



2022

CHINA  
INTERCHANGE

29-30 JULY | VIRTUAL EVENT

## Developing CDISC-CRF of special disease to promote the ecology construction of real-world data in China

Presented by Xiaoxia Peng (彭晓霞)  
Center for Clinical Epidemiology and Evidence-Based Medicine  
Beijing Children's Hospital, Capital Medical University



## Xiaoxia Peng, PhD and MPH

Professor and Director of Center for Clinical Epidemiology and Evidence-Based Medicine, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health

- Principle Investigator for project of National Funding of Science and National
- Discipline leader of Beijing high-level talent construction project of public health technology (2022)
- Principle Investigator for establishing Chinese pediatric reference intervals

# Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *The author(s) have no real or apparent conflicts of interest to report.*

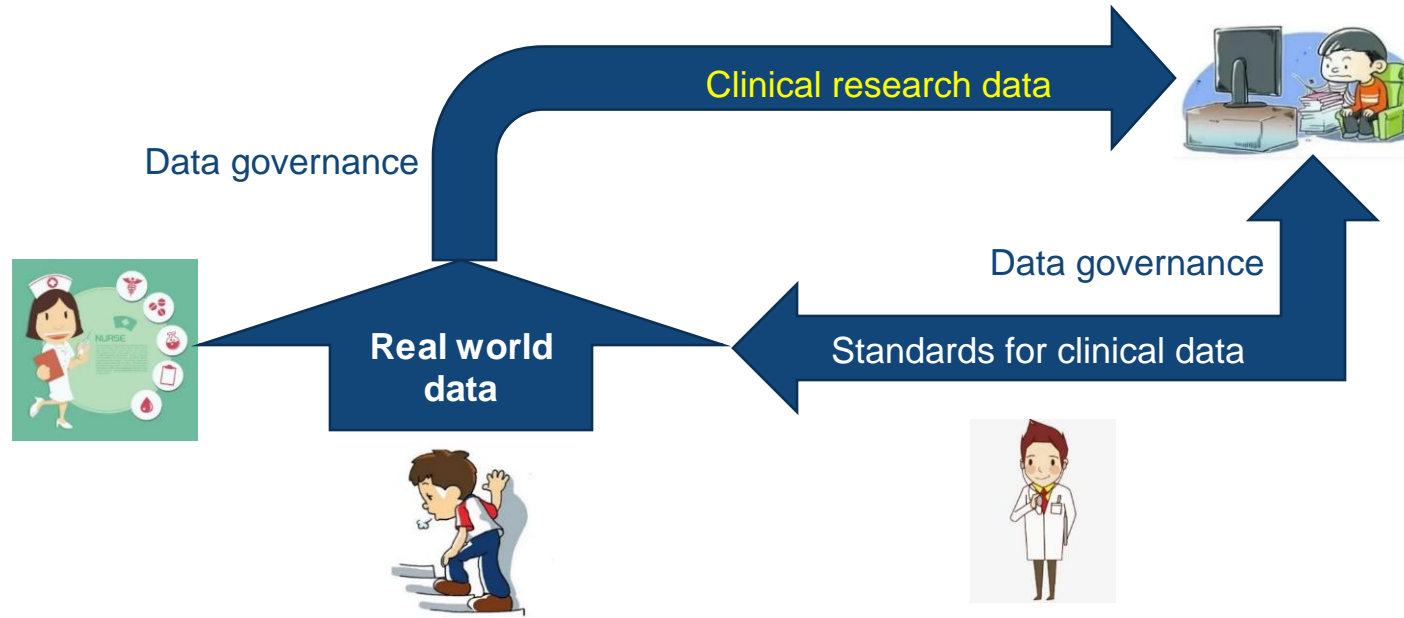


# Real-world data (RWD) and real-world evidence (RWE) are playing an increasing role in health care decisions

- FDA uses RWD and RWE to monitor postmarket safety and adverse events and to make **regulatory decisions**.
- The health care community is using these data to support coverage decisions and to **develop guidelines and decision support tools** for use in clinical practice.
- Medical product developers are using RWD and RWE to support clinical trial designs (e.g., large simple trials, pragmatic clinical trials) and observational studies to **generate innovative, new treatment approaches**.

