

전자임상시험분야의 국제 표준

가톨릭대학교 의과대학 예방의학교실

정순옥, 최인영, 김석일

=Abstract=

International Standard in Electronic Clinical Trial

Sunok Jeong , Inyoung Choi, Sukil Kim

Department of Preventive Medicine, College of Medicine, The Catholic University, Seoul, Korea

As clinical trial researchers and companies continue to partner, merge, outsource, and make submissions for marketing approvals, the needs for streamlined data exchange increases. One study estimates that the industry spends more than \$156 million annually contract research organizations (CRO) to sponsors and from electronic data capture(EDC) providers or clinical laboratories to sponsors or CRO. Without the use of standards, the activities associated with clinical development require duplicating efforts and a lot of rework, which translates into more time and higher costs. It is in everyone's best interest to improve patient safety and reduce costs by bringing new treatments to market as quickly as possible, so the value of standardization is readily apparent. In 1997, The Clinical Data Interchange Standards Consortium (CDISC) established worldwide industry standards to support the electronic acquisition, exchange, submission and archiving of clinical trials data and metadata for medical and biopharmaceutical product development. CDISC has developed six standards. We specific describes CDISC Submission Tabulation Data Model (SDTM) in clinical trial. Finally, our own regulation should be developed for streamlined data interchange as soon as possible.

Key Words: e-Clinical Trials, Electronic Data Collection, Standard, Clinical Data Interchange Standards Consortium (CDISC).

서 론

최근 다기관(Multi-center) 및 다국적(Multi-nation) 임상시험이 점차 증가하고 있다. 다기관 임상시험을 진행할 때는 자료입력(data entry)뿐만 아니라 자료의 검증(data verification), 자료정리(data cleaning)등에 많은 시간이 소요되어, 효율적인 관리가 중요한 문제로 부각되고 있다¹⁾. 특히 기존의 임상시험이 종이 증례기록지(Case Report Form: CRF)에 자료를 기재하는 방식으로 진행이

되어 왔으나, 점차 정보기술의 발달에 따라 전자자료획득(Electronic Data Capture : EDC)이나 전자증례기록지(e-CRF)의 사용도 증가하고 있다. 이런 전자 증례기록지는 임상시험 과정을 효율적으로 수행할 수 있도록 해 준다. 그러나 임상시험을 주도하는 제약회사나 임상시험 수탁기관(Contract Research Organizations:CRO), 검사수탁기관, 정부 심사기관 등에서는 기존 방법을 정보화함에 따라, 각 기관마다 고유의 자료 형식을 사용하고 있다. 따라서, 임상시험 수탁기관은 연구 의뢰자의 요구

교신저자: 김 석 일
서울시 서초구 반포동 505 (137-701), 가톨릭대학교 의과대학 예방의학교실
전화번호: 02-590-1187, 팩스: 02-532-3820, E-mail: sikimmd@catholic.ac.kr

에 맞추어 매년 새로운 자료구조를 만들어 사용해야 한다. 또한, Electronic Data Capture (EDC)의 개발자는 자료 형식을 매년 수정하고, 자료관리자는 데이터베이스 개발을 위해 사전작업을 매년 실시하여야 한다. 또한, 연구 의뢰자는 수집한 자료가 원하는 형태와 맞는지 질 관리를 하고, 정부기관 심사위원들은 서로 다른 형식으로 제출한 자료를 수작업으로 심사한다²⁾.

이와 같이 임상시험 관계자들의 비효율적인 업무방식을 개선하고, 임상시험 관계자들-연구자, 연구 의뢰자 또는 CRO와 제약회사, 정부기관-간의 의사소통이 원활히 이루어지기 위해서는 자료구조가 일관되어야 하며, 이를 지원할 수 있는 자료 형식의 표준이 반드시 필요하다³⁾. 임상시험에서 동일한 자료모형을 사용하면 시간과 비용을 절감할 수 있고 임상시험 관계자들의 업무 효율을 증가시킬 수 있다⁴⁾.

