

流式细胞术白血病免疫分型资料组织及数据管理的优化*

杜庆华, 李庆山, 许艳丽 (广州市第一人民医院 血液内科, 广州 510180)

摘要:目的 旨在探讨优化流式细胞术白血病免疫学分型及微小残留病检测的资料组织及数据管理的方法,使资料查询、回溯及归纳更为高效。**方法** 通过建立患者资料表及样本资料表,对2010年1月~2014年12月间4 530个患者的6 238份标本的基本资料及检测结果进行管理,并使用数据资料表对这些标本检测得到的原始数据文件35 563个进行管理。采用ACCESS 2010及该实验室自行开发的流式分析软件CFCS,根据流式细胞术数据标准(flow cytometry standard, FCS) 3.0版及3.1版读取流式原始数据的相关关键字,并将该信息于数据资料表中保存。使用结构化查询语言(structured query language, SQL)通过编号或姓名检索患者,通过诊断对某一病种的病例进行检索汇总,以及通过不同逻辑式对数据库进行检索操作。通过对各个表进行连接并导出排序,以人工方式检查SQL检索的准确性。**结果** 通过人工比对,每次检索均能正确无误且高效地检索出正确的结果。**结论** 使用数据库对白血病免疫分型及微小残留病检测的资料及数据进行组织管理,使用SQL进行查询操作,能方便高效地实现资料数据的回溯及归纳。

关键词:流式细胞术;免疫分型;数据管理;数据库

中图分类号:R392-33 文献标志码:A 文章编号:1671-7414(2015)06-163-02

doi:10.3969/j.issn.1671-7414.2015.06.050

Optimization for Information Organizing and Data Management in Flow Cytometry for Leukemia-Immunophenotyping

DU Qing-hua, LI Qing-shan, XU Yan-li

(Department of Hematology, the First People's Hospital of Guangzhou, Guangzhou 510180, China)

Abstract: **Objective** To investigate how to organize information and manage data in leukemia-immunophenotyping and minimal residual disease monitoring for leukemia immunophenotyping, and make data inquiry more efficient. **Methods** Managed each result of patients by creating information table of patients and information table of samples, using data file table to manage data files of flow cytometry, and read the keywords from flow cytometry data file according to Flow Cytometry Standard (FCS) 3.0 and 3.1, and saved them to the data file table. Structured Query Language (SQL) could be used to perform inquiry about a patients by patient number and patient name, cases with a certain diagnosis, and different logical algorithms in this database. Tables were linked and sorted, then perform manual search to verify the results of searching by SQL. **Results** In comparison with manual search, SQL could query the result correctly. **Conclusion** Database can be used to manage and organize information of patient and data of Leukemia-Immunophenotyping, and SQL can be used to perform inquiry efficiently.

Keywords: flow cytometry standard; immunophenotyping; data management; database

白血病免疫分型(leukemia-immunophenotyping)作为白血病MICM实验室诊断的重要部分,有着不可取代的地位,其主要使用多色流式细胞术进行检测。由于白血病免疫分型及微小残留病检测(minimal residual disease monitoring)流式数据的特殊性 & 复杂性,而流式细胞术数据分析软件为了保证其在其它临床检测及科研使用中的灵活性,并没有对白血病免疫分型及微小残留病检测的数据管理进行优化,也没有将数据与患者资料进行绑定,因此造成对数据的回溯、管理及归纳等功能严重不足。而临床工作或行微小残留病监测时往往需要对患者既往免疫分型的结果进行回溯,如果遇到疑问时甚至需要重新设门分析,资料管理手段的不足造成诸多不便。此外,如果需要对某个疾

病特定抗原表达情况进行归纳时,对流式细胞术原始数据的查找是极其费时的。虽然现有的实验室信息系统(laboratory information system, LIS)能实现部分患者资料的管理工作,但是LIS系统仅仅能管理文字和图形报告,对于原始数据的管理仍无法实现。本文将探讨使用本实验室自行开发的流式分析软件CFCS,结合数据库对患者资料及原始数据进行管理,以探讨其对资料数据回溯及归纳的准确性和效率。

1 材料与amp;方法

1.1 仪器及软件 流式细胞仪为FACS CANTO(美国BD公司),软件使用ACCESS 2010及本实验室自行开发的流式分析软件CFCS(软件著作权登记号:2010SR064983)。

* 基金项目:广东省自然科学基金,项目编号:S2013010016726。

作者简介:杜庆华(1977—),男,硕士研究生,副主任技师,研究方向:血液病分子诊断, E-mail: haven__t@sina.cn。

通讯作者:李庆山, Tel: 020-81048386, E-mail: qingshanli@hotmail.com。

1.2 对象 2010年1月~2014年12月间,本实验室白血病患者免疫分型及微小残留病检测的资料及数据,共计患者4530人,合计标本6238份,原始数据文件35563个。

1.3 方法 使用 Access 2010 建立 Patient table, Sample table 及 Data table 三个二维数据表。数据表 Patient table 内容包括患者编号、姓名、性别及出生年月。数据表 Sample table 内容包括样本编号、患者编号、年龄、病区、床号、诊断、标本类型、标本状态、检测项目、接收日期、检测日期、保存路径。Data table 用来保存流式数据的相关属性,见表1。Data table 中的记录通过“sid”与“Sample table”中的相应记录连接。“Sample table”中的记录通过“pid”与“Patient table”中的相应记录连接。“Data table”中的“file name”是该记录对应数据文件的文件名。根据该记录的“sid”从“Sample table”获取该标本所有数据的文件夹路径,文件夹路径加上文件名即可定位到文件的位置。

