Journal of the Society for Clinical Data Management

Sen A 等。儿科临床数据标准化：conect4children (c4c) 跨领域儿科数据字典的开发。《临床数据管理学会期刊》。2022年；2(3)：4，第1–10页。DOI：https://doi.org/10.47912/jscdm.218

原创研究

儿科临床数据标准化：conect4children (c4c) 跨领域儿科数据字典的开发

Anando Sen[*]、Victoria Hedley[*]、John Owen[†]、Ronald Cornet[‡]、Dipak Kalra[§]、Corinna Engel[‖]、Avril Palmeri[*]、Joanne Lee[*]、Jeane Christophe Roze[¶]、Joseph Standing[**]、Adilia Warris[††]、Claudia Pansieri[‡‡]、Rebecca Leary[*]、Mark Turner[§§] 和 Volker Straub[*]

引言： 儿科临床试验中所收集数据项的标准化是一项重要但颇具挑战的工作。制药行业对CDISC数据标准已有充分了解，但其中缺少一些儿科特异性概念的实施方案。当CDISC标准中缺少某项儿科概念时，制药公司和研究机构会采用多种方式收集儿科数据，导致标准的实施方式各不相同，也可能限制数据的再利用价值。
目的： 为克服这些挑战，conect4children联盟开发了跨领域儿科数据字典（CCPDD）。

方法： 该字典分三个阶段建立：确定范围（包括向10家行业合作伙伴和34家学术合作伙伴发送调查，以了解他们的兴趣）、创建长名单，以及为最终术语集建立共识。字典在一场研讨会上最终定稿，参会者来自学术界、医院、行业和CDISC。参会者针对每个数据项进行了详细讨论，并参与了关于是否将该数据项纳入CCPDD的最终投票。
结果： 9家行业合作伙伴和全部34家学术合作伙伴回应了调查，结果显示各方普遍对开发CCPDD感兴趣。对27个数据项进行最终投票后，3项遭到否决，6项推迟至下一版本处理，并征求了CDISC的最终意见。CCPDD首个版本包含25个数据项，于2019年8月发布。

讨论与结论： 持续使用该字典有望确保收集到标准化且可互操作的数据，之后还可汇总并重复用于其他应用。该字典已用于三项临床试验的病例报告表创建。CCPDD还将作为《儿科用户指南》的输入之一；该指南由CDISC开发。

关键词： 数据字典；儿科临床试验；数据标准化；数据标准

引言

儿科临床试验面临重大挑战，包括年龄特异性制剂和剂量、受特定疾病影响的脆弱且往往规模较小的人群、漫长的监管挑战、伦理考量，以及父母或监护人对让儿童入组的犹豫。[1–6] 汇集来自不同来源的临床数据，可形成大型队列，用于事后研究，以及研究之间有意义的比较。[7,8] 然而，针对不同目的从不同来源获取的数据，其互操作性（语义层面和语法层面）仍然是一个问题。[9,10] 例如，口腔、额头或腋下测得的体温不能视为同一项测量。因此，儿科数据的标准化是一项重要的科学挑战。一项由资助支持的儿科工作

* 英国泰恩河畔纽卡斯尔纽卡斯尔大学
† 临床数据交换标准协会（CDISC）欧洲基金会，比利时布鲁塞尔
‡ 阿姆斯特丹大学医学中心，阿姆斯特丹，荷兰§ 欧洲健康数据创新研究所，比利时根特
‖ 大学儿童医院，德国图宾根
¶ 南特大学，南特，法国 ** 伦敦大学学院，伦敦，英国
†† 埃克塞特大学，埃克塞特，英国
‡‡ 生物学与药理学评估联盟，帕维亚，意大利§§ 利物浦大学，利物浦，英国 通讯作者：Rebecca Leary (becca.leary@newcastle.ac.uk)

小组[11] 在 Health Level Seven (HL7) 内着手定义一组数据元素，以支持临床护理和研究，但由于缺乏持续资金，相关工作未能进入投票阶段。此外，若干由美国国立卫生研究院 (NIH) 资助的儿科临床试验网络正在开展工作，但并未致力于数据标准的发育。[12,13]

目前，对于如何规范儿科临床研究中许多儿科特定数据项的收集，尚未形成共识。临床数据交换标准协会（CDISC）制定的标准已为制药行业广泛了解。美国食品药品监督管理局（FDA）和日本药品和医疗器械管理局（PMDA）要求在电子监管申报中使用这些标准，中国国家药品监督管理局也倾向于在申报中采用这些标准，这体现了它们的价值。[14–16] 然而，CDISC标准目前存在两个主要缺点：1）虽然可以利用CDISC基础标准，但其中缺少儿科临床试验申办方所需的特定儿科概念；[17,18] 2）CDISC标准在学术临床研究领域的普及程度较低。[19]

在没有CDISC标准（或现有标准缺少实施示例）的情况下，制药公司会采取多种不同做法，包括采用公司特定标准（通常遵循CDISC的总体结构），也可以直接与CDISC合作，将这些标准纳入CDISC标准。因此，每个试验申办方基本上都会创建自己的数据字典，其中可能包含研究特定或机构特定标准。[20] 这些数据字典具有专有性质，通常不会共享。缺乏标准化会导致数据表示方式不一致，使互操作性（即以有意义的方式汇集和共享数据）变得非常困难，有时甚至不可能。此外，缺乏数据标准会降低试验执行效率，因为各组织在开发新的数据收集工具和标准时都必须从头开始。