이와 같은 요구를 해소하기 위해 1997년 임상시험 분야 표준 개발기구인 Clinical Data Interchange Standards Consortium(CDISC)가 설립되었다⁵⁾. CDISC는 모든 임상 시험 자료를 전자 문서화하고, 이들 전자 문서의 생성/교환, 제출, 및 저

장 방법을 지원하기 위한 표준을 정의하였다⁶⁾. 특히, 임상 시험 자료를 수집 및 저장하고, 그 자료를 처리하여 신약의 승인 여부를 판단하는 정부기관으로 보고서를 제출하는 과정에 필요한 시간과 자원을 효율적으로 줄일 수 있는 표준 자료 형식을 정의하였다. CDISC는 프로토콜 개발에서부터 자료 제출 및 분석 등의 각 단계를 지원하는 임상시험의 전반적인 표준을 개발하고 있다. 프로토콜 개발은 Protocol Representation(PR), 데이터 수집은 Clinical Data Acquisition Standards Harmonization (CDASH), 임상시험에서 적용할 자료의 형식으로 Operation Data Model(ODM)을 XML로 표현되는 전자 문서로, 제출할 임상 시험 자료의 형식을 Study Data Tabulation Model(SDTM)로, 실험실 검사는 Laboratory Standards(LAB)으로, 제출할 임상 시험 자료의 형식을 정의한 SDTM(Study Data Tabulation Model)까지 총 6개의 표준을 제시하고 있다(Table 1).

프로토콜 개발에서부터 FDA(Food and Drug Administration)에 제출하는 임상시험의 프로세스에 CDISC 표준을 적용해 보면, 프로토콜 개발에는 PR(Protocol Representation)을, 임상시험 셋팅과

Table 1. CDISC Standards

Standard	Description	Implementation Version Release date
SDTM, SEND	SDTM : Study Data Tabulation Model SEND : Standard for Exchange of Nonclinical Data	2004*
ODM	Operational Data Model	2001*
Define.xml	Case Report Tabulation Data Definition Specification	2005*
LAB	Laboratory Standards-Content standard	2002
ADaM	Analysis Data Model	2004
Protocol Representation	Collaborative effort to develop machine-readable standard protocol with data layer	In progress
Terminology Codelists	Developing standard terminology to support all CDISC standards	2006(Pkg1) Pkg2 in progress
CDASH	Data acquisition(CRF) standards	In progress

*Specification referenced in FDA Final Guidance

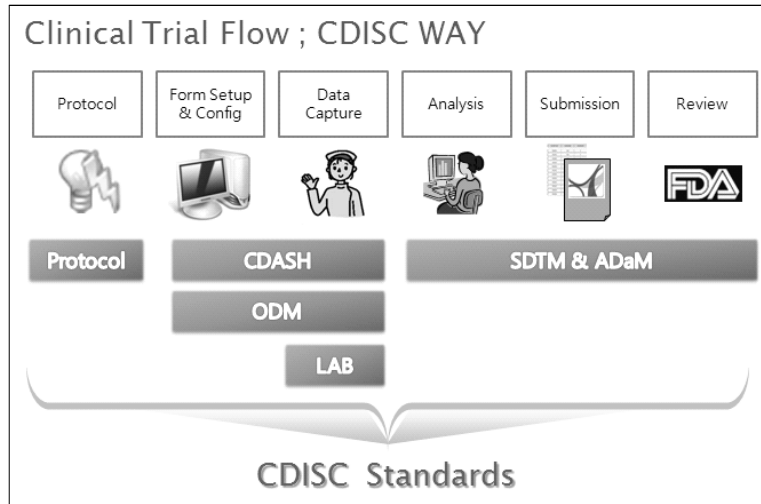


Figure 1. Clinical Trial Flow; CDISC WAY (www.CDISC.org)

데이터 수집하는 동안에는 CDASH, ODM, LAB 표준을 적용할 수 있다. 데이터 분석, 제출, 리뷰 단계에는 SDTM과 ADaM을 적용할 수 있다 (Figure 1).

1997년 미국FDA(Food and Drug Administration)에서는 표준에 근거한 임상시험 자료의 전자 접수를 허용하는 지침인21 CFR 11을 발표하였다⁷⁾. 임상자료의 전자 접수에 관한 지침 문서들은 자료 표준을 사용함으로써 임상자료의 검토를 용이하게 하였다. 2004년에는 XML 기반의 전자문서 (electronic common technical document : eCTD)를 제출할 수 있는 새로운 지침을 내놓았다. 그리고 임상시험의 제출 자료 형태를 CDISC Submission Data Tabulation Model (SDTM)으로 제출할 수 있도록 허용하였다.

2004년 CDISC의 표준 적용에 관한 조사 결과에 따르면, 데이터 제출 표준인 SDTM이 다른 표준에 비해 가장 많이 적용되었고, CDISC 표준 적용률이 2007년에는 2004년에 비해 30%정도 증가할 것이라고 예상했다 (Figure 2)⁵⁾.

