

数据解析在肿瘤放射治疗中的应用

段永璇 邹晓艳 段秀梅 李青

【摘要】 笔者在该文中介绍了肿瘤放射治疗中医学数字影像和传输数据文件的编码规则以及数据机构,综合放疗数据的产生过程和实际分布,阐述了放疗临床数据库标准化构建的相关问题,关注数据之间的逻辑关系以及数据构建过程中的细节,结合数据解析在放疗临床中的实际应用,以及软件思维进行数据解析的方法,进一步探讨了放疗数据深度整合中的相关问题和应用意义。

【关键词】 肿瘤治疗方案;放射疗法,计算机辅助;DICOM RT;数据解析;医院信息化;数据整合

Data analysis in the field of tumor radiation therapy Duan Yongxuan*, Zou Xiaoyan, Duan Xiumei, Li Qing. *Department of Information Network, Shandong Institute of Medicine and Health Information, Shandong Academy of Medical Sciences, Jinan 250062, China

Corresponding author: Zou Xiaoyan, Email: 34941269@qq.com

【Abstract】 Described in this paper are the data structure and encoding rules of digital imaging and communication in medicine radiotherapy in tumor radiation therapy. The process and the actual distribution of the radiation data integrated. Expanded the relevant issues of data base standardization construction, focus on the logical relationship of data analysis and the details of the data construction process, combined the data processing in clinic with the method of software technology in data analysis, further explore the issues related to data integration and application in radiation therapy.

【Key words】 Antineoplastic protocols; Radiotherapy, computer-assisted; DICOM RT; Data analysis; Hospital information; Data integration

放射治疗(简称:化疗)是当今医学界治疗恶性肿瘤三大主要手段之一,现代医疗器械硬件的快速发展,以及计算机数字化技术在医学领域的广泛应用,使得放疗技术异军突起。尤其是在计算机出现以后,现代信息科学技术的飞速发展,为放疗更快地朝着精确定位、精确计划、精确治疗的方向不断发展提供了有力的数字化技术保障。随着各地医院信息化建设的不断加强,传统医疗服务模式也更趋于数字化,计算机软硬件技术在放疗领域的不断深入,各种治疗方法和数据处理算法不断创新,为治疗计划系统、放疗计划的验证、实施、监督和记录系统的不断优化提供了有效的技术实施保障。数据是信息化建设的基础载体,信息化的管理本质就

是数据的管理,了解肿瘤放疗数据的内部结构及其元素结构,以及数据在临床中的产生过程,掌握放疗的医学数字影像和传输(digital imaging and communication in medicine radiotherapy, DICOM RT)标准中的数据结构以及数据解析方法,已成为临床数据集中化管理、提升临床科研水平的一个重要命题,也是提升放疗的现代化程度,加快放疗领域的信息化建设进程的必要条件。笔者就从放疗数据的基本构成、放疗数据库标准化构建、数据解析思路等方面进行综述,探讨数据解析在放疗领域的数字化建设中发挥的重要作用。

1 放疗数据的机构

国际上通用医学领域的数字传输标准是 DICOM 标准,在放疗领域的数字存储传输标准是 DICOM RT 标准。DICOM 文件是由大量的数据元素(data element)形成的数据集合,数据元素是组成数据集(data set)的基本单位。一个完整的数据元素包括

DOI: 10.3760/cma.j.issn.1673-4114.2015.06.015

作者单位: 250062 济南,山东省医学科学院山东省医药卫生科技信息研究所信息网络部(段永璇); 250004,济南军区联勤部门门诊部(邹晓艳); 250002 济南,山东省职业卫生与职业病防治研究院肾内科(段秀梅); 250001,济南市第二人民医院信息科(李青)

通信作者: 邹晓艳(Email:34941269@qq.com)

数据标签(data element tag)、值表达(value representation)、值长度(value length)和数据区(value field)^[1]。数据标签是数据元素的唯一标识,包括群编号(group number)和元素编号(element number)两组16位无符号整数。值表达描述数据类型,分显式和隐式,显式的情况下为2或4字节字符串。数据区描述数据元素的具体值,长度为偶数字节。值长度描述数据区的字节长度,为16或32位无符号整数^[2]。数据集中的数据元素严格按照数据标签的升序存放,从而包含了DICOM RT的各种字段信息(包括患者信息、诊断信息以及图像数据信息)。DICOM RT数据实体如图1所示。