这些问题令人担忧，因为在可查找、可访问、可互操作、可重用（FAIR）原则的指引下，[21] 临床研究正日益迈入数据重用与互操作的时代。2016年发布的数据管理FAIR原则，基于计算机系统能够以最少的人为干预利用数据开展研究这一能力。除儿科学领域外，罕见病领域也充分认识到标准化数据的必要性（根据欧盟委员会的定义，罕见病是指患病率低于1/2000、需要采取特别协同措施应对的严重疾病）。通过利用既往数据获取的知识，可以提高诊断的速度和准确性，加深对疾病进程的了解，并改进潜在治疗方案。罕见病对儿科学十分重要，因为大多数罕见病会影响儿童，或仅在儿童期发病。鉴于罕见病数据十分稀缺，汇集和共享数据至关重要。[22,23]

背景

为应对上述问题，conect4children（c4c）联盟于2018年成立。c4c的主要目标是解决有效开展和实施儿科临床试验所面临的诸多障碍。[24] c4c是一个有期限的公私合作联盟，由创新药物倡议（IMI）资助。IMI项目由欧盟委员会和制药行业通过其欧洲协会——欧洲制药工业和协会联合会（EFPIA）共同出资。c4c的研究领域划分为八个工作包。工作包五专注于数据标准化，其目标概述于**框1**。

数据标准化工作包的首批工作之一是开发跨领域儿科数据字典（CCPDD）。鉴于该字典旨在作为CDISC儿科用户指南（PUG）的输入，CDISC全程参与了字典的开发。这确保字典一经发布，便可立即用于儿科临床试验病例报告表（CRF）的标准化。该字典重点关注临床研究中常规收集的跨领域（或与疾病无关）数据项。本文介绍了用于编制该字典的严谨共识形成方法，并讨论了存在争议的数据项。最后，本文探讨了下一版字典及PUG扩展版的改进方向。

框1： conect4children工作包5愿景概要。

conect4children工作包5目标概要

· 标准化临床试验数据和元数据的增加，将推动有关儿科疾病的科学知识增长。这些知识将助力开发更安全的儿童用药。

· 高质量的数据驱动的知识有望推动更优质、更高效的儿科临床试验设计。临床试验将能够充分利用标准化数据的力量，采用更具创新性的方法开展儿科药物开发（例如，将相同数据再利用于多个应用，以及从此前获取的数据中外推研究结果）。

· 儿童及其家人因必须参加多项临床试验（其中许多试验重复或没有必要）而承受的负担将会减轻。[25,26]

· 基于 FAIR数据的科学将扩大其影响力，并提高社会对儿童研究的接受度。[27]

· 标准化数据有助于开展基于患者个体层面而非汇总数据的荟萃分析。

材料与方法

c4c CCPDD 的第一版历时约 16 个月（2018年5月至2019年8月）完成，包含三个主要阶段：准备（范围界定）、创建数据项长名单以及共识建立。

第1阶段——准备（范围界定）c4c调查

为了解行业和学术合作伙伴对制定 CCPDD、以标准化病例报告表的兴趣，纽卡斯尔大学 c4c 团队设计了一项简短调查。大多数问题为多项选择题，部分问题则提供了填写文字反馈的选项。该调查通过 Lime Survey（https://www.limesurvey.org/）在线创建，并发送给 c4c 的十家制药行业合作伙伴。类似的调查也发送给了34家学术合作伙伴和非 EFPIA 受益者。调查空白副本作为补充文件提供。需要指出的是，调查中的部分问题超出了 CCPDD 的范围。

CCPDD工作组的成立

调查获得积极反馈后（见“结果”），成立了一个由16名代表组成的小型工作组（WG），成员包括儿科医生、学术界人士、制药行业代表、数据标准专家和数据科学家。该小组负责与 c4c CCPDD 相关的所有战略决策。

定义范围和目的

在确定纳入字典的数据项之前，WG 先明确了其目的和范围。出于实际考虑，并为最大限度地发挥重复使用和增值的潜力，WG 同意该数据字典应着重规定在多项儿科临床试验及多个疾病领域中通常会采集的数据项。同时，大家还同意重点关注那些针对儿童可能采用特定方式采集的数据项，而不是那些在所有年龄段均以相似且统一方式采集的数据项。经决定，该字典不会强制规定采集数据项所 采用的方法（例如，不规定血压必须在坐位时测量）。不过，字典将强烈建议在病例报告表中采集限定词。这些限定词可以是方法限定词（例如坐位、卧位或站立位测量血压等）、时间限定词（例如上午、下午等），或任何其他测量相关限定词。

CCPDD工作组和c4c项目领导团队商定了以下数据项遴选标准：(1) CCPDD中的所有数据项都必须与儿科学相关；(2) 数据项必须具有跨领域性质，即不能针对特定疾病；(3) 数据项应为儿科临床试验中常规收集的数据项，其中可包括仅与某一年龄组相关的数据项（例如，青少年青春期相关数据项或与新生儿出生相关的数据项），但这些数据项不会在其他年龄组的试验中收集；(4) 临床实践中常规收集、但未用于临床试验的数据项将视为超出范围；(5) 对于已有CDISC标准的数据项，将参考相应标准。

