

国内外临床数据管理系统发展现状

孙亚林,贺佳*,曹阳

(第二军医大学卫生勤务学系卫生统计学教研室,上海 200433)

[摘要] 临床数据管理是临床试验的关键性工作,对研究结果的准确可靠性起着决定性作用,本文论述了临床数据的质量标准,介绍了临床数据管理规范化发展中的重要法规和原则,回顾了国内外临床数据管理系统的发展及其现状,并提出了积极借鉴国际先进管理模式和引入国外规范化临床数据管理软件系统,以提高国内临床数据管理水平的建议。

[关键词] 临床试验;数据管理系统;临床数据

[中图分类号] R 197.32 **[文献标识码]** A **[文章编号]** 0258-879X(2006)07-0721-05

Development of clinical data management system: current status home and abroad

SUN Ya-lin, HE Jia*, CAO Yang (Department of Health Statistics, Faculty of Health Services, Second Military Medical University, Shanghai 200433, China)

[ABSTRACT] Clinical data management, a critical part of clinical trials, plays a decisive role in assuring the accuracy and reliability of the trial results. This article discusses the quality standards of clinical data, introduces the important rules and principles for standardized clinical data management, and reviews the development and current status of clinical data management home and abroad. It is proposed that we should learn the advanced clinical data management modes from international partners and introduce standardized clinical management software from abroad, so as to improve the clinical data management in China.

[KEY WORDS] clinical trials; database management systems; clinical data

[Acad J Sec Mil Med Univ, 2006, 27(7): 721-725]

临床试验是探索新药、新技术、新疗法安全性和有效性、提高医学科学技术发展水平的重要手段,为了确保临床试验结果的准确可靠、科学可信,保护受试者权益和广大人民的用药安全,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,用以规范临床试验的整个流程。伴随着临床试验规则的制定,作为评价临床试验结果基础的临床数据同样也面临着规范化的发展趋势。为了实现临床数据的规范管理,一方面需要相关法规政策的规范和指导以及相关政府机构的监督检查;另一方面,在电子、信息科技日益发达的当代社会,临床数据管理系统对于实现高效规范数据管理起着至关重要的作用。近年来,临床数据的电子化规范管理已经成为了临床试验领域研究的重要课题,目前也已经形成了一些较为成熟的指导原则和数据管理系统。以下主要从临床数据的质量要求、临床数据管理指导原则和数据管理软件系统三个方面进行介绍。

1 临床试验数据的质量标准

1.1 数据具有可溯源性^[1] 临床研究中所有的数据都应当能够按照一定的逻辑标记对应到具体的受试者(受调查者)、研究者,各种施加于患者的干预都应当被详细的记录。

1.2 数据应当具有原始性 临床试验中所获得数据应当真正的来源于临床研究者最初调查时所获的资料,任何记录于病例报告表中的数据都应当为原始的未经研究者加工处理过的数据。

1.3 数据必须真实 所有得自临床研究的数据都必须如实反映研究的实际状况,不得有任何伪造、虚构和篡改。

1.4 数据应当准确 由病例报告表(CRF)及研究数据库等所记录和录入的数据应当保持与实际研究中所获得数据的一致性,应有相应的核查措施以确保所有记录或录入的错误能被及时改正,保证最终用于分析数据的准确性。

1.5 数据要保持完整性 应当确保每一位受试者的所有有关研究的数据都被完整地收集。因特殊原因数据缺失的应标明缺失原因,同时所有有关数据填写、修改等活动都应当详细地记录在案,以形成一系列可查的核查轨迹。