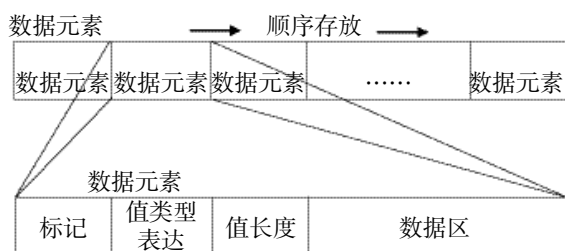


图1 放射治疗的医学数字影像和传输数据实体示意图

Fig.1 Data set and data element structures of digital imaging and communication in medicine radiotherapy

在实际的放疗临床应用中,通过DICOM RT标准,医院可以实现不同厂家医学图像设备、医学图像的放疗设备、图像工作站以及相应的放疗计划系统、质量验证系统等之间的相互通信。此外,该标准为医院实现放疗信息的数字化管理提供了方便,也为多系统的数据融合提供了重要的网络标准和协议,是完成不同来源的医学数字化影像系统间互联和影像互操作性最直接、最有效的手段。DICOM RT对象在实际的临床治疗过程中,数据的具体产

生过程如表1所示。

通过以上的数据实施步骤,就形成了完整的包括RT结构集(RT structure set)、RT计划(RT plan)、RT剂量(RT dose)和RT图像(RT image)、体外照射治疗记录(beam treatment record)、近距离治疗记录(brachy therapy treatment record)和放射治疗综合记录(treatment summary record)的记录治疗传输数据^[3]。DICOM RT数据的结构化层次如图2所示。

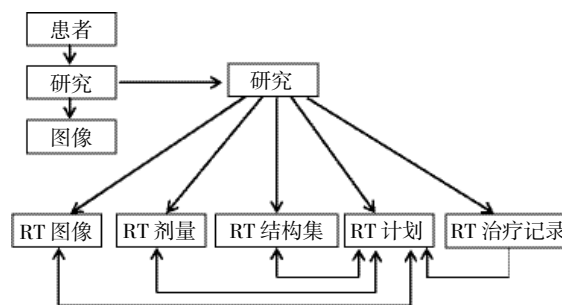


图2 放射治疗的医学数字影像和传输数据结构层次图中,RT:放射治疗。

Fig.2 Data structure of digital imaging and communication in medicine radiotherapy

DICOM RT标准沿用了DICOM的数据编码的方式,包含了DICOM定义的26种内部数据类型中的大部分,引入数据项来表达嵌套结构,1个数据集包含数个数据元素,1个具有嵌套结构的数据元素可以包含1个或多个数据项,而1个数据项又包含1个独立的子数据集,如此嵌套进行^[4]。DICOM RT数据文件的嵌套相当频繁,如RT结构集数据文件可出现4层嵌套结构^[5],这就使得DICOM RT中用来记录数据集中内容类型相同的信息对象的数据结构更复杂,如轮廓、射野等数据。

表1 放射治疗的医学数字影像和传输数据生成过程

Table 1 Data generation process of digital imaging and communication in medicine radiotherapy

序号	设备处理	数据生成
1	医学影像设备对患者进行CT扫描。	由CT、MRI等设备采集得到DICOM格式的图像序列。
2	模拟工作站请求CT使用DICOM接口发送CT图像,并对DICOM图像进行处理,进行三维重建。	生成RT结构集对象(包括病灶和受保护区域等)。
3	治疗计划系统生成初步的RT计划对象。	产生RT计划数据(包括射束几何尺寸等)。
4	利用重建技术产生RT图像对象。	产生RT图像数据。
5	治疗计划系统读取上述对象,计算剂量数据。	产生RT剂量对象,生成新的RT计划。
6	验证系统取得完整的RT计划,对直线加速器等治疗设备初始化。	产生验证图像,作为一个新的RT图像对象,可与RT图像数据进行比较,验证数据的有效性。
7	将有效数据送至治疗设备进行治疗,治疗设备和验证记录系统定期对治疗参数作记录。	产生RT治疗记录对象。