第 2 阶段：生成数据项长名单

长名单的生成过程并不正式，相关术语收集自下列不同来源。某个术语是否纳入 CCPDD，最终由第 3 阶段决定。

1. 病例报告表及d data dictionari……提供者ided b……成员 b……者 f c4c联盟——对病例报告表进行了审查，并记录了所有与疾病无关的数据项。所使用的病例报告表和数据字典示例包括拜耳全球标准病例报告表页面和血友病标准病例报告表页面、诺华Tanner分期和生命体征病例报告表，以及TREAT-NMD全球SMA/DMD登记数据字典中的基线报告、实验室数据、物理治疗和访视数据字典。2. 来自公开可用来源的数据项——在clinicalstudydata-request.com、yoda.yale.edu 和 vivli.org 上进行了关键词检索。关键词包括儿科学、新生儿、新生儿期、青少年、儿童、婴儿、儿童、婴儿、新生儿和青春期。使用clinicaltrials.gov 的高级搜索功能检索儿科研究（出生至17岁）。对试验描述、试验组、入选标准和结局指标中的文本进行了扫描，并记录了与疾病无关的数据项。3. 行业和学术合作伙伴建议的数据项——由于专有权限制，部分合作伙伴无法与c4c分享病例报告表或数据字典。因此，一些参与者提供了供考虑的数据项列表。这些数据项来自其所在机构或临床试验单位开展的真实儿科试验和研究。4. PubMed文献检索——在PubMed上开展了文献检索。对报告儿科研究结果的论文进行了分析，并记录了所有跨研究领域的数据项，以便加入长名单。具有疾病特异性的数据项未被纳入。

第 3 阶段 – 共识建立

工作组成员审阅了长名单中的数据项，并在2019年4月于葡萄牙里斯本举行的两场研讨会中的首场上进一步修订。此次研讨会的主要任务是确定哪些数据项属于跨领域数据项。通常，某个术语若出现在80%的病例报告表中，就被视为跨领域术语。然而，由于1）存在含义相似但并非同义的术语（例如，性别与表型性别），以及2）生成长名单时使用了病例报告表以外的来源（例如 PubMed、公开来源等），因此这一标准在实际应用中并非简单直接。会议还决定尽可能遵循CDISC标准。在研讨会期间及之后的一系列电话会议中，长名单经由共识逐步缩短。凡是未能达成共识的术语，均保留至下一阶段。

此外，还为后续步骤作出了一些决定。此前按年龄（新生儿、儿童和青少年）划分的字典，改为按数据项类别划分：人口统计学信息、生命体征、青春期状态及其他类别。第 2 阶段中，公司需审阅的大量材料使得无法逐一考量每个已确定的内容来源。为降低遗漏重要概念的可能性，研究团队在第 3 阶段再次联系相关公司，请其审阅长名单，查找遗漏的跨领域概念。所有建议的术语均被加入名单。此次研讨会的另一项重要举措是为 CCPDD 中的每个数据项附上使用指导意见，以帮助 c4c 研究团队提高其研究中所收集数据的协调统一程度。不过，会议决定不强制遵循这些指导意见。会议计划于2019年6月举办后续研讨会，专门完成 CCPDD 的定稿。

第二次研讨会在英国纽卡斯尔举行。除学术界和行业参会者外，本次研讨会还邀请了 CDISC 的代表，因为研讨会目标之一是敲定 CCPDD，并将其作为 CDISC PUG 的输入。参会者共 25 人（19 人现场参会，6 人线上参会）。研讨会第一部分向与会者提供了当前的数据项清单，该清单此前已经过多轮审查。本次审查更为正式，研讨会与会者受邀讨论数据项的四个方面。预计仍有一些跨领域数据项尚未由 CDISC 建模，或其现有定义需要进行大量工作才能适用于儿科。审查的四个方面是

1. **范围**：该项目是否真正具有跨领域适用性，且是否在儿科试验中常规收集？2. **单位：**该项目适用的单位是什么（如适用）？

3. **限定符：**需要哪些时间或方法限定词（例如，测量血压时患者的体位、温度采集所用的方法、与禁食相关的检测结果的确切时间）？ 4. **数值范围：**该项目适用的数值范围是什么？

讨论完第一个数据项（头围）后，与会者一致认为，虽然定义极端的最小值和最大值范围有助于保证数据质量，但还应考虑超出该范围的录入可能触发何种操作。例如，研究数据系统是否会阻止此类数值录入。由于这需要进一步达成共识，与会者同意将数值范围的纳入推迟到 CCPDD 的后续版本。因此，后续数据项均未考虑数值范围。在讨论过程中，经全体一致同意，参与者可以用类似的数据项替换某一数据项。

最终决定通过使用互动投票软件 Vevox (https://www. vevox.com/live-voting-app)进行投票予以确认。该软件展示了受访者中同意和不同意纳入该数据项的比例。线上参会者也可以参与投票。若某数据项被类似的数据项替代，则不再对此进行投票，因为替代过程本身已获得一致同意。

投票结束后，如果限定符或指导意见仍需进一步达成共识，参与者可以选择将数据项延至 CCPDD 第二轮迭代。对于引发讨论、被替换或被推迟的数据项，则提交 CDISC 进行最终评估。收到 CDISC 的反馈后，数据字典最终定稿并发布。CCPDD 的整个发育流程如**图 1**所示。