US FDA. Real-World Evidence [EB/OL]. (2020-11-30) [2020-07-21]. <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>.

# Ecosystem for clinical research data



- The adoption of data standards is one of the cornerstones supporting the leverage of real world data

# Guideline for clinical trial data management in China

## 目 录

一、概述	1
二、真实世界数据来源及现状	2
(一) 真实世界数据常见的主要来源	2
(二) 真实世界数据应用面临的主要挑战	7
三、真实世界数据适用性评价	9
(一) 真实世界数据的数据治理和数据管理	9
(二) 源数据的适用性评价	10
(三) 经治理数据的适用性评价	11
四、真实世界数据治理	15
(一) 个人信息保护和数据安全性处理	16
(二) 数据提取	16
(三) 数据清洗	17
(四) 数据转化	18
(五) 数据传输和存储	18
(六) 数据质量控制	18
(七) 通用数据模型	19
(八) 真实世界数据治理计划书	21
五、真实世界数据的合规性、安全性与质量管理体系	22
(一) 数据合规性	22
(二) 数据安全管理体系	22
(三) 质量管理体系	23
六、与监管机构的沟通	23
参考文献	25
附录1 词汇表	27
附录2 中英文词汇对照表	30

## (四) 数据转化

数据转化是将经过数据清洗后原始数据的数据格式标准、医学术语、编码标准、衍生变量计算，按照分析数据库中对应标准进行统一转化为适用真实世界数据的过程。

对于自由文本数据的转化可使用可靠的自然语言处理算法，在保障数据转化准确、可溯源的前提下，提高转化效率。

在进行衍生变量计算时，应明确用于计算的原始数据变量及变量值、计算方法及衍生变量的定义，并进行时间戳管理，以保障数据的准确性和可追溯性。

National Medicine and Production Administration. Technical guide for clinical trial data management.  
<http://www.cde.org.cn/zdyz.do?method=largePage&id=20c3cd5ea30a51f9>

# Gaps from real world data to clinical research data

**Lack of data interoperability:** According to the respondents, different platforms or hospitals may use different data standards, and it is difficult to share data across information systems in general practice. The data stored in the EMR uses data coding and data storage schemes that are incompatible with the coding and storage formats of EDC systems. The lack of interoperability between different terminologies or coding schema used in information systems causes a great challenge in EMR data usage.

Different platforms or hospitals may use different data standards, so data cannot be used across settings. For me, the value of data from other systems is limited. - Intermediate doctor 114

**Data security concerns:** The information department staff reported that any private patient-related data should not be exported to clinicians unless they adhered to the full protocol to obtain access to data. These respondents believe that they will be responsible for data leaks, and they do not know how doctors will use the data. Thus, EMRs and EDC systems cannot communicate. Strong security concerns generate complicated approval procedures for data access, leading speciality departments to use an external company platform to manage their data.

EMR contains the private information of patients, and we will not allow doctors to download the data because the hospital does not have policies. Therefore, clinicians can only collect and transcribe the data. - IT (information technology) 313

**Unstructured EMR data:** EMR is used for clinical practice but not for clinical research. The doctors reported that there are many problems when using hospital EMRs for clinical research data collection. Many laboratory test or imaging test data are stored in PDF format or image files in EMR and must be manually transcribed.

The information recorded in the EMR includes pictures, such as ECG or laboratory examination. The results in these pictures cannot be directly presented in EMR. - Senior doctor 601



# How to solve this problem underlying cornerstones?

- Which standard for clinical research data?
  - CDISC or others?
- Where begin?
  - Information system of hospital?
  - Special disease?



# The framework of developing real world data based on CDISC

Disease selection

↓

Develop CRF forms according to consensus of diagnosis and treatment for special disease

↓

Develop CDISC-CRF

↓

Manual entry      Development of single disease RWD capture platform based on CDISC-CRF      Real time electronic capture

↓

Doctor check

↓

Artificial verification error correction      Data in doubt?      Artificial verification error correction

↓

Artificial verification of filling      Data missing?      Artificial verification of filling

↓

NO

NO

Generate CDISC data operation model

## Additional data except EMR

- Questionnaire investigation
- Scale assessment
- Quality of life assessment
- .....

