

临床试验数据管理工作技术指南

2012 年 3 月 12 日

目 录

一、概述.....	1
1.1 国内临床试验数据管理现状.....	1
1.2 国际临床试验数据管理简介.....	2
1.2.1 ICH-GCP 对临床试验数据管理的原则性指导.....	2
1.2.2 21 CFR Part 11 对电子记录和电子签名的基本要求.....	3
1.2.3 “临床试验中采用计算机系统的指导原则”提供计算机系统开发的参照标准.....	3
1.2.4 GCDMP (Good Clinical Data Management Practice) 提供全面具体的数据管理要求.....	4
二、数据管理相关人员的责任、资质及培训.....	5
2.1 相关人员的责任.....	5
2.1.1 申办者.....	5
2.1.2 研究者.....	6
2.1.3 监查员.....	6
2.1.4 数据管理员.....	6
2.1.5 合同研究组织 (CRO).....	7
2.2 数据管理人员的资质及培训.....	8
三、临床试验数据管理系统.....	9
3.1 临床试验数据管理系统的重要性.....	9
3.2 数据质量管理体系的建立和实施.....	9
3.3 临床试验数据管理系统的基本要求.....	10
3.3.1 系统可靠性.....	10
3.3.2 临床试验数据的可溯源性 (Traceability).....	11
3.3.3 数据管理系统的权限管理 (Access Control).....	12
四、试验数据的标准化.....	12
4.1 临床数据标准化的现状与发展趋势.....	12
4.2 临床试验的数据标准化.....	13
4.2.1 CDISC 和 HL7.....	14
4.2.2 医学术语标准.....	16
4.2.3 临床试验报告的统一标准 (CONSORT).....	20
五、数据管理工作的主要内容.....	21
5.1 CRF 的设计与填写.....	21
5.1.1 CRF 的设计.....	21
5.1.2 CRF 填写指南.....	22
5.1.3 注释 CRF.....	22
5.1.4 CRF 的填写.....	22
5.2 数据库的设计.....	22
5.3 数据接收与录入.....	23
5.4 数据核查.....	23
5.5 数据质疑表的管理.....	25

5.6	数据更改的存档	25
5.7	医学编码	25
5.8	试验方案增补修改	26
5.9	实验室及其他外部数据	26
5.10	数据盲态审核	28
5.11	数据库锁定	28
5.11.1	数据库锁定清单.....	28
5.11.2	数据库锁定后发现数据错误.....	29
5.12	数据备份与恢复	30
5.13	数据保存	30
5.14	数据保密及受试者个人私密性的保护.....	33
5.14.1	数据保密.....	33
5.14.2	受试者个人私密性的保护.....	33
六、	数据质量的保障及评估.....	34
6.1	质量保障	34
6.1.1	质量控制.....	34
6.1.2	质量保证.....	37
6.2	质量评估	40
七、	安全性数据及严重不良事件报告.....	42
7.1	不良事件的获取、管理和报告.....	42
7.2	实验室数据	44
7.3	其他数据	45
7.4	严重不良事件数据	45
八、	参考文献	47

一、概述

临床试验数据质量是评价临床试验结果的基础。为了确保临床试验结果的准确可靠、科学可信，国际社会和世界各国都纷纷出台了一系列的法规、规定和指导原则，用以规范临床试验数据管理的整个流程。同时，现代新药临床试验的发展和科学技术的不断进步，特别是计算机、网络的发展又为临床试验及其数据管理的规范化提供了新的技术支持，也推动了各国政府和国际社会积极探索临床试验及数据管理新的规范化模式。

1.1 国内临床试验数据管理现状

我国的《药物临床试验质量管理规范》（GCP, Good Clinical Practice）对临床试验数据管理提出了一些原则性的要求，但关于具体的数据管理、操作的法规和规定目前还处于空白。由于缺乏配套的实施细则和有效的监管措施，导致我国在药物临床试验数据管理方面的规范化程度不高，临床试验数据管理质量良莠不齐，进而影响到新药有效性和安全性的客观科学评价。其次，国内临床试验数据管理系统的开发和应用尚处于起步阶段，临床试验的数据管理模式大多基于纸质病例报告表（CRF, Case Report Form）的数据采集阶段，电子化数据采集与数据管理系统应用有待提高。同时，由于缺乏国家编码标准，同类研究的数据库之间难以做到信息共享。

长期以来，由于我国是一个以仿制为主的国家，上述数据规范化管理的不足与新药临床试验科学要求之间的矛盾还未能充分凸显。但随着近几年建设创新型国家政策的提出，我国创新药物研发逐步增多，特别是国务院发布了我国中长期科学技

术发展规划纲要,其中“重大新药创制立项”是该规划在药物创新领域具体制定的策略,主要目标是研制一批具有自主知识产权的创新药物。可以想见,药物创新作为国家战略对我国临床试验数据进行规范管理的要求已非常紧迫,其发展的挑战和机遇均十分巨大。

1.2 国际临床试验数据管理简介

国际临床试验数据管理首先要符合“人用药品注册技术要求国际协调会议-临床试验质量管理规范”(ICH-GCP, International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use-Good Clinical Practice)的原则性要求。同时,各国还颁布了相应的法规和指导原则,为临床试验数据管理的标准化和规范化提供具体的依据和指导。如:美国 21 号联邦法规第 11 部分(21 CFR Part 11)对临床试验数据电子存档的规定(2003 年 8 月)、美国食品药品监督管理局(FDA, Food and Drug Administration)颁布的“临床试验中采用计算机系统的指导原则”(Guidance for Industry: Computerized Systems Used in Clinical Investigations, 2007 年 5 月)等。而且,由各国临床试验方面的学者和专家组成的临床试验数据管理组织(Society of Clinical Data Management, SCDM)经过长期的研究和讨论,还形成了一部非官方的“临床数据质量管理规范”(GCDMP, Good Clinical Data Management Practice),为现行的法规和规定尚未涉及的临床试验数据管理方面提供有关可接受的操作流程。

1.2.1 ICH-GCP 对临床试验数据管理的原则性指导

1997 年 6 月,ICH-GCP 正式颁布,对正规化临床试验进行

了系统科学的原则指导，在保护受试者利益的前提下，保证了临床试验的科学性、可靠性。虽然 ICH-GCP 中没有专门的章节阐述临床试验数据管理的要求，但其对临床试验开展过程中的研究者、研制厂商以及有关试验方案、随机化过程的记录、数据核查等都直接或间接的提出了原则性的规定，以保证整个临床试验过程中获得的各类数据信息真实可信、准确可靠。

1.2.2 21 CFR Part 11 对电子记录和电子签名的基本要求

在计算机技术飞速发展的前提下，FDA 为了保证能够让计算机技术在新药临床研究领域得到最大范围的应用，1997 年 3 月出台了有关电子记录和电子签名的法规“21CFR Part 11”。其主导思想是通过相关条文的规定提供有关电子记录与电子签名的准则，使得电子记录、电子签名具有与传统的手写记录与手写签名同等的法律效力，从而使得 FDA 能够接受电子化临床研究材料。

为了更好地理解和执行第 11 款的内容，经过几年的反复讨论并广泛征求意见，FDA 于 2003 年 8 月发布了正式的指导原则，对第 11 款的一些要求作了具体的建议，如对计算机系统的验证要求、稽查轨迹，以及文件记录的复制等。第 11 款的其它规定仍然有效。

1.2.3 “临床试验中采用计算机系统的指导原则”提供计算机系统开发的参照标准

FDA 在 1999 年 4 月颁布了“临床试验中采用计算机系统的指导原则”（Guidance for Industry: Computerized Systems Used in Clinical Trials），对计算机系统的特征、电子病例报告表（eCRF、稽查轨迹、电子记录、电子签名等作了明确的定义，成为了临

床试验中计算机系统开发的基本参照标准。FDA 于 2007 年 5 月更新该指导原则，从多个方面提出推荐，包括临床研究方案、标准操作程序、原始文件及其保存、内部安全、外部安全维护、其他系统特征以及人员培训等。该指导原则适用于计算机系统建立原始文档（电子记录），也适用于计算机系统处理自动化分析仪器收集到的原始数据，还可用于纸质文件再录入或直接将数据录入计算机系统，或者由计算机系统自动录入数据。

