

# 无纸化临床研究数据采集平台构建\*

戴维 王雅琴 魏星 廖佳 刘晓琴 杨帆 李强

(电子科技大学医学院附属肿瘤医院 四川省肿瘤医院胸外科 成都 610041)

刘芳

石丘玲

(电子科技大学医学院附属肿瘤医院 四川省肿瘤医院肿瘤中医科 成都 610041)

(1 电子科技大学医学院附属肿瘤医院 四川省肿瘤医院肿瘤预防研究中心 成都 610041

2 重庆医科大学公共卫生与管理学院 重庆 400016)

**[摘要]** 基于临床研究实践提出利用 REDCap 构建无纸化临床研究数据采集平台, 介绍平台概况, 阐述其构建和实施过程、发展趋势, 分析无纸化数据采集方式特点, 为相关研究提供参考。

**[关键词]** REDCap; 无纸化; 电子化平台; 数据采集

**[中图分类号]** R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2021.04.015

**Building of the Paperless Clinical Research Data Collection Platform** DAI Wei, WANG Yaqin, WEI Xing, LIAO Jia, LIU Xiaoqin, YANG Fan, LI Qiang, Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China; LIU Fang, Traditional Chinese Medicine Department of Oncology, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China; SHI Qiuling, 1Center for Cancer Prevention Research, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, 2 School of Public Health and Management, Chongqing Medical University, Chongqing 400016, China

**[Abstract]** Based on clinical research practice, the paper proposes to build a paperless clinical research data collection platform by using REDCap, introduces the general situation of the platform, expounds its building and implementation process and development trend, and analyzes the characteristics of paperless data collection mode, providing references for related study.

**[Keywords]** REDCap; paperless; electronic platform; data collection

**[收稿日期]** 2020-07-29

**[作者简介]** 戴维, 主治医师, 发表论文 13 篇; 通讯作者: 石丘玲, 教授, 博士生导师。

**[基金项目]** 四川省科技计划项目“基于 REDCap 的医学科研数据管理平台的构建及其应用推广研究”(项目编号: 2019YFH0070); 国家自然科学基金面上项目“基于患者报告结局(PRO)构建肺癌围手术期动态康复评价体系的研究”(项目编号: 81872506)。

# 1 引言

目前针对我国临床研究<sup>[1]</sup>，传统数据采集多基于纸质版，数据管理多基于单机版，不能满足现代临床研究需要。开展大型多中心随机对照研究、建立大样本队列研究数据库，实现研究项目实时在线质控等都需要依托强大的电子数据采集管理平台<sup>[2-3]</sup>，然而目前国内大部分医疗机构均未构建此类平台<sup>[4-5]</sup>。REDCap (Research Electronic Data Capture) 是国际通用的电子数据采集管理系统，全世界任何机构都可以在其官方网站申请免费非商业性目的使用<sup>[6-8]</sup>。本文基于临床研究对无纸化数据采集平台构建过程及经验进行介绍，以供参考。

# 2 示例研究和平台介绍

以运用患者报告结局进行肺癌患者术后症状管理的多中心随机对照研究为例<sup>[9]</sup>，介绍该平台建立和实施过程。利用 REDCap 进行数据库构建，涉及患者报告结局 (Patient - Reported Outcome, PRO) 数据和传统非 PRO 数据 (如电子病历记录、临床检查结果、基于随访的术后并发症和生存等) 采集。电子病历报告表共包含 16 个表单，分别为入组、随机化、人口统计学、术前临床特征、手术、麻醉、术后恢复、围手术期并发症、疼痛管理、症状调查表、生活质量调查表、症状及日常功能变化量表、干预的不良事件表、干预的满意度调查表、完成入组信息和随访。其中肺癌症状调查表、生活质量调查表、症状及日常功能变化量表为症状评估表；症状评估表和干预满意度调查表为 PRO 数据采集量表；干预的不良事件表单设计为在相同事件时间点可重复测量。无纸化数据采集过程，见图 1。