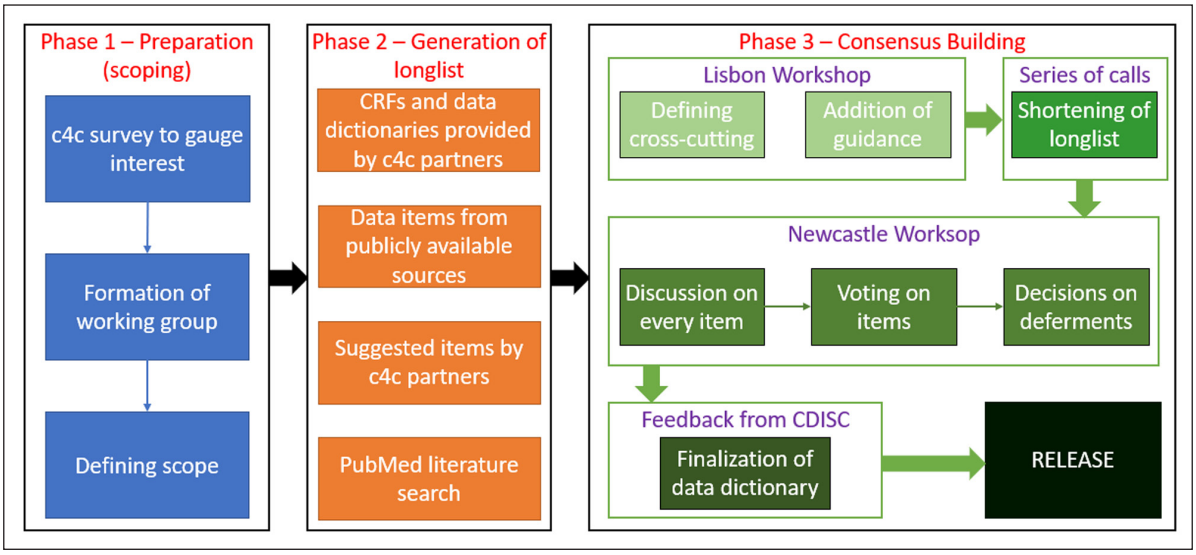

图 1：跨领域儿科数据字典的发育过程中各事件的顺序。方框以颜色区分：蓝色（第 1 阶段）、橙色（第 2 阶段）和不同色调的绿色（第 3 阶段）。事件顺序以箭头表示。较大方框内没有箭头连接的文本框表示这些事件是同时进行的。

结果

调查结果

行业调查的响应率为 90%（9/10）。调查显示，所有公司都制定了用于创建病例报告表（CRF）的标准操作规程。九家公司中有六家在创建病例报告表时采用基于文档架构的结构化标准（例如 CDISC）。基于文档架构的标准具有层级结构，可用于对临床文档进行编码。九家公司中有八家认为，建立一套大多数儿科临床试验都会收集的核心跨领域数据项将会有所裨益。所有公司都同意，病例报告表中采集的数据必须可互操作。事实上，有七家公司会汇总不同试验的数据。除一家外，所有公司都认为，儿科试验中数据项记录方式缺乏标准化和协调统一，是亟待解决的问题。

学术界和非 EFPIA 受益者的调查结果相似。回复率为 100%（34/34）。约 82%（28/34）的受访者认为，为儿科试验制定一套核心数据项会有益处，而一位受访者不认同。绝大多数受访者（85%；29/34）认为，病例报告表（CRF）采集的数据必须可互操作（两人回答“不是很重要”，三人选择“不作答”）。近 95%（32/34）的受访者认为，儿科数据存储缺乏标准化和协调统一这一问题亟待解决（两人选择“不作答”）。

调查显示，人们普遍对开发 CCPDD 感兴趣，其主要益处在于实现数据项标准化以及各机构间的协调统一。

字典编制

在这 16 个月里，数据字典经历了多次修订。随着项目的推进，最后一次研讨会在纽卡斯尔大学举行，并在会上作出了最终决定。以下是围绕几个有争议的数据项进行讨论的摘要。由于关于头围的讨论导致数值范围被排除在后续数据项之外（某些情况下，限定符则为可选项），我们在 **方框 2** 中完整呈现了由此引发的讨论。

体质指数（BMI）曾多次被纳入并从数据字典的不同版本中移除，这表明尚未达成共识。最终，该条目因以下原因从数据字典中删除：1）体质指数的正常范围取决于特定人群，因此无法标准化；2）许多临床医生质疑体质指数是否在各项试验中常规收集。经一致同意，改为纳入体表面积（BSA），因为体表面积常用于给药，因此在临床试验中会常规收集。讨论还涉及 gender、sex、genotypic/phenotypic sex 等术语的使用。虽然决定采用当前 CDISC 对 sex 的定义，但尚未确认 CCPDD 应使用哪个术语。与会者就 sex 和表型性别展开了讨论，最终决定暂缓定论，留待制定 CDISC PUG 时进一步讨论。投票结果详见**表 1**。

方框 2：关于头围的讨论摘要，该讨论促使今后的数据项不再考虑数值范围。

头围

当前 CDISC 定义：在头部最宽处测量的周长。

纳入范围：与会者一致认为，这是一个跨领域数据项，在新生儿试验中尤为常见。大家一致同意，应始终记录测量方法，因为 c4c 并未规定测量方法。头围的任何异常都需要作为单独的数据项记录，但该数据项不属于跨领域数据项。