1.2.4 GCDMP (Good Clinical Data Management Practice) **提供全面具体的数据管理要求**

临床试验数据管理组织（Society of Clinical Data Management, SCDM）是一个为了提高数据管理水平，推动这一领域发展的非营利性专业团体。所制定的 GCDMP 是为了给现行的规章和指导性文件尚未涉及的临床试验数据管理方面提供有关切实可行的操作流程。该文件给临床试验数据管理工作划分了责任。每个章节都规定了相应标准操作程序的最低标准和最高规范。最低标准与最高规范都以条目的形式概括了每一章节的主要内容。最低标准指的是能确保数据完整，可靠以及被正确处理的标准，即所谓的数据完整性要求；最高规范指的是除了确保数据完整性之外，提供更高质量、效率和功能，更低风险的数据所要遵循的要求。因此 GCDMP 是更为全面和细化的临床试验数据管理技术指导。

总之，国际社会和发达国家均已建立了临床试验数据管理的若干法规、规定和技术指导原则，以保证试验数据的质量。而我国这方面的起步较晚，发展缓慢，临床试验数据管理欠规范化，直接影响了我国新药研发与监管。目前，新药创制的国家

战略计划又对临床试验数据规范化管理提出了更加紧迫的需求。鉴于其重要性和紧迫性，在积极总结和调研临床试验数据管理工作的当前技术水平和发展趋势的基础上，特制订本技术指南。

本指南从数据管理相关人员的职责、资质和培训，管理系统的要求，试验数据的标准化，主要工作内容，数据质量的保障和评估，以及安全性数据及严重不良事件六个方面进行全面阐释，旨在对我国临床试验的数据管理工作起到规范化和指导性作用，使得国内的临床试验水平早日与国际接轨。

二、 数据管理相关人员的责任、资质及培训

临床试验数据管理工作要求临床试验研究项目团队共同努力、通力协作。研究团队中与数据管理工作相关的人员包括申办者、研究者、监查员、数据管理员，以及合同研究组织（CRO, Contract Research Organization）等。

2.1 相关人员的责任

2.1.1 申办者

申办者是保证临床数据质量的最终责任人。申办者应制定质量管理评价程序、质量管理计划与操作指南，并且应设立稽查部门，由不直接涉及试验的人员定期对质量体系的依从性进行系统性检查。

申办者的失责行为举例：

- 研究者未受到培训不能准确填写 CRF；
- 研究方案不明确或不合理。

2.1.2 研究者

研究者应确保以 CRF 或其他形式报告给申办者的数据准确、完整与及时，而且应保证 CRF 上的数据与受试者病历上的源数据一致，并必须对其中的任何不同给出解释。

研究者的错误/不当行为举例：

- 违反研究方案，如错误的访视时间；
- 源数据录入 CRF 时错误；
- 实验室仪器人为测量误差；
- 研究者造假。

2.1.3 监查员

监查员应根据源文档核查 CRF 上的数据，一旦发现其中有错误或差异，应通知研究者，以确保所有数据的记录和报告正确和完整。

临床监查中常见问题举例：

- 知情同意书缺失或签署存在问题；
- 无原始病历或原始病历中无记录（缺失或不全）；
- CRF 填写空缺、错误或不规范；
- 不良事件的记录不完整；
- 检验结果不能溯源（实验室数据、心电图、X 射线片等）；
- 试验药物管理不当。

2.1.4 数据管理员

数据管理员应按照研究方案的要求，参与设计 CRF，并建立和测试逻辑检验程序。在 CRF 接收后，录入人员要对 CRF 作录入

前的检查；在CRF数据被录入数据库后，利用逻辑检验程序检查数据的有效性、一致性、缺失和正常范围等。数据管理员对发现的问题应及时清理，可以通过比对CRF中其他部位的数据作出判断以解决问题，也可通过向研究者发放数据质疑表（Query）而得到解决。

数据管理员应参加临床研究者会议，为研究团队及时提出改善与提高数据质量的有效措施。

数据管理员的失责行为举例：

- CRF 表格设计不合理；
- 逻辑检验程序错误；
- 按照质疑表更新数据库时的错误。

2.1.5 合同研究组织（CRO）

ICH GCP 指出申办者可以将与临床试验有关的工作和任务，部分或全部委托给一个 CRO，但是，试验数据的质量和完整性的最终责任永远在申办者。CRO 应当实施质量保证和质量控制。

申办者首先应明确数据管理外包的范围，如果计划将数据管理工作外包，那么接下来就要选择合适的CRO，应对候选CRO的资质和能力等进行评价。评价CRO时应主要考虑以下因素：

- CRO的资质、以往业绩及合同履行能力
- 质量控制、质量保证的流程
- 数据管理系统的验证，以及设施条件
- 数据管理SOP以及遵守SOP的证明
- 员工资质、对SOP的掌握情况及其培训记录

- 文档修改控制过程的记录
- 文件保管系统

一旦作出选择，申办者将与CRO签订有效合同，在合同中需明确双方的责、权、利。试验申办者必要时应对CRO进行相关的培训，以保证其所提供的服务符合申办者质量标准要求。在临床试验数据管理过程中，试验申办者需对CRO进行的活动进行即时有效的管理、沟通和核查，以确保其遵守共同商定的流程的要求。申办者的质量管理计划中必须包括CRO的质量管理信息，同时必须明确流程和期望结果。

2.2 数据管理人员的资质及培训

负责临床试验数据管理的人员必须经过 GCP、相关法律法规、相关标准操作程序（SOP，Standard Operating Procedure），以及数据管理的专业培训，以确保其具备工作要求的适当的资质。数据管理专业培训应包括但不局限于：

- 数据管理部门标准操作程序和部门政策
- 临床试验数据标准化文档及存档规则
- 数据管理系统及相关的计算机软件的应用与操作能力的培训
- 法规和行业标准：GCP, SFDA 法规和指导原则, 和 ICH 指导原则
- 保密性，私密性和数据安全性培训。

数据管理人员都必须保持完整的培训记录以备核查，培训记录需提供课程名称、培训师名称、课程的日期、完成状况、受训人员及其主管的签名。如果是基于网络的培训，系统应提

供培训证书，证书上有课程的名称、参训人员姓名，以及完成培训的时间。

数据管理人员还应该通过继续教育不断提升专业素质，以保障数据管理工作的高品质完成。

三、 临床试验数据管理系统

3.1 临床试验数据管理系统的重要性

数据管理的目的是确保数据的可靠、完整和准确。数据管理过程包括数据接收、录入、清理、编码、一致性核查、数据锁定和转换。数据管理的目标是获得高质量的真实数据（Real Data）。因此，临床试验数据管理的各个阶段需要在一个完整、可靠的临床试验数据管理系统下运行，临床试验项目团队必须按照管理学的原理建立起一个体系，即数据管理系统，对可能影响数据质量结果的各种因素和环节进行全面控制和管理，使这些因素都处于受控状态，使临床研究数据始终保持在可控和可靠的水平。此处的数据管理系统不是指狭义的计算机系统，而是一种广义的数据质量管理体系（QMS, Quality Management System），它是临床试验项目管理系统的一个组成部分。

3.2 数据质量管理体系的建立和实施

数据质量管理体系的建立是应用管理科学、提高管理水平、不断发展的过程。

建立和实施质量管理体系首先需确立质量方针和目标，以确定预期结果，帮助管理者利用其资源达到这些结果。质量方针是管理者的质量宗旨和方向，质量目标是方针的具体化，是管理者在质量方面所追求的目的。

质量管理体系依托组织机构来协调和运行，必须建立一个与质量管理体系相适应的组织结构。组织机构应明确规定数据管理相关人员的责任和权限。

质量管理体系的实施和运行是通过建立贯彻质量管理体系的文件来实现的。质量管理体系文件一般由四部分组成：质量手册、程序文件、作业指导书、质量记录。质量手册的核心是对质量方针目标、组织机构及质量体系要素的描述；程序文件是对完成各项质量活动的方法所作的规定；作业指导书是规定某项工作的具体操作程序的文件，也就是数据管理员常用的“操作手册”或“操作规程”等；质量记录是为完成的活动或达到的结果提供客观证据的文件。

完成质量管理体系文件后，要经过一段试运行，检验这些质量管理体系文件的适用性和有效性。数据管理机构通过不断协调、质量监控、信息管理、质量管理体系审核和管理评审，实现质量管理体系的有效运行。

数据管理质量体系的建立、实施和运行是一个动态的过程，最重要的是要求数据管理相关人员将质量管理的理念贯彻到数据管理的日常工作之中。

3.3 临床试验数据管理系统的基本要求

3.3.1 系统可靠性

系统可靠性是指系统在规定条件下、规定时间内，实现规定功能的能力。临床试验数据管理系统必须经过基于风险考虑的验证，以保证数据完整、安全和可信，并减少因系统或过程的问题而产生错误的可能性。