1.6 引用的数据应当得到证实 临床研究的开展具有延续性,当前的临床研究往往建立在先前的研究之上,特别是对于某些罕见病或是开展大规模临床研究

[作者简介] 孙亚林,硕士。

* Corresponding author. E-mail: hejia@smmu.edu.cn

比较困难的疾病,常常需要综合先前工作中的一系列数据。在进行研究时必须通过文献评价等方式对这些数据可靠性进行证实,保证所利用数据的真实性。

为了保证临床试验中获取的数据能够达到上述的标准,当前各研究机构都应当按照 GCP 的规范,制定相应的标准操作规程(standard operation procedure, SOP),并以此作为依据实施全程的临床数据管理。同时该 SOP 也是监管机构考查各研究单位工作质量重要依据。然而,在传统的纸张型 CRF 的研究模式下, SOP 中所需规定的内容较为烦琐,同时也很难形成标准、统一、规范的数据管理规程,因此在具体的实施中存在较多的问题。但,随着计算机信息技术在临床试验数据管理领域的应用及不断发展和完善,数据管理的效率得到了快速的提高,同时也使较多先前难以执行的操作规程的实现成为可能,使临床试验的数据管理走向了电子化规范管理的时代,较好地推动了临床试验的健康发展。

而为了保证 GCP 原则的切实贯彻与落实,对于这些数据管理系统而言,同样需要相应的法规和指导原则进行规范。

2 临床数据管理系统应符合的标准

临床数据管理系统的开发首先应当符合 ICH Good Clinical Practice(ICH-GCP)^[2]; Consolidated Guideline 的规范,同时还应当符合各国相关法规的要求,在各国的规范中,以美国 FDA 的 21CFR Part 11 Rule (March 20, 1997)^[3]和 FDA 的关于“临床试验中采用计算机系统的行业指导规范”(Guidance for Industry “Computerized Systems Used In Clinical Trials, May 10, 1999)^[4]这两部法规具有代表性。同时,由各国临床试验方面的学者和专家组成的临床数据管理社团(Society of Clinical Data Management, SCDM)经过长期的研究、探索和讨论,形成了一部非官方的 GCDMP (Good Clinical Data Management Practice)^[5],为临床数据管理的国际标准化和规范化提供了依据和指导。

2.1 ICH-GCP 对临床数据管理的原则性指导

ICH-GCP 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数据管理的要求,但其对临床试验开展过程中对研究者、研制厂商以及有关试验方案、随机化记录、数据核查等的规定都直接或间接的对临床试验中数据的要求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、有关受试者的个人信息和人身权益受到良好的保护。

2.2 FDA 的 21CFR Part 11 Rule 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 5 月 FDA 出台了有关电子记录和电子签名的法规 21CFR Part 11 Rule。在该法规的制定过程中, FDA 成立了专门“电子身份识别和签名”任务组, 并广泛向社会相关领域的专业人士, 如制药企业、临床研究人员、软件开发人员等征求意见, 保证了法规的权威性和可行性。

21CFR Part 11 Rule 的主导原则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、电子签名和对于电子记录进行的手写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 能够接受电子化各类临床研究材料。

2.3 FDA 的 Guidance for Industry “Computerized Systems Used In Clinical Trials” 为了确保临床试验中在临床研究场所使用计算机化系统收集的数据具有符合 FDA 申报要求的高质量和完整性 (high quality and integrity), FDA 在 1999 年 5 月颁布了“临床试验中采用计算机系统的行业指导规范”。该规范适用于三种形式的数据: 记录于各种硬性媒质 (如纸张) 原始数据, 而后录入计算机系统的数据、直接录入计算机系统的数据和由计算机系统自动产生的数据。规范中对计算机系统的特征、电子 CRF、核查轨迹、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。