이 연구 결과는 임상시험 분야에서 표준이 왜

필요하며, 표준을 적용했을때 효익이 무엇인지 잘 보여주고 있다. 따라서 이 논문에서는 임상시험 분야의 국제표준이 CDISC과 CDISC의 표준 중 가장 많이 적용되고 있는 SDTM을 살펴보고자 한다.

CDISC Submission Data Tabulation Model(SDTM)

SDTM은 임상시험의 데이터를 제출하기 위해 정의한 표준으로 도메인에 기반한 모델이다⁸⁾. 1999년부터 표준을 개발하기 시작하였고, 2006년 4월에 최종 승인되었다. SDTM을 어떻게 현장에 적용할 것인가를 설명하는Implementation Guide가 계속 발표되고 있는데 현재는 2005년 8월에 발표된 Implementation Guide version 3.1.1에 따라 제출 데이터를 작성하도록 권고하고 있다. SDTM은 임상시험 데이터의 특성에 따라 여러 개의 도메인으로 분류되어 있는데 여기에서는 SDTM 도메인과 도메인에 포함되는 데이터의 특성에 대하여 자세 히 알아보고자 한다.

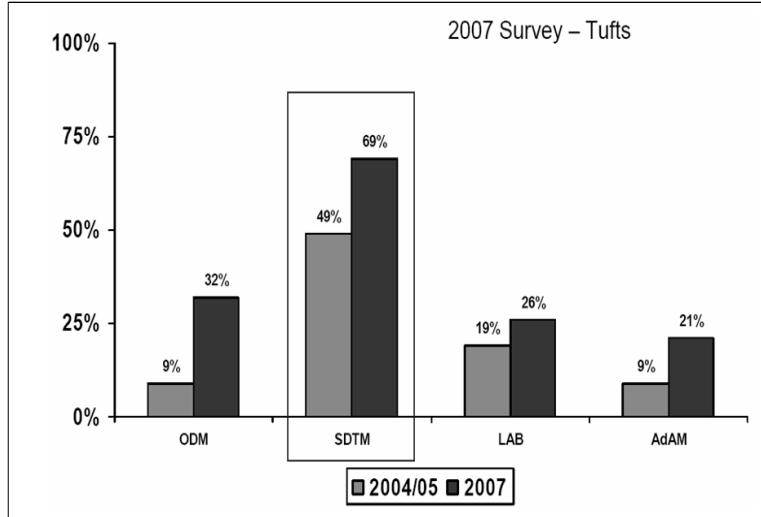


Figure 2. CDISC Standards Adoption (www.CDISC.org)

Table 2. CDSC Metadata

Attributes	Description
Variable Name	실험값 기술에 필요한 변수를 정의, 8자로 제한
Variable Label	변수명에 대한 상세 설명, 40자로 제한
Data Type	실험값에 대한 타입으로 Character나 numeric등이 있음
Controlled terminology	값이나 변수의 포맷을 표현할 때 사용하는 용어집
Origin of each variable	변수(항목)이 어디에서 유래된 것인지를 나타냄 CRF 에서 수집되거나 파생된 것
Role of the variable	변수명이 데이터 셋에서 어떻게 사용되는지 정해 놓은 것
Comments or other relevant information	코멘트나 Data와 관련된 정보를 나타냄

1. SDTM 도메인(Domain)

SDTM의 도메인은 임상시험 동안 수집되는 피험자에 대한 데이터들을 공통적인 특성을 가지고 있는 데이터들을 분류한 Dataset name이다. 각 도메인 dataset은 유일하고 두 글자 코드로 구성되어 있다. 모든 dataset은 flat file 구조로, 도메인의 변수들은 테이블의 열에 대응된다. 또한 CRF에서 수집된 하나의 실험값은 해당 테이블에서 하나의 행으로 표현된다. 각 dataset은 dataset안에 사용한

변수명에 관한 정보를 제공하는 meta data 정의에 의해 설명된다. Metadata는 'Define.xml'이라고 불리는 data definition document로 설명할 수 있다. Define.xml은 SDTM data를 설명하는 7개의 서로 다른 meta data 속성들을 일일이 열거한 것이다 (Table 2).

SDTM Domain은 임상시험을 진행하는 동안 피험자의 실험값을 모집하는 general observation classes와 특별한 목적을 가진 special-purpose datasets으로 나눌 수 있고, 현재 24 개가 정의되

어 있다. 앞으로도 도메인은 계속 추가될 예정이다⁴⁾.