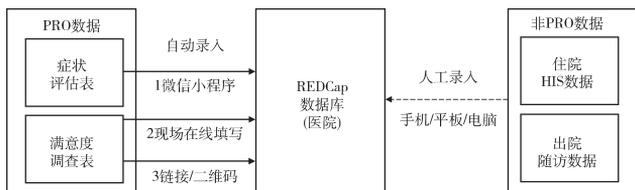


图 1 无纸化数据采集过程

# 3 平台构建和实施过程

## 3.1 利用 REDCap 建立数据库

3.1.1 操作步骤 REDCap 系统安装<sup>[10]</sup>→账号和密码登录→创建新项目<sup>[11]</sup>→进入项目设置→启用调查表<sup>[12]</sup>和纵向数据采集功能<sup>[13]</sup>→点击在线设计器→设计 16 个数据采集表单<sup>[11]</sup>→设定事件时间点<sup>[13-14]</sup>→关联表单和事件时间点<sup>[13]</sup>→启用并设置重复测量表单和事件时间点→启用随机模块<sup>[15]</sup>→设置随机化<sup>[15]</sup>→添加项目用户和设置权限<sup>[16]</sup>→设置数据访问组<sup>[17]</sup>→测试数据库<sup>[18]</sup>→转入正式启用状态。

3.1.2 注意事项 该阶段设计数据采集表单耗时最多且常需反复修改，建议启用病历自动编号和计划模块功能<sup>[13]</sup>。

## 3.2 PRO 数据无纸化采集

该研究中 PRO 数据通过 4 个电子量表采集。其中肺癌症状调查表和生活质量调查表的采集时间点为：术前、术后每天至出院 (≤14 天)、出院后每周 2 次至第 4 周或术后抗癌治疗开始时；症状及日常功能变化量表采集时间点为：出院后每周 2 次至第 4 周或术后抗癌治疗开始时。利用 REDCap 应用程序编程接口 (Application Programming Interface, API) 定制微信小程序 (ePROcell 患者端) 用于 3 个症状评估表的数据采集，患者只需要在计划时间点打开该程序即可进行量表填写。当该程序不可用时研究者可进行补救填写。填写完成后数据自动录入 REDCap 数据库。干预满意度调查表采集时间点为出院当天和出院 4 周后。出院当天患者通过研究者的移动终端直接打开 REDCap 数据库，点击填写满意度调查表。出院后通过企业微信将满意度调查表链接或二维码推送至患者手机进行填写。

## 3.3 非 PRO 数据无纸化采集

电子病历记录、临床检查结果等住院期间产生的非 PRO 数据通过医院信息系统 (Hospital Information System, HIS) 获得。基于随访的术后并发症和生存等出院后非 PRO 数据通过微信或电话随访方式获得。上述数据均通过基于网络的终端设备 (如手

机、平板电脑等) 直接录入 REDCap 数据库。

## 4 讨论

### 4.1 研究概述

本文示例研究为 PRO 症状管理的多中心随机对照研究, 重点涉及 PRO 数据的高频多时间点采集, 其数据采集平台构建和数据采集实施过程相对复杂, 较具代表性。该研究利用 REDCap 实现所有数据无纸化采集, 尤其对于规模较大、复杂且易出现录入错误的 PRO 数据, 均不需要纸质采集和人工录入, 大幅减轻研究者负担, 有利于其将主要精力用于过程质量控制。在数据溯源和安全性方面, 电子化数据采集方式优于传统纸质采集方式。REDCap 具有痕迹追踪功能, 可实时记录每次操作的用户、时间及内容等信息, 方便研究者和监管部门进行核查。REDCap 服务器置于研究机构, 可实现平台所有数据定期自动备份及单个项目随时备份。

### 4.2 PRO 数据无纸化采集方式

基于 REDCap 实现 PRO 数据的无纸化采集主要有 3 种方式, 一是利用 API 开发患者端应用程序, 对于研究者较省时省力, 但前期需要投入较多时间和经费。二是打开 REDCap 数据库中的量表现场直接填写, 适用于研究者直接面对患者的场景, 如住院期间。三是将链接或二维码转发到患者微信或邮箱, 适用于远程随访。在国外患者多有使用邮箱的习惯<sup>[7, 12, 19]</sup>, 所以远程随访时只需利用 REDCap 自动发送邮件功能将 PRO 量表定时发送到患者邮箱即可。国内患者大多使用智能手机和微信, 所以通过微信推送 PRO 量表更方便可行。在多时间点的数据采集中, 一个时间点对应的 PRO 量表是唯一的, 即每个时间点的 PRO 量表对应的链接或二维码也是唯一的。远程随访时需要人工定时将量表链接或二维码发送到患者微信, 如果量表较多工作量也较大。