临床试验数据管理系统的验证包括以下方面内容：

- 证明系统满足特定的使用目的
- 系统的访问控制，以及用户管理
- 确保数据的完整性，包括防止数据的删除或丢失
- 防止数据和文档未经授权的更改
- 稽查轨迹
- 灾难恢复计划/备份
- 系统维护和变更控制
- 用户培训

数据管理软件开发商必须对其软件进行严谨的设计、严肃的验证，以及严格的检测。数据管理软件开发商应提供系统设计与使用说明书。临床试验数据管理软件终端用户，需有在本地进行安装调试、测试和人员培训的记录。

3.3.2 临床试验数据的可溯源性 (Traceability)

临床试验数据管理系统必须具备可以为临床试验数据提供可溯源性的性能。CRF中数据应当与源文件一致，如有不一致应作出解释。对CRF中数据进行的任何更改或更正都应该注明日期、签署姓名并解释原因（如需要），并使原来的记录依然可见。

临床试验数据的稽查轨迹（Audit Trail），从第一次的数据录入以致每一次的更改、删除或增加，都必须保留在临床试验数据库系统中以保证从原始资料到申报数据全过程的透明度。稽查轨迹应包括更改的日期、时间、更改人、更改原因、更改前数据值、更改后数据值。此稽查轨迹为系统保护，不允许任

何人为的修改和编辑。稽查轨迹记录应存档并可查询。

3.3.3 数据管理系统的权限管理（Access Control）

临床试验数据管理系统必须有完善的系统权限管理。纸质化或电子化的数据管理均需要制定SOPs进行权限控制与管理。对数据管理系统中不同人员或角色授予不同的权限，只有经过授权的人员才允许操作（记录、修改等），并应采取适当的方法来监控和防止未获得授权的人的操作。

电子签名(electronic signature)是电子化管理系统权限管理的一种手段。对于电子化管理系统来说，系统的每个用户都应具有个人帐户，系统要求在开始数据操作之前先登录帐户，完成后退出系统；用户只能用自己的密码工作，密码不得共用，也不能让其他人员访问登录；密码应当定期更改；离开工作站时应终止与主机的连接，计算机长时间空闲时实行自行断开连接；短时间暂停工作时，应当有自动保护程序来防止未经授权的数据操作，如在输入密码前采用屏幕保护措施。

四、 试验数据的标准化

4.1 临床数据标准化的现状与发展趋势

近些年来随着国家大力推动医疗机构进行信息化建设，医院信息系统建设发展迅速，电子病历作为医院信息系统的重要组成部分，其应用也正成为一种趋势。但是，信息标准化的问题已经成为医疗卫生信息化进一步发展的瓶颈。目前在建立区域医疗信息化过程中存在系统分割、业务流程不统一、信息系统孤岛等问题。在医学研究领域，各医院、科研院所等都积累了大量丰富的临床数据，但不能互相共享利用。所以，为了实

现跨机构、跨区域、跨领域的临床数据资源互联互通、共享利用，必须研究建立统一的临床数据标准体系。

我国标准化研究起步较晚，随着加入WTO，以及贸易全球化和经济集团化、高新技术的迅猛发展，优先借鉴和引用国际标准(ISO¹、IEC²及ITU³的标准以及被ISO认可收入KWIC索引⁴的其它25个国际组织制定的标准)和国外先进标准(未经ISO确认并公布的其它国际组织的标准、发达国家的国家标准、区域性组织的标准，国际上有权威的企业公司标准中的先进标准)已经成为发展趋势。与临床数据相关的国际标准和国外先进标准包括：(1)数据/信息交换标准，如医院信息传输协议HL7、临床数据交换标准CDISC、放射影像信息交换标准DICOM等；(2)医学术语及代码标准，如实验室检查结果代码LONIC、临床医学术语SNOMED、疾病诊断分类代码ICD-10等；(3)临床医学文档标准，如临床文档结构标准CDA等；(4)临床医学概念标准，如HL7信息模型RIM等。

4.2 临床试验的数据标准化

临床试验数据标准化的意义在于：

- 标准化的数据格式是临床试验数据管理系统与临床试验机构建立医疗信息互通性的基础。
- 在申办者内部不同研究之间建立无缝数据交换，并为申办者之间的交流，申办者与药物评审机构之间的交流提供便利。

¹ ISO: 国际标准化组织。

² IEC: 国际电工委员会，是世界上成立最早的国际性电工标准化机构，负责有关电气工程和电子工程领域中的国际标准化工作。

³ ITU: 国际电信联盟，是电信界最权威的标准制订机构。

⁴ KWIC: 题内关键词索引，是最早出现的机编索引。

- 便于各临床试验的药物安全性数据共享。
- 方便元数据（Meta Data）的存储和监管部门的视察，为不同系统和运用程序之间数据的整合提供统一的技术标准。
- 为评审机构提供方便，从而缩短审批周期。
- 有助于数据质量的提升，可以更快地提供更高质量的数据。

4.2.1 CDISC 和 HL7

CDISC（Clinical Data Interchange Standards Consortium）是一个全球的、开放的、多学科的非盈利性组织，它建立了一系列的标准用于收集、交换、提交和归档临床研究数据及元数据，使得不同临床研究间的数据可以方便地进行交换与共享。CDISC公布的标准见下表。

标准	描述
研究数据制表模型（SDTM）	有关临床研究项目病例报告表数据表格，用于向监管部门递交的内容标准。
分析数据模型（ADaM）	有关分析数据集及相关文件，用于向监管部门递交的内容标准。
操作数据模型（ODM）	基于XML，用于获取、交换、报告和递交，以及对基于病例报告表的临床研究数据归档的内容和格式标准。
实验室数据模型（LAB）	用于在临床实验室和研究申办者/CRO间进行数据转移的内容和格式标准。
病例报告表表格数据定义规范	基于XML的内容和格式标准，用于

(CRTDDS) -(Define.xml)	CDISC SDTM数据集数据定义的规范文件，提供给FDA参考。该标准也称为Define.xml，是ODM的拓展。
非临床数据交换标准 (SEND)	SDTM的拓展标准，用于递交临床前研究的数据。
方案表述 (PR)	用于支持临床研究方案信息交换的内容和格式标准。该部分与HL7联合制订。
试验设计模型 (TDM)	关于事件的计划顺序和试验处理计划的结构定义的内容标准。
临床数据获取的标准协调 (CDASH)	以CDISC为指导，联合开发的，用于病例报告表中基础数据收集字段的内容标准。该标准基于SDTM。
专业术语	全部CDISC模型/标准所涉及的标准词汇和编码集。
词汇表	CDISC词典，用于解释与临床研究信息电子获取、交换、报告相关的术语及其定义。简称和缩写也列在其中。

从上表可以看出，CDISC标准集从临床研究的方案设计开始，覆盖数据采集、分析、交换、提交等环节，为整个临床研究过程提供标准化的规范参考。尽管它们仍在不断地被测试、更新和完善，CDISC的标准化理念与基本构架对于我国的临床试验具有极大的借鉴价值。对于我国临床试验数据标准化工作，可以参照已有的CDISC标准并根据中国的实际状况来逐步建立。

HL7是美国国家标准局授权的标准开发组织之一，是从事医疗服务信息传输协议及标准研究和开发的非盈利组织。HL7的目的是开发和研制医院数据信息传输协议及标准，优化临床及其

管理数据信息程序，降低医院信息系统互连成本，提高医院信息系统之间数据信息共享程度。为此，HL7制定了一套标准化卫生信息传输协议，便于医疗领域不同应用之间的电子数据传输，将允许各个医疗机构在异构系统之间进行数据交互。也就是说，HL7实现了医疗数据的共享与交换，CDISC制定的标准实现了临床研究数据的共享与交换。近年来，将医疗信息与临床研究数据集成与共享，已成为临床研究数据管理的一种趋势。但当基于HL7标准的医疗数据要转化为基于CDISC标准的临床研究数据时，就需要一种转换标准。

CDISC研发了一种域分析模型BRIDG（Biomedical Research Integrated Domain Group），旨在使其制定的临床研究数据标准可以与HL7进行交换。BRIDG于2005年被HL7规范化临床研究信息管理技术委员会所采纳，它是CDISC与HL7间的桥梁，正逐渐发展成为一套独立的、用于在医疗信息与临床研究信息间进行交换的国际标准。