## Real world data sources

- Electronic medical record systems
- Laboratory information system
- Prescribing information system
- medical record system
- .....



# Select applicable diseases

- For this disease, whether there are major problems that need further study?
- For this disease, is there a clinical practice guideline or consensus for diagnosis and treatment?
- For this disease, is the framework feasible?

## Diagnosis and treatment of Consensus for Neuroblastoma

附件 6

### 儿童神经母细胞瘤诊疗规范

(2019 年版)

#### 一、概述

神经母细胞瘤 (neuroblastoma, NB) 是婴幼儿最常见的颅外实体肿瘤, 占儿童恶性肿瘤的 8%~10%。NB 是一组临床表现及预后差异很大的疾病, 从肿瘤播散、转移、患儿死亡, 到肿瘤发展成熟为良性的节细胞神经瘤或自发消退等不同临床转归。NB 来源于未分化的交感神经节细胞, 故凡有胚胎性交感神经节细胞的部位, 都可发生肿瘤。肾上腺是最常见的原发部位, 其次是腹部交感神经节、胸部交感神经节、颈部交感神经节和盆腔交感神经节, 约 1% 的病人未能发现原发肿瘤。NB 可转移至淋巴结、骨髓、骨骼、硬脑膜、眼眶、肝脏和皮肤, 少数情况下也会转移至肺部和颅内。

儿童 NB 治疗难度大、单一的治疗预后差, 临床需要包括外科、内科、放疗科、移植科以及影像科、病理科、营养科、心理科、疼痛科等多学科的综合诊疗模式, 才能规范 NB 的诊治。

#### 二、适用范围

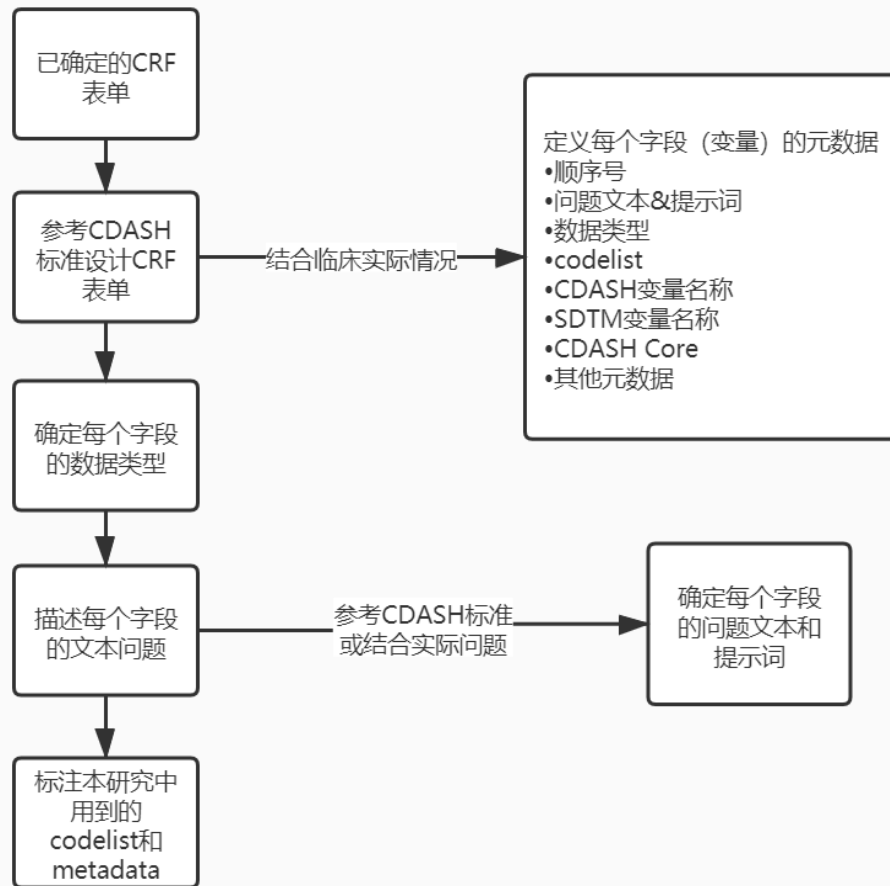
经肿瘤组织病理学确诊, 或经影像、骨髓、尿儿茶酚胺代谢产物等检查确诊的儿童 NB。