### 4.3 非 PRO 数据无纸化采集方式

基于 REDCap 实现非 PRO 数据无纸化采集主要有两种方式。一是利用 REDCap API 和医院信息系统对接, 实现临床数据自动化采集<sup>[20]</sup>, 是最优方案; 二是采用传统人工录入方式。即使在 MD 安德

森癌症中心等国外较先进医院, 出于安全等因素考虑, 也仅对小部分电子病历数据实行自动提取。国内虽然技术上可实现 REDCap 和 HIS 对接, 但受资金、安全、病历质量等其他因素限制较难实现, 因此第 2 种方式较为现实和可行。

### 4.4 临床研究数据采集平台发展过程和趋势

临床研究数据采集平台发展过程和趋势可分为 5 个阶段。第 1 阶段特点为: 纸质版采集数据, 人工录入 PRO 和非 PRO 数据。第 2 阶段特点为: 无纸化采集数据, 人工录入 PRO 和非 PRO 数据。第 3 阶段特点为: 无纸化采集数据, 自动录入 PRO 数据, 人工录入非 PRO 数据。第 4 阶段特点为: 无纸化采集数据, 自动录入 PRO 数据, 自动录入和人工纠错非 PRO 数据。第 5 阶段特点为: 无纸化采集数据, 自动录入 PRO 数据, 自动录入和机器纠错非 PRO 数据。该示例研究目前处于第 3 阶段, 后续将逐渐过渡到第 4 和第 5 阶段, 以最大化地减轻研究者数据录入负担。

## 5 结语

利用 REDCap 建立无纸化临床数据采集平台开展临床研究有助于减少数据录入工作量, 方便可行。REDCap 可用性较好、功能较强, 目前尚处于发展阶段, 有待进一步深入研究与应用。

## 参考文献

- 1 曹焯, 王欣, 曹玉, 等. 我国研究者发起的临床研究管理现状调查与分析 [J]. 中国新药与临床杂志, 2018, 37 (7): 395 - 400.
- 2 曹焯. 哈佛医学院 DF/HCC 临床研究管理介绍——临床研究的“中心化”管理 [J]. 中国新药杂志, 2015, 24 (18): 2128 - 2133.
- 3 刘晋. 电子数据采集系统在临床科研管理中的实践与思考 [J]. 中华医院管理杂志, 2019, 35 (6): 503 - 506.
- 4 谢高强, 李英山, 姚晨. 电子数据采集对我国临床研究的机遇和挑战 [J]. 中国新药杂志, 2013, 22 (6): 620 - 623.
- 5 姚宏伟, 张忠涛. 基于规范化的临床实践和数据管理做好中国结直肠外科的临床研究 [J]. 中华消化外科杂志, 2019, 18 (1): 31 - 34.
- 6 Harris PA, Taylor R, Thielke R, et al. Research Electronic Data Capture (REDCap) — a metadata - driven methodolo-