4.2.2 医学术语标准

（1）WHOART 术语集

WHOART是一个精确度较高的用于编码与药物治疗过程中的临床信息的术语集，主要在参加WHO药品监测项目的成员国中免费使用，世界上的制药企业和临床研究机构也在使用。

WHOART术语集自开发30多年以来一直是不良反应术语合理编码的基础。

WHOART涵盖了几乎所有的在不良反应报告中所需的医学术语，但仍小而精干，可以以行列表的形式打印出来。由于新药和新的适应证会产生新的不良反应术语，术语集的结构是灵

活可变的，允许在保留术语集结构的基础上纳入新的术语，同时又可以不丢失之前术语间的关系。

WHOART包含4级术语，分别是系统器官分类（System Organ Class, SOC）、高级术语（High level term, HT）、首选术语（Preferred Term, PT）和收录术语（Included terms, IT）。

（2）MedDRA

MedDRA是在ICH主办下创建的国际医学术语集。MedDRA用于医疗产品整个研发与应用周期的行政管理，对医学信息进行分类、检索、报告与信息交流。ICH于1997年在英国医药管理局开发的MedDRA测试版的基础上，经各方协调一致，发行了MedDRA 2.0 的执行版本，指定了MedDRA 的维护与客户服务机构（MSSO）在ICH 管理委员会（management board, MB）的监督与指导下工作。该机构与国际制药工业协会联合会（IFPMA）签约负责MedDRA 的维护、发行以及进一步开发。MedDRA 现每半年更新一次（每年的3 月和9 月）。

MedDRA作为新药注册用医学词典，适用于政府注册管辖下所有的医疗和诊断产品的安全报告。在新药注册环节中需要用到MedDRA 的有临床研究、不良反应的自发性报告、注册报告、受政府注册管理的产品信息。MedDRA 用于药品上市后不良反应监测中，用于药品不良反应的报告和数据分析等。

目前，美国、欧盟、日本、加拿大、澳大利亚等国家在其不良事件报告系统中使用MedDRA，其中欧盟、日本等还要求制药企业在提交不良反应报告中使用时MedDRA编码。

MedDRA包含5级术语，分别是系统器官分类（System Organ Class, SOC）、高级别组术语（High Level Group Term, HLGT）、

高级别术语 (High Level Term, HLT)、首选术语 (Preferred Term, PT) 和低级别术语 (Low Level Term, LLT)。

(3) 世界卫生组织药物词典 (WHO Drug词典)

世界卫生组织药物词典是医药产品方面最综合的电子词典，为WHO国际药物监测项目的重要组成部分。WHO Drug词典自1968年发布以来，已被制药公司、临床研究机构和药物监管部门广泛使用，用于编码和分析临床试验报告中的合并用药、上市后不良反应报告以及其他来源的报告中提及的药品。乌普萨拉监测中心 (UMC) 为WHO协作中心成员，负责该词典的维护及更新。

UMC提供的WHO Drug词典包括 4 种：世界卫生组织药物词典 (WHO DD)、世界卫生组织药物词典增强版 (WHO DDE)、世界卫生组织草药词典 (WHO HD) 和综合词典 (combined dictionary)。

WHO Drug词典采用解剖学治疗学及化学分类系统对药物进行分类。

(4) 解剖学治疗学及化学分类系统 (ATC, Anatomical Therapeutic Chemical)

ATC是世界卫生组织对药品的官方分类系统。ATC系统由世界卫生组织药物统计方法整合中心 (The WHO Collaborating Centre for Drug Statistics Methodology) 所制定，第一版在1976年发布。1996年，ATC系统成为国际标准。现在ATC系统已经发布2006版。

ATC分类系统根据药物作用的器官或系统、药物的治疗学、

药理学和化学特性，将药物分为5个级别，由字母与数字间隔而成，共有7位，其中：第一级由一位字母组成，表示解剖学上的分类；第二级由两位数字组成，表示治疗学上的分类；第三级由一位字母组成，表示药理学上的分类；第四级由一位字母组成，表示化学上的分类；第五级由两位数字组成，表示化合物上的分类。

(5) ICD-10

ICD-10，全称为“The International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10)”，即国际疾病伤害及死因分类标准第十版，是世界卫生组织（WHO）依据疾病的某些特征，按照规则将疾病分门别类，并用编码的方法来表示的系统。现有版本包括15.5万种代码，并记录多种新型诊断及预测，与ICD-9版本相比较，该版本增加了1.7万个代码。2010年WHO发布了最新的ICD-10更新版本。

WHO目前只提供4位编码的ICD-10。各国在引用的时候可以添加附加码来增加疾病数量。澳大利亚于1998年发布了首部5位编码的ICD-10AM。接着加拿大在2000年，法国在05年，泰国在07年，韩国在08年都出了自己的本地化修改版本。美国将在2013年10月正式启用6位编码的ICD-10。

根据WHO的规定，各国的本地化版本都可以对照转换成标准的ICD-10编码以便国际间交流。

ICD分类依据疾病的4个主要特征，即病因、部位、病理及临床表现（包括：症状体征、分期、分型、性别、年龄、急慢性发病时间等）。每一特性构成了一个分类标准，形成一个分类轴心，因此ICD是一个多轴心的分类系统。

当对一个特指的疾病名称赋予一个编码时，这个编码就是唯一的，且表示了特指疾病的本质和特征，以及它在分类里的上下左右联系。

4.2.3 临床试验报告的统一标准（CONSORT）

大量证据显示随机对照临床试验(RCT, Randomised Controlled Trial) 的报告质量不理想。报告不透明，则读者既不能评判试验结果是否真实可靠，也不能从中提取可用于系统综述的信息。最近的方法学分析表明，报告不充分和设计不合理与对治疗效果产生评价偏倚有关。这种系统误差对RCT 损害严重，而RCT正是以其能减少或避免偏倚而被视为评价干预措施的金标准。

为了提高RCT的报告质量，一个由临床试验专家、方法学专家和期刊编辑组成的工作组制定了临床试验报告的统一标准 (CONSORT, Consolidated Standards of Reporting Trials) 声明。CONSORT声明于1996年首次发表，2001年作了一次修订，方法学研究的不断扩充最终凝练成了CONSORT 2010声明。

CONSORT声明(或简称CONSORT)由报告RCT必备的基本项目清单和描述整个试验过程中受试者流程的流程图组成，主要针对的是两组平行设计的RCT报告。CONSORT中的大多数条目也与很多其他设计类型的临床试验相关，如非劣效性试验、等效性试验、析因设计试验、群组试验，以及交叉设计试验等。

CONSORT的目的是指导作者如何提高其临床试验报告的质量。临床试验的报告需要清晰、完整和透明。读者、审稿人和编辑还可以利用CONSORT来帮助评估和解释RCT报告。但CONSORT不是用于质量评价的工具，其内容更多地着眼于那些

与临床试验的内部和外部真实性相关的条目。

自1996年发表以来, CONSORT 已经得到400多种期刊和若干编辑组织如国际医学期刊编辑委员会的拥护。RCT报告质量的提高与期刊是否采用CONSORT有关。

CONSORT 2010声明、其说明与详述文件, 以及相关网站(www.consort-statement.org), 对于改进随机临床试验报告必将有所裨益。

五、 数据管理工作的主要内容

在进行临床试验数据管理之前, 必须由数据管理部门根据项目实际情况制定数据管理计划(DMP , Data Management Plan)。数据管理计划应包括以下内容和数据管理的一些时间点并明确相关人员职责。数据管理计划应由数据管理部门和申办方共同签署执行。

5.1 CRF 的设计与填写

5.1.1 CRF 的设计

临床试验主要依赖于CRF来收集试验过程中产生的各种临床试验数据。CRF的设计必须保证收集试验方案里要求的所有临床数据。CRF的设计、制作、批准和版本控制过程必须进行完整记录。

CRF的设计、修改及最后确认会涉及多方人员的参与, 包括申办者、申办者委托的CRO、研究者、数据管理和统计人员等。一般而言, CRF初稿由申办者或CRO完成, 但其修改与完善由上述各方共同参与, 最终定稿必须由申办者或申办者委托的CRO完成。

5.1.2 CRF 填写指南

CRF填写指南是根据研究方案对于关键字段和容易引发歧义的条目进行特定的填写说明。

CRF填写指南可以有不同的形式，并可以应用于不同类型的CRF或其他数据收集工具和方式。对于纸质CRF而言，CRF填写指南应作为CRF的一部分或一个单独的文档打印出来。对电子CRF或EDC系统而言，填写指南也可能是针对表格的说明，在线帮助系统，或是系统提示以及针对录入的数据产生的对话框。