# Develop CRF forms according to diagnosis and treatment of consensus for Neuroblastoma

Domain name	Contents
Participants	Inclusion and exclusion criteria
	Informed consent
	Randomization
Basic information	Demography
	Medical history
	Vital signs
	Physical examinations
Diagnosis	Confirmed diagnosis of NB
Therapy	Medicine researched
	Drugs combined in therapy
	Non-drug therapy
	Drug delivery
	Surgery
	Radiotherapy
Follow-up	Visiting time
	Effective assessment of NB
	Survival follow-up
	Death information

	primary lesion
	metastatic lesion
	Adverse events
	Biomarks of NB
	Morphological examination of bone marrow
	Minimal bone marrow lesions
	MIBG scanning
	Electrocardiogram
	Echocardiography
	Hearing examination
	Myocardial enzyme
	Immunoglobulin
	Coagulation function
	CD
CRP	
End of research	Blood cell analysis
	Serum biochemistry
	Stool routine
End of therapy	Urine routine
	End of research
	End of therapy

# Developing CDISC-CRF forms

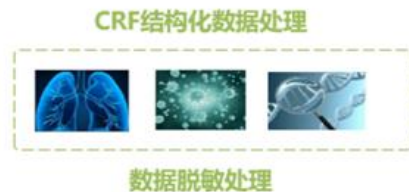
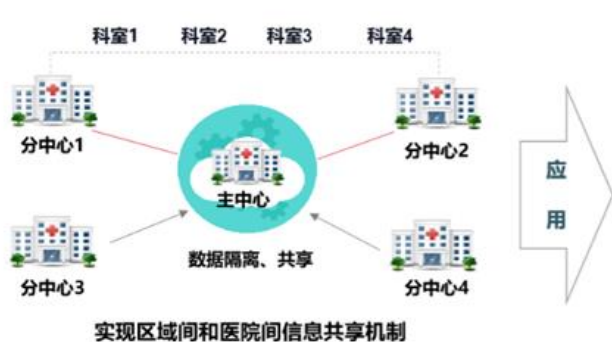
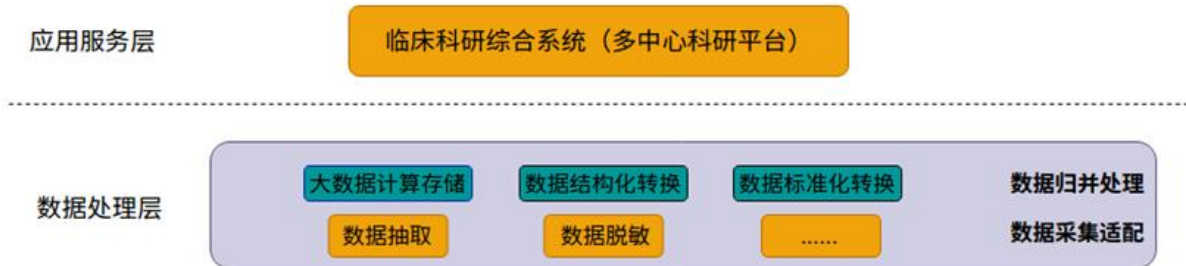