单位：可能使用厘米、毫米或英寸。

限定符：测量日期很重要，但测量时间并不重要。日期通常会与测量值一同记录，因此无需进一步提出建议。因此，决定对某些数据项将限定符设为可选。

数值范围：大多数与会临床医生认为，在儿科数据字典中使用数值范围可能会带来问题，原因包括：1）不同儿科年龄组需要采用不同的数值范围；2）研究方案会提供定义；3）儿童在常规就诊时的数值常常超出正常范围。接下来还需要考虑，录入超出该范围的数值后可能触发什么处理。例如，系统是否会阻止此类录入？是否会触发警报？这些问题还需要进一步达成共识，也可以纳入 CCPDD 第二轮迭代的讨论。因此，研讨会决定暂不考虑数据项的数值范围。

经讨论，囟门闭合、妊娠相关事件和粪便采样这三个条目被认定不属于范围，并在投票中遭到否决。在此前的讨论中，这些条目同样存在争议。囟门闭合被认为仅与某些新生儿疾病密切相关，但大多数与会者认为它不属于跨领域内容。同样，与会者一致认为，各项研究中粪便样本的采集频率不足，因此不应视为跨领域内容。对于妊娠相关事件，讨论主要围绕如何区分孕产妇妊娠与可能怀孕的青春期女孩相关条目展开。孕产妇妊娠相关条目可能属于跨领域内容（例如胎龄），但青春期儿童的妊娠相关事件被认定不属于范围。

有六个数据项——婴儿喂养、发育、阿普加评分、合并用药、合并症和常规血液检查——在投票期间获得通过，但随后被推迟至 CCPDD 的下一次迭代。主要原因是，这些数据项的儿科特异性还需要进一步评估。对于婴儿喂养，

表 1：关于是否将数据项纳入 CCPDD 的投票结果。颜色说明：绿色——多数票赞成纳入，随后已纳入；红色——多数票反对纳入，原因是该数据项超出范围；黄色——多数票赞成纳入，但需要延期处理或进一步审查。

Item	Yes %	No %	Deferred	Notes
Date of birth	100	0		
Estimated gestational age	100	0		
Phenotypic Sex/Sex	100	0		Questions remained whether to use sex or phenotypic sex
Date of death	100	0		
Head circumference	100	0		
Blood pressure	100	0		Systolic and diastolic blood pressure
Body mass index				Replaced by body surface area
Body surface area				Replaced item – no voting
Height	100	0		
Weight	100	0		
Total body length	91.7	8.3		
Heart rate	100	0		
Pulse rate	81.8	18.2		
Respiratory rate	75	25		
Oxygen saturation	81.8	18.2		
Temperature	83.3	16.7		
Diagnosis	91.7	8.3		
Fontanelle closure	36.3	63.6		
Pubertal status	75	25		Split into Tanner Staging (stages for male genitalia, male pubic hair, female breast, female pubic hair), testicular volume and menarche (date and age)
Pregnancy related events	18.2	81.8		
Infant feeding	70	30	X	
Stool sampling	44.5	54.5		
Development	100	0	X	
APGAR score	100	0	X	
Concomitant medications	92.3	7.7	X	
Comorbidities	84.6	15.4	X	
Standard blood tests	100	0	X	

与会者一致认为，“是否母乳喂养”属于范围之内，但若不包括喂养时长和程度（全母乳、部分母乳、不母乳），信息就不完整。这些内容需要明确定义。儿童发育测量是一项跨领域工作，但涉及面很广，包括功能、社会和运动发育。还需要评估若干子项，包括问卷、评分和量表，并确定现有 CDISC标准是否涵盖这些主题。研讨会召开时，基于 AGPAR评分的数据表示方式仍在制定中。与会者曾讨论是否应在此版本的 CCPDD 中纳入一些常见血液检查——肌酐、电解质、葡萄糖和全血细胞计数；最终一致同意在 CCPDD 下一轮迭代中加入完整的血液检查项目。与会者还向 CDISC 征求了有关添加合并用药的进一步意见。

研讨会最终决定将青春期阶段分为三个子类别：Tanner分期[28]（男性生殖器、男性阴毛、女性乳房和女性阴毛的分期）、睾丸体积以及初潮（日期和年龄）。争议之一是睾丸体积是否应与Tanner分期分开。

研讨会结束后，在审阅 CDISC 纳入的项目后，对 CCPDD 作出了最终修改。撤销了删除体质指数的决定，保留了体质指数和体表面积。作出这一决定的理由是，虽然体表面积对年龄较小的患者更为相关，但体质指数是年龄较大的儿科患者常规收集的数据项。各方同意将表型性别一词纳入 CCPDD，并在制定 CDISC PUG 期间讨论是否更新 CDISC 对性别的定义。由于需要与“合并症”作出适当区分，“诊断”被推迟至下一次迭代再加入。此前被推迟至下一次迭代的合并用药，则提前纳入了第一版，因为目前 CDISC 对合并用药数据的建模似乎已足够完善。包含 25 个术语的最终字典于 2019 年 8 月 29 日 以 Microsoft Excel 文件形式发布。虽然禁止分享完整字典（原因见下文），但在**图 2**和**3**中分享了两个快照。