保证临床试验中心在入选受试者之前获得CRF及其填写指南，并对临床试验中心相关工作人员进行方案、CRF填写和数据提交流程的培训，该过程需存档记录。

5.1.3 注释 CRF

注释CRF 是对空白的CRF的标注，记录CRF各数据项的位置及其在相对应的数据库中的变量名和编码。每一个CRF中的所有数据项都需要标注，不录入数据库的数据项则应标注为“不录入数据库”。注释CRF作为数据库与CRF之间的联系纽带，帮助数据管理员、统计人员、程序员和药物评审机构了解数据库。注释CRF可采用手工标注，也可采用电子化技术自动标注。

5.1.4 CRF 的填写

临床研究者必须根据原始资料信息准确、及时、完整、规范地填写CRF。CRF数据的修改必须遵照标准操作程序，保留修改痕迹。

5.2 数据库的设计

临床试验方案设计具有多样性，每个研究项目的数据收集

依赖于临床试验方案。临床试验数据库应保证完整性，并尽量依从标准数据库的结构与设置,包括变量的名称与定义。就特定的研究项目来说，数据库的建立应当以该项目的CRF为依据，数据集名称、变量名称、变量类型和变量规则等都应反映在注释CRF上。

数据库建立完成后，应进行数据库测试，并由数据管理负责人签署确认。

5.3 数据接收与录入

数据可以通过多种方式进行接收，如传真、邮寄、可追踪有保密措施的快递、监查员亲手传递、网络录入或其它电子方式。数据接收过程应有相应文件记录，以确认数据来源和是否接收。提交数据中心时应有程序保证受试者识别信息的盲态。

数据录入流程必须明确该试验的数据录入要求。一般使用的数据录入流程包括：双人双份录入，带手工复查的单人录入，和直接采用电子数据采集（EDC，Electronic Data Capture）方式。数据录入方式和采用时间的选择取决于资源技术水平。

5.4 数据核查

数据核查的目的是确保数据的有效性和正确性。在进行数据核查之前，应列出详细的数据核查计划，数据核查包括但不限于以下内容：

- 确定原始数据被正确、完整地录入到数据库中：检查缺失数据，查找并删除重复录入的数据，核对某些特定值的唯一性（如受试者ID）；
- 随机化核查：在随机对照试验中，检查入组随机化实施

情况；

- 违背方案核查：根据临床试验方案检查受试者入选/排除标准、试验用药计划及合并用药（或治疗）的规定等；
- 时间窗核查：核查入组、随访日期之间的顺序判断依从性情况；
- 逻辑核查：相应的事件之间的逻辑关联来识别可能存在的数据错误；
- 范围核查：识别在生理上不可能出现或者在研究人群的正常变化范围外的极端数值；

数据管理人员应对方案中规定的主要和次要有效性指标、关键的安全性指标进行充分的核查以确保这些数据的正确性和完整性。

数据核查应该是在未知试验分组情况下进行，数据质疑表内容应避免有偏差或诱导性的提问，诱导性的提问或强迫的回答会使试验的结果存有偏差。

数据核查可通过手动检查和电脑程序核查来实现。数据核查程序应当是多元的，每个临床研究人员有责任采用不同的工具从不同的角度参与数据库的疑问清理工作。

有时，在数据清理过程中无需研究中心批准，数据管理员便可对数据按照事先特许的规定进行修订，主要是指由具备一定资历的数据管理人员对于明显的拼写错误进行更正，或根据研究中心提供的计量单位进行常规的数值转换。这些数据清理惯例必须在数据管理计划中明确详细地列举，并明确地告知研究中心，同时保留可溯源性。

5.5 数据质疑表的管理

数据核查后产生的质疑表以电子或纸质文档的形式发送给申办方临床监察员由其整理并转交给研究者。研究者对疑问做出书面回答后，申办方临床监察员将已签字的质疑表复印件返回到数据管理部门。

数据管理员检查返回的质疑表后，根据质疑表对数据进行修改。质疑表中未被解决的质疑将以新的质疑表形式再次发出。

质疑表发送和返回过程将重复进行，直至数据疑问被清理干净。

数据管理部门保存质疑表电子版。由研究者签名的质疑表复印件待研究完成后连同CRF一起返还给申办方。

5.6 数据更改的存档

错误的的数据在数据清理过程中会被纠正。数据质疑表或数据核查文件作为数据更改的记录必须由研究者签名。在电话会议中认可的数据更改，应由批准更改的研究中心代表和数据中心代表同时签署讨论和批准的记录。

5.7 医学编码

临床试验中收集的病史、不良事件、伴随药物治疗建议使用标准的字典进行编码。编码的过程就是把从CRF上收集的描述与标准字典中的项目进行匹配的过程。医学编码员须具备临床医学知识及对标准字典的理解。临床试验中采用的字典应在研究方案中明确规定。CRF上出现的项目不能够直接与字典相匹配时，应当通过数据质疑表与研究者的沟通以获得更详细的信息来进行更确切的编码工作。医学编码应在锁库前完成。

广泛使用的标准字典有MedDRA， WHOART， ICD10等。数据管理部门应建立标准流程，适时更新字典并保证医学和药物编码在不同版本字典之间的一致性。临床研究使用的字典版本应储存在数据库里。

5.8 试验方案增补修改

试验方案增补修改是药物临床试验的一个有机组成。当有新的信息或者申办者/管理当局要求时就需要修改试验方案。不是所有的试验方案增补都需要变更CRF，需要制定相应的流程处理此种情况。纸质CRF修改后需要一段时间送达研究中心，到那时研究中心通常已经接收到机构/伦理审查委员会（IRB/IEC）对方案修订的批准。如果是使用EDC系统，修改eCRF是远程操作即时获得，所以对eCRF的修改必须在IRB 批准之后。

5.9 实验室及其他外部数据

在临床试验的组织实施过程中，有一些临床试验方案中规定采集，但是在研究者的研究基地以外获得的，由其他供应商（如中心实验室）提供的外部数据。外部数据类型比如：

- 生物样本分析数据：实验室数据、药代动力学/药效学数据、生物标记物的检测数据等。
- 外部仪器检测数据（如血生化、心电图、血流仪、生命体征监测、影像学检查等）
- 受试者的记录。

下列这些方面可能会影响外部数据的完整性，在建立数据库期间应注意：

- 关键变量的定义和必需内容

- 数据编辑和核查程序
- 记录格式和文件格式（例如，SAS、ASCII）
- 数据传输
- 数据库更新
- 数据储存和归档

为了确保有足够的信息可供用于外部数据的鉴别和处理，选择关键变量（唯一地描述每一个样本记录的数据）时必须谨慎。若无关键变量，将会对患者、样本和访视与结果记录的准确配对造成困难。

本地实验室数据的收集一般经过纸质CRF，需强调本地实验室使用的分析单位与正常值范围，重视对缺失数据、异常数据，以及重复数据等的检查。中心实验室数据的收集主要通过电子化的文件形式传输。在研究开始之前，数据管理员要为中心实验室制定一份详细的数据传输协议，对外部数据的结构、内容、传输方式、传输时间以及工作流程等作具体的技术要求。

数据管理员在外部数据加载入数据库前应对其进行检查。数据进入数据库后，启动逻辑检验程序，数据管理员也需要对其作相应的关联检查和医学审查，并对发现的问题启动质疑。

对于实验室和其他外部数据审查中发现的问题，临床研究监查员要对这些数据作100%的源数据核查。临床研究监查员对外部数据的监查与对研究中心的监查过程相同。

实验室数据的报告可采用实验室变化表和实验室数据分析图表等。

5.10 数据盲态审核

无论临床试验过程是开放或盲法操作，在临床试验数据库锁定前，应由申办方、研究者、数据管理人员和统计分析师在盲态下共同最终审核数据中未解决的问题，并按照临床试验方案进行统计分析人群划分、核查严重不良事件报告与处理情况记录等。

如双盲临床试验还需检查紧急揭盲信件和临床试验总盲底是否密封完好，如有紧急揭盲情况发生，需有紧急揭盲理由及处理报告。

5.11 数据库锁定

数据库锁定是临床研究过程中的一个重要里程碑。它是为防止对数据库文档进行无意或未授权的更改，而取消的数据库编辑权限。数据库锁定过程和时间应用明确的文档记录，对于双盲临床试验，数据库锁定后才可以揭盲。