注:表中,DICOM:医学数字影像和传输;RT:放射治疗。

2 标准化数据库构建

通过以上对 DICOM RT 数据的逻辑结构和元素分析可以看出,由于放疗数据较其他数据更为复杂,数据结构层次及附加信息量较其他数据更为庞大,因此放疗数据的整合在放疗领域一直是一个较难解决的问题,一方面临床数据涉及的数据内容复杂,零散分布于各个业务系统,异构数据多,数据量级大;另一方面受临床医师的专业及实际情况的限制,很难单方面去完成此类专业数据的深度整合。但随着大数据时代的来临,放疗领域和医院内部的数据整合又势在必行,因此必须重新正视放疗数据的整合与数据规范化建设的问题。DICOM RT 文件包含的数据类型复杂,既包括文本字段、图像等常规数据,又包含计划、治疗、剂量体积直方图、放射剂量等非常规数据^[6],对数据的传递性及关联性要求较高,因此需要对数据集之间的数据关系做明确而详细的规划,在理解 DICOM RT 数据集的数据架构的基础上,分析数据之间的逻辑关系,包括数据纵向的传递关系,以及横向的关联性,设计必要的数据归类方法,考虑数据之间的相互关系,将数据进行划分和归类,为数据库的构建提供参考。

DICOM RT 是对 DICOM 的补充扩展, DICOM 标准中的数据是针对信息对象定义(information object definitions, IOD)而来的,因此 DICOM RT 标准与 DICOM 标准一样采用实体-联系方法的基本信息模型对实体进行抽象描述,使用 IOD 的形式来建立放疗领域的数据库模型,并用服务类的方式实现对放疗信息对象的操作^[7],从而完成应用层功能。不管是信息的存储还是不同设备间的信息交换都是以 IOD 实例来进行的,所以数据库的设计应尽量保持 IOD 的完整性,体现 IOD 之间的关系,可按照患者、研究、系列和图像 4 个层次来进行检索和管理,保持数据的完整性。基于以上思想,数据库的框架结构应尽量与 DICOM RT 标准信息模型保持一致。此外,由于 DICOM RT 的文件信息是按文件存储和通讯的,相对于 DICOM 信息对象(CT、MR、PET 等),DICOM RT 信息对象的数据更庞大,除了普通的图像信息,还包括剂量信号(剂量体积直方图、等剂量线等)、射野信息(准直器、挡块)等。而且这些数据常常又对应着大量浮点数据

文件,不利于信息的检索和存档,因此需要通过数据的标准化建设和解析,构建相应的数据库。数据库设计过程中,应遵循逻辑结构、信息对象关系、元素属性、属性值的表达方式等,尽量与 DICOM RT 标准保持对应。通过数据解析手段将临床放疗中产生的数据进行分析,把重要的图像信息及诊断信息从 DICOM RT 文件中提取出来,并按照一定的逻辑结构相对地存储在标准化数据库中。

3 数据解析方法

通过以上方法,可以搭建起一个标准的基于 DICOM RT 数据架构的放疗数据库,如何将 DICOM RT 数据更好地提取和存入,就需要通过数据解析的手段进行实现。DICOM 的数据集是描述现实世界医学信息对象实例的手段,是 DICOM 文件的基本内容组成,它采用树状结构,并且符合二进制编码规则^[8],这在临床应用中的 DICOM RT 数据编译转换和解析的实现提供了数字化实现途径。因此,可以按照 DICOM RT 的数据传输标准^[9],对各元素的数据地址顺序读取,并对此类数据进行数据解析,通过定位数据文件中包含的数据元素,读取数据元素中的数据标示符,从数据值域中获得对应的关键数据,并将其映射存储到数据库的相应字段中,从而获得文件的信息内容。以此思路为依据,可以通过编写程序实现数据解析,将 Tag 值作为一条主线,相当于一个指针。程序在解析、调用数据时,首先找到需要信息的 Tag 值,读取此 Tag 后的 Length 值,再根据其 Length 值从后继信息中找出所需内容并将其显示。在实际数据的解析过程中,根据 DICOM 文件头和数据集合,通过读取文件头中 80H 位的“DICM”标识来判断是否为 DICOM 数据^[9],这样就能在一堆混合文件中将其数据逐个解析出来,然后根据 DICOM RT 的特殊标志位,萃取数据集中患者姓名、图像组名(series number)、图像序列号(instance number)等基本信息^[10]。通过以上方式的层层推进,进一步完成对放射剂量、剂量体积直方图等信息的解析,最终通过数据整合,实现 DICOM RT 的数据解析。程序的解析模块如图 3 所示。