Sub-topic	CDISC mapped Item	CDISC units (CDISC Codelist) C-CODE value in brackets	CDISC SDTM Domain	CDISC Qualifier Variables	Guidance	Source of information
Tanner staging	Male Genitalia Stage	N/A	Questionnaires, Ratings, and Scales (QS)		Tanner Staging is modelled in the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf
	Male Pubic Hair Stage	N/A	Questionnaires, Ratings, and Scales (QS)		Tanner Staging is modelled in the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf
	Female Breast Stage	N/A	Questionnaires, Ratings, and Scales (QS)		Tanner Staging is modelled in the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf
	Female Pubic Hair Stage	N/A	Questionnaires, Ratings, and Scales (QS)		Tanner Staging is modelled in the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf
Menarche	Date of menarche		Medical History (MH) Clincial Events (CE)		See the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules Refer to the followng CDISC documentation on recommendations on how to collect dates/times in	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf
	Age of menarche		Reprodictive System Findings (RP)		See the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf
Testicular Volume	Testicular Volume	**UNIT Codelist (C71620)** mL (C28254)	Reprodictive System Findings (RP)	Location (RPLOC) = TESTES Laterality (RPLAT) = LEFT or RIGHT Method (RPMETHOD) = e.g., ORCHIDOMETERY	See the CDISC Therapeutic Area Data Standards for Type 1 Diabetes - Pediatrics and Devices Modules	https://www.cdisc.org/system/files/members/standard/ta/TAUG- T1D Pediatrics and Devices v2.0.pdf

图2：跨领域儿科数据字典中青春期状态部分的截屏。

CDISC mapped Item	CDISC units (CDISC Codelist) C-CODE value in brackets	CDISC SDTM Domain	CDISC Qualifier Variables	Guidance	Source of information CDASH = Data Collection SDTM = Data Tabulation
Height	cm (C49668) m (C41139) in (C48500) ft (C71253)	Vital Signs (VS)	unit, standardised unit	Specify standardised unit	CDASHIG v2.1 - https://www.cdisc.org/system/files/members/standard/foundational/CDASHIG%20v2.1-Final_Rev.pdf Section 8.3.18 SDTMIG v3.3 - https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3/html#Vital+Signs
Total body Length	cm (C49668) in (C48500) ft (C71253)	Vital Signs (VS)	Vital Signs Position of Subject (VSPOS)? - not needed if new definition accepted as it states "recumbent"	Specify standardised unit	CDASHIG v2.1 - https://www.cdisc.org/system/files/members/standard/foundational/CDASHIG%20v2.1-Final_Rev.pdf Section 8.3.18 SDTMIG v3.3 - https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3/html#Vital+Signs
Weight	kg (C28252) g (C48155) LB (C48531) *No CDISC controlled terminology exists for stone*	Vital Signs (VS)		Specify standardised unit	CDASHIG v2.1 - https://www.cdisc.org/system/files/members/standard/foundational/CDASHIG%20v2.1-Final_Rev.pdf Section 8.3.18 SDTMIG v3.3 - https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3/html#Vital+Signs
BMI	kg/m2 (C49671)	Vital Signs (VS)		Raw data values must be collected (Height and weight)	CDASHIG v2.1 - https://www.cdisc.org/system/files/members/standard/foundational/CDASHIG%20v2.1-Final_Rev.pdf Section 8.3.18 SDTMIG v3.3 - https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3/html#Vital+Signs
Body surface area	m2 (C42569)	Vital Signs (VS)	Method used: Haycock Gehan and George Boyd Fujimoto Shuter and Aslani Schlich	Need to define the options - which are used/not used in paediatrics	CDASHIG v2.1 - https://www.cdisc.org/system/files/members/standard/foundational/CDASHIG%20v2.1-Final_Rev.pdf Section 8.3.18 SDTMIG v3.3 - https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3/html#Vital+Signs

图3：跨领域儿科数据字典中生命体征部分的截屏。红色文字可能会在字典第2版中修改。

讨论

发布后，CCPDD 立即被采纳为以下三项可行性验证儿科临床试验创建病例报告表的工具：泊沙康唑新给药方案囊性纤维化并伴有曲霉菌感染的儿科患者（cASPerCF，ClinicalTrials.gov标识符：NCT04966234，EudraCT编号：2019-004511-31）、川崎病冠状动脉瘤预防

试验（KD-CAAP，EudraCT编号：2019-004433-17）以及对早产儿动脉导管使用对乙酰氨基酚进行预防性治疗（TREOCAPA – ClinicalTrials.gov标识符：NCT04459117，EudraCT编号：2019- 004297-26）。另外四项由行业资助的试验原计划于2021年启动，但因COVID-19疫情而推迟；制药公司也错开了“首位患者首次访视”的时间安排。

将 CDISC 纳入开发过程，确保 CCPDD 包含 CDISC 实施说明和示例，因此 CCPDD 一经发布便可立即使用（无需进行任何格式转换或翻译）。用于收集跨领域数据项的标准程序将使来自不同来源的汇总数据能够实现互操作，并有望促进这些数据在未来研究中的共同再利用。对于儿科学这类领域而言，数据的互操作性尤为重要，因为常规研究方法受到人群规模有限、监管挑战和伦理考量的制约。因此，长期使用该数据字典将有助于实现 FAIR原则所强调的互操作性和可再利用性。数据的可发现性和可访问性仍由数据提供者负责。