5.11.1 数据库锁定清单

数据库锁定时，必须有证据显示数据库的数据编辑权限在定义好的时间点之前收回，并将这一证据记录在文件中。为了减少数据库锁定之后重新开启数据库的需要，应事先定义好一个有组织的程序，并且严格遵守这个程序，以保证完成所有数据处理，完成数据质量等级评定，通知了试验相关工作人员，并且所有相关人员批准锁定试验数据库。数据管理员应制定数据库锁定清单。数据库锁定清单应涵括以下内容：

- 所有的数据已经收到并正确录入数据库。
- 所有的数据质疑表已经解答并进入数据库。

- 非病例报告表数据（例如，中心实验室电子数据）已经合并到试验数据库中，并完成了与试验数据库的数据一致性核查。
- 已完成医学编码。
- 已完成最终的数据的逻辑性和一致性验证结果审查。
- 已完成最终的明显错误或异常的审查。
- 已完成数据质量审核，并将质量审核中发现的错误发生率纪录在文档中。
- 根据标准操作程序更新并保存了所有试验相关文档。

一旦完成上面所述步骤，就应书面批准数据库锁定，并由试验相关人员签名及签署日期，试验相关人员有：数据管理人员，生物统计师，临床监查员代表，研究者代表等。一旦获得数据库锁定的书面批准文件，就应收回数据库的数据编辑权限，并将收回数据编辑权限的日期记录在文档中。

针对期中分析，应严格按照方案中规定时间点或事件点进行分析，期中分析数据库锁定过程与最终分析的数据库锁定要求可能有所不同，但是所有数据库锁定的要求以及采取的步骤都应记录在文件中，还应报告截止至进行期中分析时的数据情况、时间情况及终点事件情况等。

5.11.2 数据库锁定后发现数据错误

如果数据库锁定后发现数据错误，应仔细的考虑处理并纪录这些错误数据。最重要的是，应评估这些数据错误对安全性分析和有效性分析的潜在影响。然而，并非所有发现的数据错误都必须更正数据库本身。数据错误也可以记录在统计分析

报告和临床报告文档中。尽管一些申办者选择更改发现的数据库中的所有错误，但一些申办者可能只更改对安全性/有效性分析有重要影响的数据错误。最重要的是，申办者应事先确定一个程序来决定应处理哪些数据错误和记录这些数据错误。

如果一个数据库锁定后又重新开锁，这个过程必须谨慎控制，仔细记录。重新开锁数据库的流程应包括通知项目团队，清晰地定义将更改哪些数据错误，更改原因以及更改日期，并且由主要研究者，数据管理人员和统计分析师共同签署。数据库的再次锁定应遵循和数据库首次锁定一样的通知/批准过程。

5.12 数据备份与恢复

在整个研究的数据管理过程中，应及时备份数据库。通常是在另外一台独立的计算机上进行备份，并根据工作进度每周对备份文件进行同步更新。最终数据集将以只读光盘形式备份，必要时，未锁定数据集也可进行光盘备份。

当数据库发生不可修复的损坏时，应使用最近一次备份的数据库进行恢复，并补充录入相应数据。

相关计算机必须具有相应的有效防病毒设置，包括防火墙、杀病毒软件等。

5.13 数据保存

数据保存的目的是保证数据的安全性、完整性和可及性（accessibility）。

保证数据的安全性主要是防止数据可能受到的物理破坏或毁损。在进行临床试验的过程中，把所有收集到的原始数据（如CRF和电子数据）存储在安全的地方，诸如受控的房间，保证

相应的温度、湿度，具有完善的消防措施，防火带锁文档柜。这些原始文档是追踪到原始数据的审核路径的一部分，应如同电子审核路径对数据库的任何修改或备份所做记录一样，严格进行保护。建议数据至少保存 10 年。

数据的内容及其被录入数据库的时间、录入者和数据在数据库中所有的修改历史都需要保存完整。保证数据的可及性是指用户需要时能够自如登录和获取数据，以及数据库中的数据可以按照需要及时传输。

在临床试验完成后，应对试验过程中的文档进行存档。下表中总结了临床试验数据归档保存的各类型信息：

归档内容	要求
临床试验数据	试验中收集的所有数据。这些数据既包括记录在病例报告表上的数据也包括非病例报告表收集的数据（例如实验室检查结果，心电图检查结果以及受试者电子日记）。
外部数据	外部收集并将导入至临床试验数据管理系统（CDMS）的数据，包括所有导入的数据及其文件和用于外部数据质量控制的所有文件。
数据库元数据信息	临床试验数据结构相关信息。这类典型信息是表、变量名、表单、访视和任何其他相关对象，也包括编码列表。
数据管理计划书	数据管理计划的微软Word或PowerPoint文档可以转成PDF格式文件或打印成纸张文件归档保存。
编码词典	如果数据是使用公司内词典或同义词表自动编码，那么使用的词典和统一词表都应归档保存。
实验室检查参考值范	实验室检查的参考值范围。如果临床试验研究过程

围	中使用多个版本的参考值范围 ,那么每个版本的参考值范围都应归档保存。
稽查轨迹	试验稽查轨迹的整个内容 , 并使用防修改的方式。
逻辑检验 , 衍生数据 变更控制列表	以工作清单、工作文件、工作报告的形式提供逻辑检验定义和衍生数据的算法 ,以及它们的变更控制记录。
数据质疑表	所有数据质疑表 ,传递数据质疑表的相关邮件及数据质疑表解答的复印件。纸张形式的数据质疑表可以扫描归档保存 , 并且为扫描文件添加索引。
程序代码	数据质量核查程序的代码 ,衍生数据的代码以及临床试验数据统计分析的程序代码。程序代码文档应归档保存。最理想情况是 ,这些文件以在线方式保存 , 并编制索引或超链接。
病例报告表的映像 PDF格式文件	对于纸张的病例报告表临床试验来说 ,CRF映像文件通常可以通过扫描方式获得 ,并将这些扫描文件转成PDF格式。对于电子数据采集的临床试验来说 ,电子表单的PDF格式映像文件可以通过EDC/M应用创建。
其他	其他与数据管理相关的文件

下表中列举不同类型的临床试验数据及其常用的归档格式。

格式	描述
CSV	以逗号为分隔符的ASCII文本文件 , 可以使用文本编辑器、文字处理器以及Excel电子表格软件编辑。
XML	以ASCII技术为基础 , 便于不同系统间结构化信息的转换。
SAS 版本5传输文件	SAS公司提供的开源格式文件。通常用来提交临床试验数据。

Adobe PDF	应用广泛的文本输出格式。
-----------	--------------

对于使用纸质病例报告表的临床试验，临床试验研究中心应保存一份纸质病例报告表的复印件。对于使用电子数据试验，临床试验数据管理系统的供应商应为临床研究中心提供一份所有电子病例报告表的PDF文件格式以备案。

5.14 数据保密及受试者个人私密性的保护

5.14.1 数据保密

数据保密是药物研发过程中必须遵守的基本原则，参与药物研发的机构应建立适当的程序保证数据库的保密性，包括建立及签署保密协议以规范相应人员的行为，以及建立保密系统以防止数据库的泄密。

5.14.2 受试者个人私密性的保护

临床试验受试者的个人私密性应得到充分的保护，受保护医疗信息包含：姓名、生日、单位、住址；身份证/驾照等证件号；电话号码、传真、电子邮件；医疗保险号、病历档案、账户；生物识别（指纹、视网膜、声音等）；照片；爱好、信仰等。

个人私密性的保护措施在设计数据库时就应在技术层面考虑，在不影响数据的完整性和不违反 GCP 原则的条件下尽可能不包括上述受保护医疗信息，比如：数据库不应包括受试者的全名，而应记录下全名的缩写。以中文姓名为例，应该采用该受试者姓的首字母和名字的首字母。

六、 数据质量的保障及评估

临床试验数据的质量不仅直接影响试验结果的客观性和可靠性，更关系到研究报告，以及整个临床研究的结论。建立和实施质量保障和评估措施对于保证临床试验数据的质量是非常关键的。

6.1 质量保障

质量保障需要确定组织机构，明确从事数据管理工作人员应该具备的资质要求、责任和权限等；质量保障必须具备一定的资源，包括人员、设备、设施、资金、技术和方法等；为了保证组织机构能按预定要求正常进行，标准操作程序的制定非常重要，它是数据管理人员工作的行为规范和准则，明确规定各项工作由哪个部门、团队或个人做，怎样做，使用何种方法做，在何种环境条件下做等；质量保障还应有机制确保它能被遵照执行、工作人员不执行规范或操作失控时得到警告，内部质量审核和稽查等是常用的机制，保证质量持续改善。

