库物理存储结构、数据采集软件及数据共享审批流程进行验证。字段映射、典型病例对照及专家复核仅作为标准条款可实施性的举证材料，不作为标准正文新增要求。

（三）技术经济论证及预期效果

本标准主要通过数据字典配置、字段映射、编码转换和采集校验等方式实施，可结合既有电子病历系统与科研平台开展适配。主要投入包括临床与信息人员协同、历

史数据清洗、eCRF 模板开发、系统改造、多中心验证及培训。实际工作量取决于各机构的数据基础、信息系统条件及需迁移的历史记录规模。

实施后预期可减少同一字段的重复定义及跨机构、跨系统交换中的反复解释，降低数据清洗和编码转换工作量，提高多中心神经母细胞瘤临床科研数据的可比性与追溯能力。通过统一诊断分期、疗效评价及不良事件分级等核心数据元的语义与格式，可为国家级注册研究平台建设、国际多中心协作及长期生存与生活质量关联研究提供更稳定、更高质量的数据基础，助力儿童神经母细胞瘤精准诊疗与科研创新。

四、采用国际标准的程度及水平

未采用或参考国际和国外先进标准。

五、重大分歧意见的处理经过和依据。

无。

六、贯彻标准的要求和措施建议

（一）组织保障与宣贯培训

标准正式发布后，建议牵头单位联合各参与机构，面向临床医师、科研人员、数据管理员及信息化工程师等核心用户群体开展系统性宣贯培训。培训内容应聚焦十二类数据模块的业务边界、数据元语义内涵、专用值域代码表的规范使用，以及历史数据映射中的常见难点与解决方案。各实施单位宜设立专项工作组，明确业务统筹负责人、数据字典维护专员及数据质量核查员，建立常态化的问题收集、反馈响应与标准迭代更新机制，确保标准落地过程中的疑问能够及时得到权威解答与处理。

（二）技术实施与数据治理

实施单位应构建本地信息系统字段与标准数据元的映射关系表，详细记录源系统、原始字段名、目标数据元标识符、转换逻辑、计量单位及适用前提，并完整保留原始值与转换过程日志以备追溯。对于代码型数据元，须严格采用本标准附录 A 规定的值域代码表；数值型数据元应配置合理的取值范围与单位校验规则；日期时间型数据元需统一格式与时区基准，严禁以默认值填充未知或缺失信息，应通过专用标识予以标注。

针对神经母细胞瘤诊疗过程中重复发生的检查、治疗及随访事件，应结合患者唯一标识、就诊流水号、时间节点及事件关联关系进行结构化组织，避免多次记录相互

覆盖或混淆。涉及危险度分组、疗效评价等复合指标时，宜同步记录所依据的分期系统版本、评估标准及计算逻辑，确保结果可解释、可复核。具体技术实施细则应通过配套的数据采集方案或电子病例报告表（eCRF）设计指南予以明确。

在数据采集、汇聚与科研应用全生命周期中，应遵循最小必要原则配置数据访问权限，严格落实受试者隐私保护、操作审计留痕、数据备份及使用审批等安全管理措施。跨机构数据共享、国际合作研究或公开发布等行为，须另行履行伦理审查、数据安全评估及合规授权程序，本标准仅规定数据结构规范，不替代相关法律法规及管理制度的要求。

（三）持续维护与效果评价

建议建立基于实际应用反馈的标准动态评估机制，定期检视数据元覆盖完整性、语义一致性、代码有效性及跨系统交换可用性。当新增、修订或废弃数据元及值域代码时，应完整记录版本号、变更原因、影响范围及向后兼容策略，严禁将已分配的标识符重新赋予不同含义，防止破坏已有数据的语义稳定性。凡涉及标准核心技术内容调整的，应按照团体标准管理规定履行正式的维护或修订程序，确保标准持续适应儿童神经母细胞瘤临床科研的发展需求与技术演进。

七、其他应予说明的事项。

无。

儿童神经母细胞瘤临床数据集规范标准起草组

2026 年 09 月 20 日

CSBME