# Example of CDISC-CRF forms

{ }CDASH  
CORE

CDISC-CRF 模板, 以神经母细胞瘤诊断 (Diagnosis of Neuroblastoma (NB)) 为例	
What was the initial diagnosis date of NB? 病初确诊日期 [Initial diagnosis date] {O} <MHSTDAT> <MHSTDTC>	____/____/____(YYYY/MM/DD) < >SDTM Annotation
What was the category of initial diagnosis? 病初确诊类型 [Category of initial diagnosis] {O} <MHDIACAT>	(DIACAT) <input type="checkbox"/> Clinical diagnosis 临床诊断 <input type="checkbox"/> Pathologic diagnosis 病理确诊
What was the INRG stage at screening diagnosis? INRG 分期 [INRG Stage] {R/C} <MHSTAGE>	(STAGE) <input type="checkbox"/> L1 <input type="checkbox"/> L2 <input type="checkbox"/> M <input type="checkbox"/> MS ( ) Codelist name
What was the INRG group at screening diagnosis? INRG 分组 [INRG Group] {O} <MHRGROUP>	(RGROUP) <input type="checkbox"/> Very low risk 极低危 <input type="checkbox"/> Low risk 低危 <input type="checkbox"/> Medium risk 中危
What was the INSS stage at screening diagnosis? INSS 分期 [INSS Stage] {R/C} <MHSTAGE>	(STAGE) <input type="checkbox"/> Stage I <input type="checkbox"/> Stage II <input type="checkbox"/> Stage III <input type="checkbox"/> Stage IV <input type="checkbox"/> Stage IVS <input type="checkbox"/> Unknown [ ] prompt
What was the evidence of clinical diagnosis? 临床诊断依据 {O} <MHDIADTL> [Evidence of clinical diagnosis]	(DIATL) <input type="checkbox"/> Imaging findings 影像学 <input type="checkbox"/> Clinical feature 临床表现 <input type="checkbox"/> Increased catecholamine metabolism 血或尿儿茶酚胺代谢增高 <input type="checkbox"/> Bone marrow examination 骨髓检查 <input type="checkbox"/> Other, specify _____ 其他, 请说明 _____ {O} <MHDIADTO>

# Data capture platform for one disease based on CDISC-CRF



## 预期目标 (Expected Objectives)

病例覆盖广 (Wide Case Coverage)

数据类型全 (Complete Data Types)

科研方向深 (Deep Research Direction)

# Online test for the data platform

## 神经母细胞瘤专病队列数据采集平台

神经母细胞瘤专病队列数据采集平台旨在以标准化的数据采集表采集神经母细胞瘤临床数据。本数据采集平台参考神经母细胞瘤临床诊疗规范和临床数据交换标准协会 ( Clinical Data Interchange Standards Consortium, CDISC ) 标准开发而成。

CDISC是一个全球性的、开放的、多学科的非盈利性组织。该组织建立了一系列标准,用于收集、交换、提交和归档生物医学研究数据及元数据,从而提高数据质量,优化医药开发与研究流程,方便监管机构分析和评审各企业提交的数据,促进研究组织等更好地交换和分享信息。

本数据采集平台可以满足研究者、研究中心协调员、研究监督员、数据录入人员、医学编码员和统计员等用户的科研需求。此外,本平台还可以推动神经母细胞瘤的临床规范化,并极大推进不同医院之间的治疗决策分享与数据共享。

登录



请输入用户名



请输入密码

登录

[还没有账号?赶紧注册](#)

[忘记密码](#)

# A structured and standard real world data platform

首都医科大学附属北京儿童医院  
Beijing Children's Hospital, Capital Medical University

神经母细胞瘤专病队列数据采集平台

admin | 退出

多中心科研

课题研究

课题概况

CRF 管理

系统管理

课业研究

课题概况

病历录入

病历审核

成员管理

CRF取数

审核列表

数据审核

A 入选排除标准

B 知情同意

C 人口学资料

D 病史

E 生命体征

F 体格检查

G 神经母细胞瘤诊断

H 药物治疗相关

I 手术及放射治疗

J 访视周期

K 神经母细胞瘤治疗

C 人口学资料

查看修改记录

数据收集日期 [Collection date]

When was the date of data collection ?

{R/C}<DMDAT><DMDTC>

出生日期[date of birth]

What is the subject's date of birth ?