质量保障和改善来源于质量控制（QC, Quality Control）、质量保证（QA, Quality Assurance）和纠正预防措施（CAPA, Corrective Action and Preventive Action）等活动。

6.1.1 质量控制

ICH E6将质量控制定义为“质量保证系统内所采取的操作技术和活动，以查证与临床试验有关的活动都符合质量要求。”

临床试验数据的质量控制适用于数据处理的每一个方面，如临床研究中心、数据监查、计算机系统生命周期过程和数据的处理过程。

(1) 临床研究中心和质量控制

所有临床研究人员应具有资质并受到培训。制定质量控制程序，例如：

- 安全性：临床研究人员受到培训，且遵照权限管理程序。
- 设备：临床研究人员遵照程序确保设备和数据安全并适当储存。
- 数据私密性：确保遵照程序保护受试者私密性。
- 质量审核：临床研究人员对数据进行内部审核。
- 存储和归档：确保数据和文件存储归档。

(2) 监查和质量控制

临床数据监查是质量控制中最常考虑的环节，包括：

- CRF数据审核
- 电子数据完整性：确认电子数据是充分、完整和准确的。
- 程序化的数据核查：确认方案依从性、受试者安全性。
- 可溯源性
- 原始数据审核：确认原始文件完整以发现未报告数据（如不良事件）。
- 计算机系统的适当使用：确认工作人员受到培训，使用权限管理，且能正常使用计算机系统完成分配的任务。

(3) 计算机系统的生命周期过程和质量控制

如使用计算机系统，须让其满足试验和工作人员的需求。在系统生命周期的每一步都需执行质量控制，以确保所有要求都被记录、测试和满足。例如：

- 要求：确保系统的运行和维持涵盖了所有用户以及技术的、商业的和监管部门的要求。
- 系统验证过程：确保系统遵循确定的程序进行验证，且记录完整准确。
- 变更控制：系统的生命周期过程中所有的变更都需评估和测试。

(4) 数据处理过程和质量控制

通常从CRF的质量开始控制，考虑的因素包括：设计恰当、遵从方案、数据收集环境和培训等。质量控制核查举例：

- 数据录入系统
- 数据有效范围核查
- 逻辑核查
- 安全性核查

在数据管理中，数据管理员的两个不同工作性质决定了两种质量控制方式：过程质控（in-process QC）和实时在线质控（on-line QC）。

对于设计工作的质量控制，如CRF设计、数据库的设计以及逻辑检验的建立等，一般多采用过程质控的方法。过程质控提供产品的生产过程中的每一个阶段质量状况，以保证每一阶段的质量都是可靠的。例如，逻辑检验的质量控制就是通过录入不同的测试数据来检查该逻辑检验的计算机程序能否正确地捕捉到“问题”数据。如果不能，则该逻辑检验需要修改并再次测试，直到正确为止。当逻辑检验可以捕捉到“问题”数据后，该检验就可以进入生产环境。

临床试验进行阶段的质量控制，一般多采用实时在线质控。实时在线质控是计算某一时间点数据的错误率来评估数据的质量。例如，实时在线质控报告显示有3个受试者已经按计划完成了整个试验，但这些受试者的某一访视的实验室检查数据仍未录入。因此，要求数据管理员及时发现这些实验室数据并适时启动质疑机制。

6.1.2 质量保证

ICH E6将质量保证定义为：“为保证试验的进行和数据产生、留档（记录）以及报告都符合临床试验质量管理规范（GCP）和适用的监管要求所建立的所有有计划、成体系的行为。”

大部分申办者或CRO等都有独立的数据质量保证部门，其主要任务是建立质量管理体系，即制定质量方针、质量手册与计划、标准操作程序等，评估数据管理过程是否达到规定的要求，是否按程序执行，同时稽查数据质量。

（1）标准操作程序（SOP）

SOP是为达到均一性，完成一个特定职责而制定的详细书面说明。制定SOP的意义在于尽可能控制各种主、客观因素对临床试验结果的影响，尽可能降低临床试验的误差或偏差，并确保研究资料的真实可靠，以提高临床试验结果的质量。一般来说，数据管理的SOP均会包括以下内容：

- 数据管理计划
- CRF设计
- CRF填写指南
- 数据库的建立与设计
- 逻辑检验的建立

- CRF追踪
- 数据录入
- 数据核查与清理
- 外部电子数据的管理
- 医学编码
- SAE一致性核查
- 数据库的质量控制
- 数据库的锁定与解锁
- 数据的保存与归档
- 数据的安全性
- CRO的选择与管理
- 人员培训等。

SOP的建立应能覆盖数据管理的所有过程，但重要的是对所建立的SOP的遵守。SOP制订不会一步到位，需要在实践中不断地完善和发展。

(2) 稽查

申办者还应设立稽查部门，由不直接涉及试验的人员定期对质量体系的依从性进行系统性检查，以判定试验的执行、数据的记录、分析和报告是否与已批准的试验方案、SOP、以及GCP相一致，了解误解或错误发生的原因，并提出预防及改正的建议。数据管理稽查要求稽查人员不但要具有稽查的经验，而且要熟悉数据管理的过程以及相应的计算机程序，特别要熟悉监管部门对于临床试验数据的标准和要求。

数据管理稽查是评价整个数据管理的质量系统，它包括三个层次：

- 具有符合监管部门要求的数据管理SOP。
- 应提供书面文件记录对SOP的遵守情况（如遵守数据库锁定SOP时产生的过程记录）。
- 在以上基础上，还有其他客观的证据支持数据处理过程能够产生可信赖的高质量数据，并可用于统计分析和申报等。

对于临床试验数据的稽查，一般关注四个部分：研究文件、数据、统计分析数据、临床研究报告。

与数据管理稽查有关的文件主要有：数据管理员的简历和培训记录、数据管理各岗位的描述与要求、数据管理计划、接收CRF的记录、数据核查整理的记录和清单、数据库的变更控制记录、逻辑检验的变更控制记录等。

数据稽查的主要内容包括：CRF数据的缺失、CRF数据与数据库数据的差异等。医学编码直接影响到临床试验数据的报告，因而也是稽查的重点。

有些数据管理员也负责统计分析数据列表或汇总图表的建立和质量控制。稽查列表或汇总图表的主要内容有：图标的标题、注脚及其样式是否符合统计分析计划的要求；列表数据与CRF中的数据是否一致；汇总图表数据与列表数据是否一致等。

临床研究报告的稽查注重报告的书写和统计学分析均遵守研究方案和统计分析计划的要求，以及研究报告内容的准确性（所有数据与源文件一致）和报告格式的合理性（符合临床试验报告的要求）。

（3）纠正和预防措施（CAPA）系统

根本原因的分析以及纠正和预防措施是质量系统的基础，CAPA是质量持续改善的核心。

纠正措施是指针对已存在的不符合或不期望的现象，消除其根本原因所采取的措施，防止重复出现（Recurrence）。预防措施是指针对潜在的不符合或潜在不期望的现象，消除其原因所采取的措施，防止发生（Occurrence）。

深刻了解数据管理系统和数据管理工作过程有利于建立有效的CAPA系统，从而加强质量管理体系，保证数据管理所有过程的产出都是符合临床试验的目的，以及确保受试者安全以及数据的完整性。衡量CAPA系统内的某个系统或某个过程是否符合试验目的，需要全面了解数据管理相关的投入、产出、控制和资源等。对一个临床试验质量管理体系的有效性和效果的评估，包括定义相关的评价措施以及反馈。

6.2 质量评估

获得高质量的真实数据是临床试验数据管理的目的。良好的数据质量应该达到以下要求：

- ALCOA：可归因性(Attributable)，易读性(Legible)，同时性(Contemporaneous)，原始性(Original)，准确性(Accurate)。
- ALCOA+：持久性(Enduring)，完整性(Complete)，一致性(Consistent)。

数据完整、准确、真实和可靠是达到良好数据质量的最基本的要求。

评估数据质量的指标可以包括：

- 录入和报告数据的时间
- 监查员或稽查员确认有问题的观测的数量，或纠正的数量
- 解决质疑问题所需的时间
- CRF审核所需时间
- 数据错误的数量

临床试验中所收集的数据的错误必须尽可能少，使其能支持该临床试验得出的发现或结论。通过发现临床试验数据在转录、转移和处理中的错误，对数据质量进行定量，并评估其对临床试验结果正确性的影响是必要的。