General observation classes에는 Interventions, Events, 그리고 Findings이 있다. 일반도메인은 Figure 3와 같이 수집되는 데이터의 특성에 따라 Interventions, Events, 그리고 Findings로 다시 구분된다.

- Interventions class: 피험자에게 의도적으로 처치한 임상치료, 약물치료, 수술치료와 연관된 데이터로서, 연구 프로토콜에서 정의되거나(예: 시험약, Exposure) 연구 평가 과정에서 발생한 데이터(예: 병용약물, Concomitant Medications), 기타 피험자가 복용하고 있는 기호식품에 관한 데이터(예: 기호식품, Substance Use)가 속한다.

- Events class: 임상시험 동안 발생하는 예기치 않은 사건에 관한 데이터인 이상반응(Adverse Events)이나 시험종결(Disposition), 또는 임상시험 이전에 발생한 사건에 관한 데이터인 과거병력(Medical History)이 포함된다.

- Findings class: 신체검사, 결과, 심전도 검사, 설문조사와 같은 미리 계획된 세부적인 각종 검사 및 설문 데이터가 포함된다.

특수목적도메인 (Special-purpose domain)은 Demographics (DM 인구학적 조사), Comments (CO 코멘트), Subject Elements (SE)와 Subject Visits (SV)와 같이 피험자 단계의 도메인을 포함한다. 그리고 Trial Design Model (TDM) datasets은 Trial Arms (TA)과 Trial Elements (TE)가 속하며 피험자의 데이터가 아니라 임상시험 설계와 관련된 정보를 표현하는 도메인들이다.

SUPPQUAL 도메인은 일반용도 도메인에 정의되어 있지 않는 변수를 사용자가 추가로 정의할 수 있도록 하기 위한 도메인이다 (Figure 3).

2. 변수명(Variables)

도메인에 기록될 각각의 실험값은 도메인에서

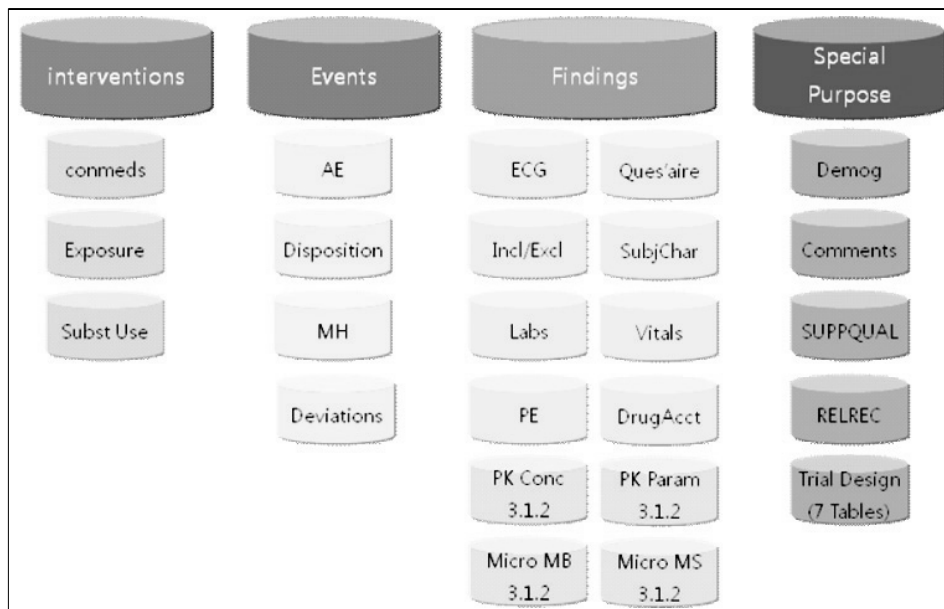


Figure 3. SDTM Domains

정의한 변수들로 기술된다. 변수명은 각각의 역할이 정해져 있어, 변수명이 각각의 실험값과 실험값을 어떻게 사용하는지에 관해 정확히 알려준다.

역할은 다섯 개로 분류된다. Identifier는 임상시험, 피험자, 도메인 그리고 레코드의 일련번호와 같이 임상시험을 구분할 수 있는 식별자의 역할이다.

Topic은 실험값의 초점을 두고 상세히 설명하는 것이고, Timing은 실험값의 시간을 설명한다. Qualifier는 결과가 문자값인지 수치값을 갖는지를 부가적으로 설명해 주는 것으로 단위나 부가적인 설명을 말한다. Role은 임상시험 설계 모델이나 임상시험 수행의 시작과 끝을 정의하는 알고리즘을 표현한다.