尽管跨领域数据项是一个合适的起点，但主要挑战在于将字典扩展至跨领域术语之外。为此，CDISC 目前正在开发 PUG，旨在涵盖临床试验中所用儿科术语的数据收集和制表示例。CCPDD 是 PUG 的输入之一，但也会使用其他输入。PUG 已进入公开评审的最后阶段，预计将于 2022年底在 CDISC网站上发布。[29]

尽管创建 CCPDD 时采用了严格的流程，但该字典仍存在一些局限。共识构建的本质意味着，另一组专家可能会得出一套不同的数据项。虽然各方尽力达成一致共识，但某些决定确实需要通过投票作出，例如关于粪便采样的投票。使用 Microsoft Excel 进行发育，导致 CCPDD 结构扁平、可视化效果不佳。因此，c4c 正在探索 CCPDD 的下一版是否可以基于临床建模工具（CMT）开发。[30] 该工具已在项目期间进行试点，其中包含思维导图可视化功能，并支持导出为常见的数据格式（如 JSON、XML）。

上述 CCPDD 下一版目前正在开发中。除新增数据项以及可能采用 CMT 外，这一版还将调整控制流程，以便持续添加相关的新项目。此外，c4c 的其他几项工作也可能在未来扩展 CCPDD。这些工作包括探索哪些真实世界数据项最适用于儿科 IV 期（上市后监测）试验，以及在罕见病研究中重复使用对照组数据。[20]

结论

c4c 正在编制 CCPDD，以满足儿科研究领域的一项需求：提高儿科临床试验中所收集数据的标准化程度。团队采用高度协作的共识建立方法，确保字典中纳入最相关的数据项。CCPDD 聚焦于常见且具有跨领域适用性的数据项，以最大限度地发挥其实施的潜在影响。

免责声明

本出版物反映作者的观点，IMI、欧盟、EFPIA 或任何关联合作伙伴均不对其中所含信息可能被如何使用承担责任。

致谢

本项目获得创新药物倡议第二联合 undertaking 根据资助协议编号 777389 提供的资金支持。该联合 undertaking 获得欧盟地平线2020年研究与创新计划以及 EFPIA 的支持。

利益冲突

作者声明不存在任何利益冲突。

参考文献

1. **Joseph PD, Craig JC, Caldwell PHY.** 儿童临床试验。英国临床药理学杂志。2015；79(3)：357–369 12305 2 10 1111。DOI：https://doi.org/。/bcp.。**Lim A, Cranswick N.** 临床试验、伦理问题与患者招募：澳大利亚视角。儿科与围产期药物治疗。2003；5：183–187。DOI：https://doi。org/10 146300903774115775 3 1185。/。**von Baeyer CL.** 儿童疼痛强度自我报告：量表选择、局限性与解读。疼痛研究与管理。2006年秋；11(3)：157–62。DOI：https://doi。org/10 197616 4 1155 2006。/ /。**Hwang TJ, Tomasi PA, Bourgeois FT.** 欧盟《儿科法规》框架下临床试验完成及结果报告的延迟：一项队列研究。《公共科学图书馆·医学》。2018年；15(3)：e1002520–e1002520。DOI：https://doi。org/10 1002520 5 1371。/journal.pmed.。**Caldwell PH, Murphy SB, Butow PN, Craig JC.** 儿童临床试验。《柳叶刀》。2004年8月28日至9月3日；364(9436)：803–11。DOI：https://doi。org/10 0 6 1016 0140 6736 04 16942。/S - ( ) -。**Green DJ, Burnham JM, Schuette P,** 等。提交给美国食品药品监督管理局的儿科疗效试验主要终点。J Clin Pharmacol。2018年；58(7)：885–890。DOI：https://doi.org/10 1109 7 1002。/jcph.。**OECD.** 无国界数据：促进知识与创新。2019年。8. **Seneviratne MG, Kahn MG, Hernandez-BoussardT.** 整合异构临床数据以促进知识发现。太平洋生物计算研讨会。2019年；24：439–443。9. **Simpson E, Brown R, Sillence E,** 等。了解长期健康问题患者使用数字技术共享患者生成的健康数据时的障碍与促进因素：叙述性综述。Front