{R/C}<BRTHDAT><BRTHDTC>

出生时间[Birth time]

When was the subject's birth time ?

{O}<BRHTIM><BRTHDTC>

年龄[age]

What was the age of the subject ?

{O}<AGE><AGE>



# Paper and Computer Software's Registration

## 开发专病 CDISC-CRF 助力我国真实世界数据生态建设

彭晓霞<sup>1,2</sup> 苏健<sup>1</sup> 蔡思雨<sup>1</sup> 彭瑞玲<sup>3</sup> 彭亚光<sup>1</sup> 马晓莉<sup>1</sup> 倪鑫<sup>1</sup>

<sup>1</sup>国家儿童医学中心,首都医科大学附属北京儿童医院,临床流行病学与循证医学中心,北京 100045;<sup>2</sup>海南省真实世界数据研究院,海口 571435;<sup>3</sup>北京医普科诺科技有限公司 100089

通信作者:彭晓霞,Email:pengxiaoxia@bch.com.cn

**【摘要】** 真实世界数据是指在日常医疗活动、生活与工作等环境下产生的数据。真实世界数据在临床与公共卫生研究中一直被广泛应用,但真实世界数据质量相关问题,如不完整性和准确性等会影响真实世界研究的真实性。为了应对真实世界源数据缺乏标准化带来的挑战,本文基于当前被广泛应用的数据标准,即临床数据交换标准协会(CDISC)开发的 CDISC 标准开发了专病 CDISC-病例报告表(CRF),以提高真实世界源数据标准化水平,助力我国真实世界数据生态建设。我们阐述了如何应用数据标准弥补真实世界数据到真实世界证据之间的裂痕;设计了基于专病 CDISC-CRF 建设真实世界数据生态的流程,重点介绍了 CDISC-CRF 表单的开发技术;并就基于专病 CDISC-CRF 建设真实世界数据的应用前景及意义进行了描述。

**【关键词】** 真实世界数据; 临床数据交换标准协会标准; 病例报告表

**【基金项目】** 国家重点研发计划(2016YFC1000105);北京市医院管理局儿科学术协同发展专项(XTCX201812)



# Further goals

868

*e-Health – For Continuity of Care*  
C. Lovis et al. (Eds.)

© 2014 European Federation for Medical Informatics and IOS Press.

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

doi:10.3233/978-1-61499-432-9-868

## Interconnection of Electronic Medical Record with Clinical Data Management System by CDISC ODM

Yasushi MATSUMURA<sup>ab</sup>, Atsushi HATTORI<sup>b</sup>, Shiro MANABE<sup>ab</sup>, Toshihiro TAKEDA<sup>a</sup>, Daiyo TAKAHASHI<sup>b</sup>, Yuichiro YAMAMOTO<sup>ab</sup>, Taizo MURATA<sup>c</sup>, Naoki MIHARA<sup>a</sup>

<sup>a</sup>Medical Informatics, Osaka University Graduate School of Medicine, Osaka, Japan

<sup>b</sup>MKS Ltd, Osaka, Japan

<sup>c</sup>Division of Medical Informatics, Osaka University Hospital, Osaka, Japan

**Abstract.** EDC system has been used in the field of clinical research. The current EDC system does not connect with electronic medical record system (EMR), thus a medical staff has to transcribe the data in EMR to EDC system manually. This redundant process causes not only inefficiency but also human error. We developed an EDC system cooperating with EMR, in which the data required for a clinical research form (CRF) is transcribed automatically from EMR to electronic CRF (eCRF) and is sent via network. We call this system as “eCRF reporter”. The interface module of eCRF reporter can retrieve the data in EMR database including patient biography data, laboratory test data, prescription data and data entered by template in progress notes. The eCRF reporter also enables users to enter data directly to eCRF. The eCRF reporter generates CDISC ODM file and PDF which is a translated form of Clinical data in ODM. After storing eCRF in EMR, it is transferred via VPN to a clinical data management system (CDMS) which can receive the eCRF files and parse ODM. We started some clinical research by using this system. This system is expected to promote clinical research efficiency and strictness.