변수명은 8글자로 제한되어 있으며, 이는 통계 분석 프로그램인 SAS의 변수명이 8글자로 제한되어 있기 때문이다. 변수 레이블은 40자, 데이터는 200자로 내외이다.

변수명은 도메인을 의미하는 영문 두 글자로 포함하여 ‘도메인 이름+변수명’으로 구성된다. 예를 들면, Vital sign domain의 검사항목명은 VSTEST로 표현되며, 검사결과값은 LBTEST로 표현된다

(Table 3). 따라서, 변수명을 보면 그 값을 의미하는 바를 추론할 수 있다. 그러나, 특수목적도메인 중에서 인구학적 조사인 DM과 부가적 측정치인 SUPQUAL은 예외로서 변수명에 도메인명이 나타나지 않는다.

위에 제시된 변수들을 이용하여 임상시험 데이터를 Table 4와 같이 표현할 수 있다. 본 예제는 ABC123(STUDYID)이라는 연구에 피험자ABC-0001(USUBJID)에 관한 병용약물(CM)에 데이터를 측정된 것으로 데이터가 6개(CMSEQ) 수집된 것을 알 수 있다. 또한 이 피험자는 ASPIRIN (CMTRT)이라는 약물(명)을 100(CMDOSE)mg(CMDOSU)씩 2004년 1월 1일(CMSTDTC, CMENDTC)에 1번, 2일에 1번, 3일에 1번, 7일에 2번, 9일에 1번 복용하였음을 보여준다. CMDOSFRQ는 약을 먹는 주기에 대한 정보를 갖고 있다. ONCE는 한번 복용하는 것이고, QD는 Once a day로 시간에 관계없이 하루에 한번 복용하고, BID는 Twice a day로 시간에 관계없이 두 번 복용하고, PRN은 필요시에 따라 복용하는 것이다. CMDOSFRQ, EXDOSFRQ, SUDOSFRQ의 CODELIST에 관해서 CDISC TERMINOLOGY에 정의되어 있다. 이와 같이

Table 3. Domain ‘CM’ dataset

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMTRT	CMDOSE	CMDOSU	CMDOSFRQ	CMSTDTC	CMENDTC
1	ABC123	CM	ABC-0001	1	ASPIRIN	100	MG	ONCE	2004-01-01	2004-01-01
2	ABC123	CM	ABC-0001	2	ASPIRIN	100	MG	ONCE	2004-01-02	2004-01-02
3	ABC123	CM	ABC-0001	3	ASPIRIN	100	MG	ONCE	2004-01-03	2004-01-03
4	ABC123	CM	ABC-0001	4	ASPIRIN	100	MG	ONCE	2004-01-07	2004-01-07
5	ABC123	CM	ABC-0001	5	ASPIRIN	100	MG	ONCE	2004-01-07	2004-01-07
6	ABC123	CM	ABC-0001	6	ASPIRIN	100	MG	ONCE	2004-01-09	2004-01-09
7	ABC123	CM	ABC-0002	1	ASPIRIN	100	MG	QD	2004-01-01	2004-01-03
8	ABC123	CM	ABC-0002	2	ASPIRIN	100	MG	BID	2004-01-07	2004-01-07
9	ABC123	CM	ABC-0002	3	ASPIRIN	100	MG	QD	2004-01-09	2004-01-09
10	ABC123	CM	ABC-0003	1	ASPIRIN	100	MG	PRN	2004-01-01	2004-01-09

SDTM을 이용하여 데이터를 표현하면 임상시험 관계자간의 데이터에 관한 의사소통이 원활해짐은 물론 서로 다른 표현방식을 사용하여 데이터 입력이 완료된 후, 데이터 검토시 단위를 맞추지 않아도 되고, Query를 감소시킬 수 있다.

‘Core’ 변수는 Dataset에 필수 항목 여부를 표현하는 것으로 3개로 구분된다.

Required variable은 도메인에서 꼭 필요한 것으로 null을 허용하지 않는다.

Expected variable은 도메인에서 필요하나 일부는 null을 허용하고 값이 null일 수도 있다. 그러나 특별한 상황이나 예상하는 값이 원래 null인 경우는 그대로 표시한다.

Permissible variables은 도메인에 필요에 따라 포함된다⁸⁾.