2.4 SCDM 的 Good Clinical Data Management Practice 临床数据管理社团是一个为了教育与科学, 以及发展临床数据管理这一学科, 提高临床数据管理水平为目的非营利性专业团体。其制定 GCDMP 是为了给现行的规章和指导性文件尚未涉及的临床数据管理方面提供有关可接受的操作流程。其意图是使数据管理的整个过程保持与现行规章制度的一致性。可以说是对现行的各项法规和指导原则的进一步细化和具体化, 为临床数据管理的整个流程提供了一套可供参照的系统模式。该规程可以服务于各类研究参与者如: 数据管理人员, 数据处理人员, 统计人员, 现场调查人员, 临床专业人员, 依从性审核人员, 管理机构的工作人员等所有的临床试验参与人员。该文件从 20 个章节给临床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规程的最低标准和最高规范。最低标准与最高的规范都以条目罗列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 是较为全面和细化的临床数据管理指导原则。

2.5 临床数据电子化规范管理参考规范的特点 (1)原则性强,具有普遍指导意义。对于临床数据管理中计算机系统的有关规定指明了临床研究中计算机系统的一般标准,没有规定具体应当采用的操作系统类型,软硬件的具体配置,给计算机系统的开发带来了比较广阔的空间。对于电子记录和电子签名的规定同样没有细化到具体采用形式的层面。(2)内容全面,涉及数据管理的各个环节。上述规范覆盖了硬件支持、软件设置、权限管理、核查轨迹、人员身份认证和培训、规程设置、系统维护、数据录入、存储、转换等各个方面,内容全面完整,为规范化临床数据管理软件系统的开发和电子化数据管理实施提供了十分全面具体的标准,具有很好的可操作性。(3)中心明确,重点突出。规范以提供高质量规范化的数据管理为中心,为了却确保数据的真实性、可靠性、完整性、私密性和安全性,重点对人员的权限管理、系统的核查轨迹、计算机系统的物理和逻辑安全性进行了详细的规定。指明了临床数据电子化管理和临床数据管理软件系统的开发的重点问题。

3 临床数据管理系统的发展

临床数据管理系统主要包括硬件系统和软件系统。就硬件系统而言,在计算机技术飞速发展的当今,不存在太多的问题。因而在临床数据电子化规范管理的进程中最主要的是开发规范化的临床数据管理软件系统,其发展伴随着计算机信息技术、特别是数据库管理技术的发展,临床数据管理软件系统主要经历了以下几个阶段。

3.1 层次与网状数据库发展阶段 数据库技术的主要功能是将数据进行集中存储,统一管理。在临床试验中引入数据库技术极大地提高了临床数据管理的效率。最早的数据库诞生于 20 世纪 60 年代,当时 IBM 推出的数据库管理的标志性产品 IMS (Information Management System),该数据库为基于层次模型数据库管理系统,并作为商品化软件投入市场。但由于 IMS 是将数据组织成层次的形式来管理,对非层次数据使用虚拟记录,大量指针的使用降低了数据使用的效率,同时数据库管理系统提供的数据库模型及数据库语言比较低级,数据的独立性也比较差,给使用带来了很大的局限性。为了克

服这些缺点,美国数据库系统语言协会 (Conference On Data System Language, CODASYL) 下属的数据库任务组 (DataBase Task Group, DBTG) 对数据库的方法和技术进行了系统研究,并提出了著名的 DBTG 报告。该报告确定并建立了数据库系统的许多基本概念、方法和技术,成为网状数据模型的典型技术代表。

3.2 关系数据库模型阶段 由于网络性数据模型存在的不足,数据库专家们接着开发出了关系型数据库 (Relational DataBase Management System, RDBMS) 模型^[6]。由于关系模型的简单易理解及其所具有的坚实理论基础,整个 20 世纪 70 年代开始的十几年时间内,数据库界集中围绕关系数据库进行了大量的研究和开发工作,为关系数据库概念的实用化打下了坚实的基础。20 世纪 80 年代以来,关系系统逐渐代替网状系统和层次系统而占领了市场。实践证明,由于关系模型具有严格的数学基础,概念清晰简单,数据独立性强,因此成为了当前数据库管理系统的主流^[7]。目前在临床试验数据管理领域应用的数据管理系统大多都是关系型数据库系统,如 Excel、Access、FoxPro 等等。