错误的数字是指不能代表数据的真值。数据错误地来源或原因可以包括临床研究中心的抄写错误、数据处理错误、不明确的问题产生无意的错误，或要求的时间窗外进行的数据采集等。错误的根本原因包括理解错误、操作错误、管理不当、疏忽和欺骗等。

发现错误的主要方法有源数据核查、逻辑检验、数据核实、汇总统计、CRF与数据库核对等。

评估数据质量最常用的方法是计算错误数据的发生率，即错误率。错误率=发现的错误数 / 所检察的数据项总和。

对于CRF中关键指标核查，将对数据库进行100% 的复查，与CRF及疑问表进行核对，发现的所有错误将被更正。对于非关键指标的核查，如果总病例数大于100，将随机抽取10% 的病例进行复查；如果小于100例，则抽取例数为总病例数的平方根进行复查。将数据库与CRF及疑问表进行核对，可接受的错误率为：

数值变量不超过0.2%；文本变量不超过0.5%。如错误率超过此标准，将进行100%核对。

关键指标、非关键指标的界定，由研究者、申办者以及统计人员共同讨论决定。

七、 安全性数据及严重不良事件报告

临床试验的一个关键目标是，确定、研究、建立或证实一种研究产品的安全性特征。试验中安全性数据的管理和报告应该支持该目的。在临床试验数据的管理和报告中，安全性数据经常是最具挑战性的。

许多研究机构设立了数据安全和监查委员会（DSMB, Data Safety and Monitoring Board），对临床试验中安全性数据进行核查。该委员会由独立于实施研究、中期分析和数据核查的人员组成，可以设盲，也可以不设盲。数据监查委员会可由以下原因作出暂停正在实施的研究的建议或决定：（1）极其显著的疗效；（2）不可接受的安全性风险；（3）无效。该委员会也可以推荐对正在实施的研究进行更改，比如，下调剂量、去除出现不可接受的安全性风险的研究组别。

申办方的医学监查员也应对安全性数据进行检查。

7.1 不良事件的获取、管理和报告

临床研究中的安全性数据既是一种丰富的信息来源，同时其管理和报告也是最具挑战性的。临床不良事件经常含有临床研究中最重要的安全性信息。为了保证收集、编码、分析和报告的方法有利于获得可靠的结论，需要理解不良事件数据的特征及其局限。ICH的一些指导原则为工业如何管理和报告临床试验

安全性数据提供了指南，如E1A、E2A、E2B、E2C、E3、E5、E6和E9。

当对安全性数据进行获取、管理和报告时：

- 设计CRF需要对安全性数据的收集给予足够的重视。
- 对严重程度进行定义，理解其用途和局限性。
- 保证正常值范围正确地与实验室数据关联，在对不同研究中心实验室数据进行汇总时需要注意正常值范围是否相同。
- 在实验室数据的分析和报告中，对类别变化（状态从正常转化为异常）及其变化幅度均进行考虑。
- 不良事件数据获取的精确性与数据的分析和报告方法相关。

不良事件收集的精确度大致可分三种：

- 高度精确：受试者在隔离的环境里(如I期试验房间)，由护士或医生在床边持续的观察，不良事件发生和缓解的时间可以被精确地记录。
- 中度精确：医院里的医疗监护和观察经常但不连续，通常不良事件发生和缓解的日期可以准确地被记录。
- 低度精确：门诊病人的不良事件发生和缓解的日期准确度低，更相关的是记录事件的发生频率，例如单次发作，间歇发作，持续发作，和最严重的程度。

在临床试验中，不良事件数据的收集一般都是以开放方式由研究者填写的文字，应将它们按照标准字典或术语集进行归类，并编码。字典的选择要符合研究的目的，最好是具有代表

性，符合业界标准。字典的安装、维护和升级要有一套标准的操作程序。字典升级前应评价升级对现有临床试验数据编码的影响，同时提出解决方案。选择适当的编码方法与编码程序。编码者要有一定的临床知识，并获得相应的培训。编码后的数据需要进行质量控制检查。

可以应用很多方式对安全性数据进行展示和报告。为了保证充分地对属于药物反应的结果进行报告，需要判断和科学选择，以确定数据的趋势和显著特征。对药物反应进行辨别是安全性数据处理和报告的驱动目的。

7.2 实验室数据

实验室数据的特点需要在数据管理时进行考虑。数据的存储单位应能够清楚地反映数据的值；在许多数据库中，单位和数据是分开的。管理实验室数据最具挑战性的方面之一是连接数据到合适的正常值范围。在获得数据时，如果数据不是通过电子方式到达数据管理员那里，数据值和其合适的正常值范围的连接将花费很大精力。当正常值范围不可得到时，可以使用参考值范围，它是从正常值范围衍生得到的范围，可以从研究中或者从参考书中得到。然而，使用参考值代替正常值的文档对数据库用户必须是清楚的。

为了方便地对研究之间的数据进行连接，经常可以应用实验室数据的标准化技术，以达到这一目的。标准化技术一般包括，当数值为正常时，将数据转化为无单位的数值“0”和“1”，当数值低于正常范围的下限时，将数值转化为小于“0”，当数值高于正常范围上限时，将数值转化为大于“1”。实验室本身应该达到所在国家或地区规定的质控标准。

7.3 其他数据

除不良事件（AEs）和实验室值以外，安全性数据还有其它形式。专项检测数据（如心电图，脑电图）的采集需要对来自这些检测的常见数据及其格式、精度和特殊属性要有理解。

体格检查在临床试验中是很常见的。广义上讲，体格检查是一种筛查方法；如果一个非预期的、明显的临床异常在体格检查中被发现，通常利用专项检测来确定这个事件。在这种情况下，从专项检测中得到的数据具有更大的可靠性。

不鼓励使用开放式文字描述数据。对‘其他数据’的管理依赖于信息的形式。对于体格检查和专项检查，自由文本注释是允许的。文本框可通过文字处理而不是数据录入系统进行计算机化，从而与数据库连接，但不作为数据库本身的一部分。此时纠错方式是校对，而不是双录入。有些申办者为了避免将冗长的文本注释计算机化，而采用编码的方式，如 0=“无注释”，1=“有注释，不相关”，2=“有注释，相关”，3=“有注释，关键”。

7.4 严重不良事件数据

临床研究过程中，监管部门要求及时报告严重不良事件（SAEs）。一些公司通常设置专门小组处理和报告SAEs，此时报告的内容、格式及时限的要求与CRF填写可能不完全相同，这些安全性数据及其报告一般在药物警戒数据库中完成，该数据库通常包括各种不同来源的安全性数据，可能是不完整的、重复的、零碎的，或不准确的。而临床试验数据库中基于CRF填写的SAEs是经过严格的数据管理程序处理的，包括清理、质疑以及确保准确度的验证。两个数据库中SAEs的记录可能存在差异，

为确保SAEs数据的一致性必须对临床试验的数据库与药物警戒数据库的一致性进行核查。该核查必须在临床试验中定期进行。核查范围包括但不仅局限于下列内容：

- 方案
- 研究者
- 受试者代号(随机号)
- 受试者姓名缩写, 出生日期, 性别, 种族
- 严重不良事件的病例号
- 不良事件诊断
- 报告的严重不良事件名称
- 编码术语或首选术语
- 不良事件开始日期, 结束日期
- 死亡日期, 死亡原因及尸检结果
- 不良事件结局
- 不良事件严重程度
- 因不良事件对研究药物采取的措施, 伴随药物名称及开始和停止使用日期
- 申办者和研究者对不良事件进行评估, 评估内容应包含: 严重程度、因果关系、是否预期、有无破盲。对不良事件评估需使用标准词典编码术语。
- 不良事件是否报告给有关部门? 何时报告?

八、 参考文献

1. 国家食品药品监督管理局. 药物临床试验质量管理规范 (局令第3号) . 2003.
2. International Conference on Harmonisation. Guideline for Good Clinical Practice, E6. 1996.
- 3.FDA. Guidance for Industry: Computerized Systems Used in Clinical Investigations.2007.
- 4.FDA.Guidance for Industry:Part 11, Electronic Records; Electronic Signatures - Scope and Application.2003.
- 5.Society for Clinical Data Management. Good Clinical Data Management Practices(GCDMP).2007.
- 6.CDISC.Introducing the CDISC Standards: New Efficiencies for Medical Research.2009.
- 7.Drug Information Association(DIA). Computerized Systems in Clinical Research: Current Data Quality and Data Integrity Concepts. 2011.
8. 颜崇超. 医药临床研究中的数据管理. 北京:科学出版社, 2011.