결 론

임상시험의 비효율적인 업무 방식을 개선하고 시간과 비용을 절감하기 위해서는 정보기술과 표준의 적용은 필수 불가결하다. 이런 새로운 패러다임을 수용하기 위해서는 임상시험을 진행하는 스폰서나 병원, 또는 연구자들의 인식에 전환이 필요할 뿐만 아니라, 국제표준인 CDISC에 이해가 수반되어야 한다.

CDISC는 계속 진화하고 있으며, 2007년에는 더욱 빠른 속도로 표준을 제정하여 발표하고 있다. 또한, CDISC에서는 2010년까지 domain analysis model인 BRIDG (Biomedical Research Integrated Domain Group) Model을 미국 FDA와 HL7, NCI와 공동으로 개발 중에 있다⁹⁾.

CDISC 전자제출 표준(SDTM)은 표준 기반의 리뷰를통로써 FDA 리뷰 프로세스를 향상시키고, 통계 분석가들의 많은 요구들이 합의하는데 도움이 되고 있다. 또한 표준을 쉽게 배울 수 있고 방대한

지식이 없어도 데이터를 직접 검토할 수 있다.

향후에는, CDISC 표준들은 스폰서가 무슨 데이터를 수집할 것인가, 어떻게 데이터를 관리하고 최종 데이터 분석을 어떻게 할 것인가에 관한 명백한 가이드를 제시할 예정이다. 스폰서는 표준을 적용함으로써, 이익을 낼 수 있고 업무처리 방식이 최적화된 프로세스로 전환될 수 있다. FDA나 제약회사가 최대의 이익을 창출할 수 있도록 CDISC는 서서히 표준을 발전시켜 적용해 나갈 것이다⁶⁾.

최근 아시아지역 국가들은 신약 개발에 있어 앞서서 안 될 존재로 부상하고 있고 향후에는 다국적 프로젝트의 리더로 자리매김해 나갈 것이다. CenterWatch’s survey 결과, 아시아 국가들은 2005년에서 2006년 사이에 50%의 성장률을 보였다¹⁰⁾.

다국적 임상시험 수행에 있어 아시아 국가에도 기회가 증가하고 있으므로, 우리나라의 장점인 빠른 수행력과, 정보기술의 발달을 최대한 부각시키고 연구자의 질이나 사이트의 환경적인 요소, 언어 장벽을 보완한다면 다른 아시아 국가들에 비해 경쟁적 우위를 선점할 수 있을 것이다. 또한 다국적 임상시험 수행시 데이터를 수집할 때는 센터와 사이트간의 커뮤니케이션이 중요하므로, 가이드나 규정이 필요하다. 스폰서 자체 가이드도 있을 수 있지만, 국제표준을 적용한다면, 스폰서 입장에서는 매 스테디마다 가이드를 제정할 필요가 없을 뿐만 아니라, 커뮤니케이션도 효율적으로 이루어질 것이다. 우리나라의 사이트나 연구자도 사전에 국제표준인 CDISC을 이해한다면 다국적 임상시험 수행이 좀 더 수월해 질 수 있다.

Acknowledgements

본 연구는 보건복지부 보건의료기술진흥사업의 지원에 의하여 이루어진 것임 (A060093)

참 고 문 헌

1. Dimenas E, Johansson D, Palmblad M, Wrangstadh M. *Clinical operations online (COOL) -a World Wide Web-based approach to running clinical trials: Results from two international multicenter gastrointestinal trials.* *Drug Info J* 2001;35: 745-753.
2. Kush R. *Data Standards for Clinical Development-Progress Update.* *APPLIED CLINICAL TRIALS* 2002;11: 35-44.
3. CDISC. Available at http://www.fda.gov/cder/regulatory/ersr/2003_06_18_standards/sld010.htm
4. Kush RD, hardiso CD. *How Necessary are data standards?* *Scrip Magazine* 2004.
5. *Clinical Data Interchange Standards Consortium (CDISC).* Available at <http://www.cdisc.org>
6. Kubick WR, Ruberg S, Helton E. *Toward a Comprehensive CDISC Submission Data Standard.* *Drug Info J* 2007;41: 373-382.
7. *21 CFR Part 11: Electronic Records; Electronic Signatures.* Available at http://www.fda.gov/ora/compliance_ref/part11/
8. CDISC Standard. *Study Data Tabulation Model Implementation Guide: Human Clinical Trials. V3.1.1.* Available at <http://www.cdisc.org/models/sdtm/v1.1/index.html>
9. BRIDG Veneer Site. Available at <http://www.bridgproject.org/>
10. *Asia-Pacific Countries Set for Global Clinical Project Leadership.* In: *The CenterWatch Monthly; 2007.*