4 总结

放疗已逐步迈入图像引导的精确放疗时代,数

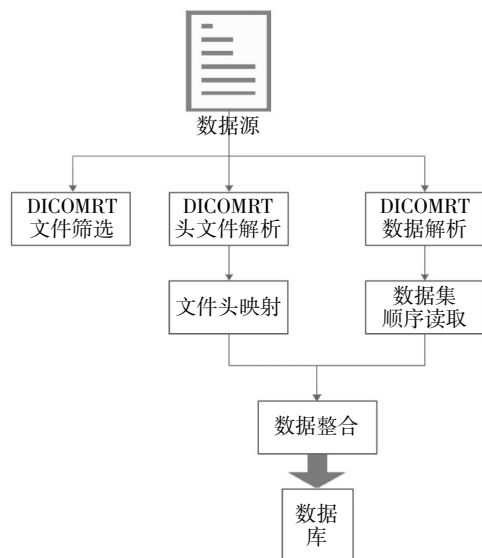


图3 放射治疗的医学数字影像和传输数据解析程序模块 图中, DICOM RT: 放射治疗的医学数字影像和传输

Fig.3 Data analysis module of digital imaging and communication in medicine radiotherapy

据作为基础性保障显得尤为重要, 通过加深对放疗数据的解析和理解, 结合有效的技术手段, 利于保证数据在放疗信息系统中的有效传输, 更好地使数据在临床治疗中发挥积极作用, 并利用其演绎出更多的优化算法, 保证肿瘤放疗的精准性。此外, 不同系统、治疗设备间的信息数据整合, 为医师查阅患者治疗信息以及对放疗计划进行质量保证、质量控制提供了数字化基础, 为今后的放疗数据分析及处理方法的深入研究, 提供了必要的实践基础。以数字化资源为基础, 遵循循序渐进、规范标准和实用为主等基本原则, 抓好医院数据管理和信息化的建设工作, 从技术、管理、应用等各个层面真正把

医院的信息孤岛连通起来, 把海量的数据整合利用起来, 从而有利于加快医院信息化进程, 有利于加强放疗的质量保证, 有利于放疗数据的集中管理, 有利于全面提升临床工作。

参 考 文 献

- [1] 於文雪, 张惠, 罗立民. DICOM 标准在放射治疗中的应用[J]. 中国医疗器械杂志, 2002, 26(5): 352-355.
- [2] Yu C, Yao Z. XML-based DICOM data format [J]. J Digit Imaging, 2010, 23(2): 192-202.
- [3] 刘芳, 曹瑞芬, 裴曦, 等. DICOM-RT 解析及在精确放射治疗计划系统中的应用[J]. 中国医学物理学杂志, 2012, 29(2): 3263-3266.
- [4] 吴磊, 於文雪, 罗立民. DICOMRT 数据解析库的开发与应用[J]. 医疗装备, 2006, 19(4): 11-13.
- [5] 蒲立新, 曲建明. DICOM 放射治疗中结构集文件的创建[J]. 中国数字医学, 2011, 6(2): 102-105.
- [6] 夏德国, 雷力, 周凌宏. 基于 DICOM 结构化报告模型和可扩展标记语言的放疗计划电子病历[J]. 中国组织工程研究, 2012, 17(17): 3188-3192.
- [7] 雷力, 周凌宏, 夏德国. DICOM-RT 标准解决放射治疗信息系统标准化通讯模型的构建[J]. 中国组织工程研究, 2012, 17(17): 3178-3182.
- [8] National Electrical Manufacturers Association. ISO12052 -2015 Digital imaging and communications in medicine (DICOM) [S]. Rosslyn VA: National Electrical Manufacturers Association, 2015.
- [9] National Electrical Manufacturers Association. ISO12052 -2009 DICOM-information object definitions[S]. Rosslyn VA: National Electrical Manufacturers Association, 2009.
- [10] Gorthi S, Bach Cuadra M, Thiran J. Exporting contours to DICOM-RT structure set[J/OL]. Insight J, 2009[2015-08-07]. <http://hdl.handle.net/1926/1521>.

(收稿日期: 2015-08-11)