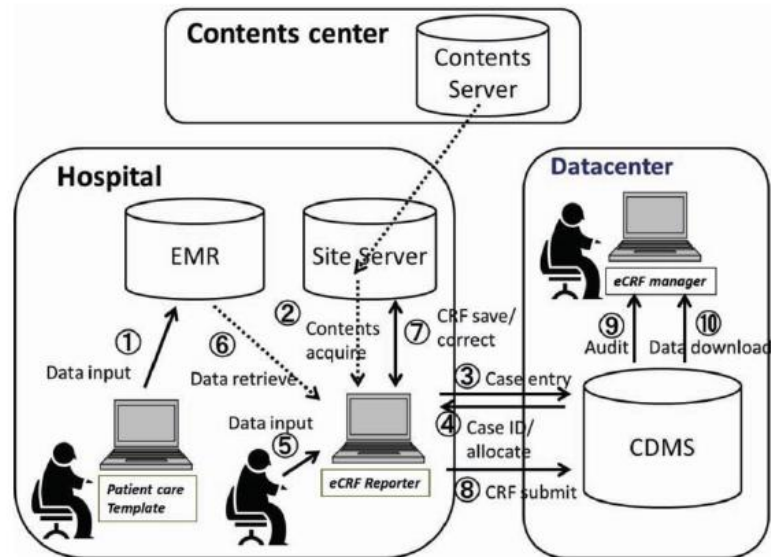


Figure 1. Data flow of the system



## Sections

## Abstract

[↑ Back to top](#)

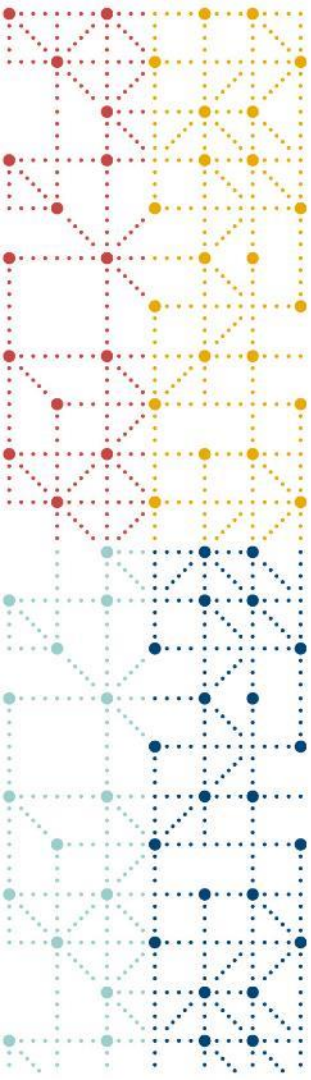
## The Use of CDISC Standards for Real-World Data (RWD): Expert Perspectives from a Qualitative Delphi Survey

Rhonda Facile; Erin Elizabeth Muhlbradt; Mengchun Gong; Qing-Na Li; Vaishali B. Popat; Frank Pétavy; Ronald Cornet; Yaoping Ruan; Daisuke Koide; Toshiki, I. Saito; Sam Hume; Frank Rockhold; Wenjun Bao; I. Sue Dubman; Barbara Jauregui

### ABSTRACT

#### Background:

Real World Data (RWD) and Real World Evidence (RWE) have an increasingly important role in clinical research and health care decision making in many countries. In order to leverage RWD and generate reliable RWE, a framework must be in place to ensure that the data is well-defined and structured in a way that is semantically interoperable and consistent across stakeholders. The adoption of data standards is one of the cornerstones supporting high-quality evidence for clinical medicine and therapeutics development. CDISC data standards are mature, globally recognized and heavily utilized by the pharmaceutical industry for regulatory submission in the US and Japan and are recommended in Europe and China. Against this backdrop, the CDISC RWD Connect Initiative was initiated to better understand the barriers to implementing CDISC standards for RWD and to identify the tools and guidance needed to more easily implement CDISC standards for this purpose. We believe that bridging the gap between RWD and clinical trial generated data will benefit all stakeholders.



**Thank You!**