- gy and workflow process for providing translational research informatics support [J]. *J Biomed Inform*, 2009, 42 (2): 377-381.
- 7 Cantor Z, Plint AC, Kamil Y, et al. REDCap: a pediatric ED experience [J]. *Am J Emerg Med*, 2016, 34 (10): 2048-2049.
  - 8 Harvey LA. REDCap: web-based software for all types of data storage and collection [J]. *Spinal Cord*, 2018, 56 (7): 625.
  - 9 Dai W, Zhang Y, Feng W, et al. Using Patient-reported Outcomes to Manage Postoperative Symptoms in Patients with Lung Cancer: protocol for a multicentre, randomised controlled trial [J]. *BMJ Open*, 2019, 9 (8): e030041.
  - 10 耿辉, 贺海蓉, 曾宪涛, 等. 多中心临床数据采集系统 REDCap 系统应用及架设 [J]. *中国循证心血管医学杂志*, 2017, 9 (9): 1025-1028.
  - 11 贺海蓉, 郑婕, 耿辉, 等. 如何利用 REDCap 数据采集系统实现病例报告表的设计 [J]. *中国循证心血管医学杂志*, 2018, 10 (2): 133-137.
  - 12 Tomko RL, Gray KM, Oppenheimer SR, et al. Using REDCap for Ambulatory Assessment: implementation in a clinical trial for smoking cessation to augment in-person data collection [J]. *Am J Drug Alcohol Abuse*, 2019, 45 (1): 26-41.
  - 13 李蒙, 田国祥, 贺海蓉, 等. 如何利用 REDCap 实现重复测量数据的录入及管理 [J]. *中国循证心血管医学杂志*, 2018, 10 (6): 647-650.
  - 14 侯倩, 江真, 刘卿, 等. REDCap 数据库在嗓音外科临床研究中的初步应用 [J]. *中国耳鼻咽喉颅底外科杂志*, 2017, 23 (4): 338-341.
  - 15 贺海蓉, 田国祥, 任晓东, 等. 如何利用 REDCap 实现临床试验的随机化操作 [J]. *中国循证心血管医学杂志*, 2018, 10 (9): 1039-1042.
  - 16 耿辉, 贺海蓉, 郑婕, 等. 多中心临床数据采集系统 REDCap 权限设置与管理 [J]. *中国循证心血管医学杂志*, 2018, 10 (1): 11-13.
  - 17 高凡, 田国祥, 贺海蓉, 等. REDCap 实现多中心研究数据管理的方法 [J]. *中国循证心血管医学杂志*, 2018, 10 (4): 392-395.
  - 18 耿辉, 田国祥, 贺海蓉, 等. 如何安全实现 REDCap 系统的数据录入、存储 [J]. *中国循证心血管医学杂志*, 2018, 10 (3): 262-264.
  - 19 Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-reported Outcomes during Routine Cancer Treatment: a randomized controlled trial [J]. *J Clin Oncol*, 2016, 34 (6): 557-565.
  - 20 Campion TR, Sholle ET, Davila MA. Generalizable Middleware to Support Use of REDCap Dynamic Data Pull for Integrating Clinical and Research Data [J]. *AMIA Jt Summits Transl Sci Proc*, 2017 (2017): 76-81.

(上接第 69 页)

## 参考文献

- 1 唐维红, 唐胜宏, 廖灿亮. 跨入 5G 时代的中国移动互联网—《中国移动互联网发展报告 (2020)》发布 [J]. *中国报业*, 2020 (17): 32-35.
- 2 张鹏, 赵动员, 梅蕾. 移动社交网络信息传播研究述评与展望 [J]. *情报科学*, 2020, 38 (2): 170-176.
- 3 赵加奎, 陈虹仪, 庄颖, 等. 三种呈现方式在新媒体健康传播中的效果评价 [J]. *健康教育与健康促进*, 2020 (4): 411-413, 421.
- 4 张明, 刘嘉浩, 陈晨. 疫情防控期间新媒体传播的功能及特征研究 [J]. *新闻研究导刊*, 2020, 11 (20): 106-107.
- 5 潘维华, 林静, 孙秀芸. 从新冠肺炎疫情看当下网络舆情应对新变革 [J]. *新闻研究导刊*, 2020, 11 (19): 140-141.
- 6 邱雅, 王华彬. 网络舆情分析中智能爬虫的设计 [J]. *电脑知识与技术*, 2011, 7 (33): 8301-8302.
- 7 袁畅, 郝志峰, 蔡瑞初, 等. 一种分布式的舆情分析系统架构 [J]. *电信科学*, 2013 (7): 66-71.
- 8 刘斌. 基于 Twitter 大数据处理的境外舆情分析系统设计与实现 [J]. *电脑知识与技术*, 2020, 16 (27): 30-33, 42.
- 9 习海旭, 蒋红芬, 程志凡, 等. 特定事件下网络舆情的情感分析与可视化方法 [J]. *情报理论与实践*, 2020, 43 (9): 132-136, 143.
- 10 肖倩, 谢海涛, 刘平平. 一种融合 LDA 与 CNN 的社交媒体中热点舆情识别方法 [J]. *情报科学*, 2019, 37 (11): 27-33.
- 11 张连峰, 周红磊, 王丹, 等. 基于超网络理论的微博舆情关键节点挖掘 [J]. *情报学报*, 2019, 38 (12): 1286-1296.
- 12 彭程, 祁凯, 黎冰雪. 基于 SIR-EGM 模型的复杂网络舆情传播与预警机制研究 [J]. *情报科学*, 2020, 38 (3): 145-153.
- 13 周琦萍, 杨芳. 基于 SIS 模型的网络舆情无监督预警机制研究 [J]. *情报科学*, 2019, 37 (8): 51-55.
- 14 湛志华. 基于大数据的网络舆情分析系统 [J]. *现代电子技术*, 2017, 40 (24): 15-17.