3.3 面向对象数据库模型 虽然关系数据库在当今的数据库技术中占主导地位,但由于关系模型不能用一张表表示出复杂对象的语义(如自然界的各种实体或各种行为),不擅长于数据类型较多、较复杂的领域(如多媒体领域声音、图片、视频等等),特别是对于临床试验中所要求的数据,除了记录简单的数据类型外,还需要记录操作的程序、图片、视频等信息。这使得数据库模型又进入了新的研究阶段——面向对象数据库的研究,面向对象数据库 (Object-Oriented DataBase, OODB)^[6]是指支持面向对象特性的数据库,它提供了面向对象的建模方法,编程语言和数据库语言。1989 年关于面向对象数据库的国际会议在东京举行,第一次定义了面向对象数据库管理系统所应实现的功能,包括了:支持复杂对象、支持对象标识、允许对象封装、支持类型或类、支持继承、避免过早绑定、计算性完整、可扩充、能记住数据位置、能管理非常大型的数据库、接收并发用户、能从软硬件失效中恢复、用简单的方法支持数据查询等,这些特性都非常适合临床试验数据管理的特点。

3.4 面向对象与关系模型的融合 面向对象数据库技术是一项新兴的技术,其发展远不如关系数据库成熟,其主要不足在于,没有统一的数据模式和形式化理论,因此缺少严格的数据逻辑基础。而且面向对象数据库技术目前尚无统一的标准,这些都是

制约面向对象数据库发展的关键因素。

而作为当代数据库技术主流的关系数据库管理系统 RDBMS 却恰好弥补了以上面向对象数据库技术的不足。随即出现了将面向对象技术引入关系数据库管理系统的对象关系数据库管理系统(Object Relational DataBase Management System, ORDBMS)^[7]。对象关系数据库集成了关系数据库技术的优点和面向对象技术的模型拓展能力,成为了当今数据库领域的重要发展方向。目前,市场上已经出现了许多商品化的对象数据库管理系统,如 Informix Universal Server, IBM DB2 Universal Database, Oracl 9i, UniSQL/X, OSMOby Unisys, Sybase Adaptive Server 等^[8]。但从总体上来说,对象关系数据库技术仍然需要进一步发展和完善,特别是在临床数据管理领域的专业数据库软件还并不多。

3.5 网络数据库技术 Web DB(网络数据库)是基于 Internet 和 Intranet 的数据库技术,其主要目的在于使用 Web 浏览器界面存取数据库内容,以 Web 资源充实本地数据系统,同时通过网络资源共享,实现数据远程利用和管理。一些厂商已开始扩展 DBMS 的数据类型,凡 Web 上有的数据类型都作为 DBMS 的内部数据类型,Web 页面、HTML、URL、图形图像都存储在同一个集成式数据库中,处理 Web 数据的机制都将成为 DBMS 的内部功能^[9]。同时网络技术的发展又将计算机系统从集中式的主机系统发展到客户/服务器系统乃至现在的分布式多层网络系统。这也是目前临床试验网络化发展的重要技术推动,特别是在全球新药研究合作日益深化的环境下,支持网络数据库技术的临床试验数据管理系统必将发挥重要作用。然而目前困扰数据库发展的问题主要集中在数据源的异构方面,包括了模式异构、数据异构和语义异构,即各种数据类型之间存在不兼容性、不同的数据源可能包含不同的数据类型以及同一语义在不同的数据库中表现为不同数据形式等等,因此,网络数据库管理技术仍有待于进一步的发展规范。

3.6 专业化的临床数据库管理软件系统 如今,许多研究者从实践的角度对数据库技术进行了研究,提出了适合各专业应用领域的数据库技术,如工程数据库、统计数据库、科学数据库、空间数据库、地理数据库等。这类数据库在原理上较传统的数据库也没有太大的变化,但是它们却与一定的应用相结合,从而加强了系统对有关应用的支撑能力,尤其表现在数据模型、语言、查询方面。