公共卫生。2021年；9：641424。DOI：https://doi。org/10 641424 10 3389 2021。/fpubh.。。。**Kochhar S, Knoppers B, Gamble C, Chant A, Koplan J, Humphreys GS.** 临床试验数据共享：挑战所在。*BMJ Open*。2019年；9(8)：e032334。DOI：https://doi.org/10.1136/ bmjopen-2019 032334 11 -。**Janczewski M.** 儿科护理健康信息技术功能概要——项目。2022年11月15日，更新于2022年5月12日。https://confluence.hl7.org/display/EHR/Pediatri c+Care+Health+IT+Functional+Profile+-+Project。访问日期：2022年11月15日。12. **儿科试验网络（PTN）**。更新于2019年5月9日。https://www.nichd.nih.gov/research/ supported/ptn。访问日期：2019年11月15日。13. **临床试验（ECHO IDeA 州儿科临床试验网络）**。 https://echochildren.org/idea-states-pediatric-clinical-trials-network/。访问日期：2022年11月15日。14. **FDA**。 研究数据标准：须知事项。https://www.fda.gov/files/drugs/published/ - St udy-Data-Standards--What-You-Need-to-Know.pdf。访问日期：2022年2月9日。15. **Bobbitt DR, LeRoy B, Palmer A,** 等。临床数据标准与研究科学、技术和数据来源的新世界。国家公共卫生研究所期刊。2019年；68(3)：194–201。DOI：https://doi.org/ 10 20683 68 3_194 。/jniph.。。。16 **Facile R, Muhlbradt EE, Gong M,** 等。在真实世界数据中使用临床数据交换标准协会（CDISC）标准：来自定性德尔菲调查的专家观点。*JMIR* 医学信息学。2022年1月27日；10(1)：e30363。DOI：https://doi.org/10 30363 17 2196。/。**Picoraro JA, Lee D, Heller CA,** 等。克罗恩病与结肠炎基金会儿科炎症性肠病临床创新会议：描绘儿科炎症性肠病的未来。炎症性肠病 疾病。2019年1月1日；25(1)：27–32。DOI：https://doi。org/10 205 18 1093。/ibd/izy。**Aurich B, Apele-Freimane D, Banaschewski T,** 等。c4c：儿科药物警戒：儿童用药研究与开发中的方法学考量——c4c专家组白皮书。英国临床药理学杂志。n/a(n/a)。DOI：https://doi.org/10 15119 19 1111。/bcp.。**Takahara S, Saito TI, Imai Y, Kawakami T,Murayama T.** 学术界研究者发起的临床试验中临床数据交换标准协会/研究数据制表模型的用例分析。名古屋医学科学杂志。2022年2月；84(1)：120–132。DOI：https://doi.org/10.18999/ nagjms.84.1.120

20. **Sen A, Palmeri A, Lee J,** 等. 通过学术界—产业界合作了解儿科数据标准面临的挑战：一项 conect4children (c4c) 定性研究。《国际卫生规划与管理杂志》；2022年。DOI：https://doi.org/10 3592 21 1002. /hpm. . **Wilkinson MD, Dumontier M, Aalbersberg IJ,** 等. 科学数据管理与 stewardship 的 FAIR指导原则。《科学数据》。2016年；3(1)：160018。DOI：https://doi.org/10.1038/ s data.2016.18 22. **Courbier S, Dimond R, Bros-Facer V.** 共享并保护我们的健康数据：基于证据的方法，探讨罕见病患者对数据共享和数据保护的看法——定量调查与建议。《Orphanet罕见病杂志》。2019年7月12日 14 1 175 10 1186 ; ( ): . DOI：https://doi.org/ . / s13023-019-1123-4 23. **Austin CP, Cutillo CM, Lau LPL,** 等. 罕见病研究的未来 2017–2027：IRDiRC视角。《临床与转化科学》。2018年；11(1)：21–27。DOI：https://doi.org/10.1111/ cts.12500 24. **Turner MA, Hildebrand H, Fernandes RM,** 等. conect4children（c4c）联盟：改善欧洲儿童用药临床研究的潜力。《药物医学》。2021年3月；35(2)：71–79。DOI：https://doi.org/10.1007/ s40290-020-00373-6 25. **Khodiyar V.** 更好的研究带来更好的数据——FAIR如何推动开放研究向前发展。Springer Nature；2021年。26. **Jacobsen A, de Miranda Azevedo R, Juty N,** 等. FAIR原则：阐释与实施注意事项。《数据智能》。2020年；2(1–2)：。。 10 29 10 1162 _ _00024 – 。27 DOI：https://doi.org/ . /dint r . **Owen R, von Schomberg R, Macnaghten P.** 尚未完成的旅程？对负责任的研究与创新十年的反思。《负责任创新杂志》。2021年；8(2)：217–233。DOI：https://doi.org/10 1948789 28 1080 23299460 2021 . / . . . **Emmanuel M, Bokor BR.** *Tanner*分期。StatPearls出版社，Treasure Island (FL)；2022年。29. **CDISC**。 儿科用户指南 v1.0。CDISC。2022年。https://www.cdisc.org/standards/therapeutic-areas/pediatrics/pediatrics-user-guide-v1-0。访问日期：2022年8月12日。30. **Amadi C, Leary R, Palmeri A,** 等. 如何使用临床数据建模工具来表示与儿科临床试验相关的数据项？从Conect4children（c4c）联盟汲取经验。《应用科学》。2022年；12(3)：1604。DOI：https://doi.org/10.3390/app12031604

本文引用格式： Sen A, Hedley V, Owen J, Cornet R, Kalra D, Engel C, Palmeri A, Lee J, Roze JC, Standing J, Warris A, Pansieri C, Leary R, Turner M, Straub V. 儿科临床数据标准化：conect4children (c4c) 跨领域儿科数据字典的开发。临床数据管理学会期刊。2022年；2(3)：4，第1–10页。DOI：https://doi. org/10.47912/jscdm.218

提交： 2022年8月9日 **接受：** 2022年11月23日 **发表：** 2022年12月28日

版权： © 2022年 SCDM 根据署名-非商业性使用-相同方式共享（知识共享署名-非商业性使用-相同方式共享）许可协议，以开放获取方式发布 JSCDM 内容。该许可协议允许他人以非商业方式混合、改编和基于本作品创作新作品，但须注明 SCDM 和作者，并以相同条款授权其新作品。请参阅 https://creativecommons.org/licenses/by-nc-sa/4.0/。

《临床数据管理学会期刊》是一份由临床数据管理学会出版的同行评审开放获取期刊。 **开放获取**