在临床试验领域,各国的数据库技术研究人员与

本国、国际的临床试验管理有关法规相结合,开发出了适应本国国情的临床试验数据管理系统,提高了临床试验的效率和数据管理的质量。比较有代表性的包括:较为成熟的 Oracle 公司的 Oracle Clinical^[10], Phase Forward 公司的 Clin Trial^[11], Nextphase 公司的 CT Series, Clinsource 公司的 Trial XS 和 StudyBuilder 公司的 StudyBuilder^[12]等。这些软件系统具有规范化的研究设计,电子 CRF 的设计,电子数据获取,全面的临床试验信息管理,在线受试者招募,基于 Web 技术的加密数据传输等功能,对于实现临床数据的规范化管理起到了极为关键的作用。

总的来说,国际的临床数据管理正日益地走向规范化,相应的临床数据管理软件系统也在不断发展和更新。目前国内的临床试验数据管理系统的开发和应用尚处于起步阶段,整体的数据管理也仍处于传统的纸张型数据管理模式阶段,电子化和规范化程度都较低。因此,积极的引入国外的临床数据管理经验和工具,对于借鉴学习、自主开发和规范提高有着重要的意义。因此一方面我们要充分认识国内临床数据管理上存在的不足和问题,积极学习先进国家的法规和做法,并适当引入国外符合规范的临床数据管理软件系统。另一方面目前以 SFDA(State Food and Drug Administration)为代表国家药品监管部门和国内的各制药企业、研发机构、研究单位、数据管理和统计分析部门都日益认识到了临床数据规范化管理的重要意义,有些单位甚至已经开发出了相应的数据管理软件并投入使用。相应地,在国家政府部门的重视下,通过国内临床试验各类参与者的共同努力,在不久的将来,中国的临床试验数据管理水平将赶上国际先进水平,国内的临床数据管理软件系统将得到国际标准的认可,并在临床数据管理中发挥关键性的作用。

[参考文献]

- [1] ICH-6. Good clinical practice[EB/OL]. (1997-04-06). <http://www.druggcp.net/regulation/detail>.
- [2] 马雪峰. 临床研究数据管理与分析系统[N]. 中国高新技术产业导报, 2003-08-07(8).
- [3] Food and Drug Administration. 21CFR PART11. Fedral register[EB/OL]. (2004-04-05). <http://www.21cfrpart11.com/files/library/compliance/isp.riskbasedapproach21cfr.pdf>.
- [4] Food and Drug Administration. Guidance for computerized systems used in clinical trials[EB/OL]. (1999-04-09). <http://www.fda.gov/ora/compliance.ref/bimo/ffinalcct.pdf>.
- [5] Good Clinical Data Management Practices Committee. Good clinical data management practices version 3[EB/OL]. (2003-09-06). <http://www.scdm.org>.
- [6] 王 珊, 楼文武. 数据库技术的发展[J]. 工程设计 CAD 及自动化, 1998, 3: 16-20.
- [7] 向海华. 数据库技术发展综述[J]. 现代情报, 2003, 12: 31-33.

- [8] 聂 哲. 数据库发展新技术[J]. 计算机时代, 2002, 12: 6-7.
- [9] 陈 楠. 面向对象数据库技术的产生与发展[J]. 计算机时代, 1997, 12: 9-10.
- [10] Oracle Clinical Co. An oracle white paper[EB/OL]. (2002-08-16). <http://oracle.ittoolbox.com/products.html>.
- [11] Phase Forward. Product overview[EB/OL]. (2004-06-08). [tp://www.phaseforward.com/products.html](http://www.phaseforward.com/products.html).
- [12] Study Builder Limited. Inside study builder[EB/OL]. (2003-01-05). <http://www.studybuilder.com>.
- [收稿日期] 2006-06-27 [修回日期] 2006-07-03
- [本文编辑] 尹 茶