表1 数据表 Data table 的内容

数据表	字段	注释
Data table	did	数据编号
	sid	样本编号
	filename	数据文件名
	btim	数据记录开始时间
	etim	数据记录结束时间
	Cytometer	流式细胞仪名称
	operator	操作者
	FnS	第n通道的抗体
	FnN	第n通道的荧光素

CFCS 能实现当一个样本的数据分析后自动导入到数据库中,在样本信息表的 PATH 字段自动填入分析文档的路径,并自动创建文件夹,把该样本所分析的原始数据放置其中。根据国际流式细胞术发展协会(international society for advancement of cytometry, ISAC)所制定的流式数据标准(flow cytometry standard, FCS) 3.0 及 3.1 版^[1,2]的文件结构描述,使用 CFCS 读取每个原始数据中数据库参数信息表所需要的关键字(Keywords)并填入表中。

在 CFCS 中提供的图形化检索工具,对指定患者资料、指定诊断、检测指定抗原原始数据的检索以及不同逻辑式检索进行检索。

以下 SQL 语句用于检索接收日期在 2011 年 01 月 01 日~2014 年 12 月 31 日期间的所有诊断为 CML 的记录: select PatientTable. ID, 姓名, 性别, 年龄, 病区, 床号, 诊断, 标本, 接收日期, 检测日期, 标本状态, 备注, 路径 from PatientTable, SampleTable where Sample. PID = PatientTable. ID AND 诊断 = 'CML' AND 接收日期 < # 2014-12-31 # AND 接收日期 > # 2011-01-01 #。以下 SQL 语句用于检索接收日期在 2011-01-01~2014-12-31 期间的所有诊断为 CLL 且使用第 5 通道检测 CD19 的数据: select PatientTable. ID, 姓名, 性别, 年龄, 病区, 床号, 诊断, 标本, 接收日期, 检测日期, 标本状态, 备注, 路径, DataTable. filename from PatientTable, SampleTable, DataTable, where SampleTable. PID = PatientTable. ID AND 诊断 = 'CLL' AND F5S = 'CD19' AND 接收日期 < # 2014-12-31 # AND 接收

日期 > # 2011-01-01 #。

2 结果 对指定患者资料、指定诊断、检测指定抗原原始数据的检索以及不同逻辑式检索的结果均和手工检索结果完全一致。

3 讨论 虽然现在的流式细胞仪可检测的颜色已经能达到 10 色以上,但是进行白血病患者免疫分型检测时,仅使用一管对所有所需抗原进行检测仍未能实现。故对标本进行检测时,必须分开多管进行检测,以致每个标本需对应多个原始数据,每个数据所检测的抗原均不相同。此外,由于白血病患者免疫分型的复杂性,每个标本所需检测的抗原不尽相同,这样也增加了原始数据管理与查找的难度。

为了满足用户对分析的需求,ISAC 自 1984 年开始规范并完善了流式数据文件的标准^[1~4]。各流式生产厂商均可按该标准生成标准的 FCS 文件,以方便用户使用第三方软件对其数据进行分析,以弥补仪器自带软件分析功能的不足。FCS 文件除了记录上机所有细胞的数据外,还以关键字的形式记录了仪器的厂家、型号、状态、时间、每个通道的荧光素以及抗体。在关键字中,以“\$”开始的关键字为标准 FCS 关键字,例如“\$CYT”为仪器的名称,“\$TOT”为该文件所记录的总细胞数。“\$PnN”记录的是该通道所检测的荧光素,而“\$PnS”则是记录该通道所检测的荧光抗体^[1]。由于仪器除了检测荧光参数外还要检测物理参数如 FSC, SSC 及 Time 等,当物理参数排在前面时,“\$PnN”及“\$PnS”中的 n 并不代表荧光通道的编号。当需要筛选出检测某个特定抗原的数据时,一般需要逐个打开原始数据文件,识别所有关键字后再进行查找判断,这种方法效率极低,特别是文件数量大的时候。若不识别关键字直接查找文件内的字符串时,判断效率会较高,但容易检出无关数据。例如查询包含 CD3 的数据,CD33, CD34 等子串包含 CD3 的均会命中。而通过数据库经过索引后使用二分查找法,能准确且极为高效地检索到所需要的数据。当数据由两台流式细胞仪生成,其荧光通道的参数可能差异较大,此时可把“Data table”中的荧光参数用一个独立的数据表进行储存,这样可增加数据库的灵活性。

总之,使用数据库对白血病患者免疫分型患者信息及原始数据进行统一管理,可方便高效地对数据进行回溯及归纳总结。

参考文献:

- [1] Seamer LC, Bagwell CB, Barden L, et al. Proposed new data file standard for flow cytometry, version FCS 3.0[J]. Cytometry, 1997, 28(2): 118-122.
- [2] Spidlen J, Moore W, Parks D, et al. Data file standard for flow cytometry, version FCS 3.1[J]. Cytometry Part A; the Journal of the International Society for Analytical Cytology, 2010, 77(1): 97-100.
- [3] Murphy RF, Chused TM. A proposal for a flow cytometric data file standard[J]. Cytometry, 1984, 5(5): 553-555.
- [4] Seamer L. Data file standard for flow cytometry, FCS 3.0[J]. Curr Protoc Cytom, 2001(chapter 10): 326-327.

收稿日期: 2015-03-17

修回日期: 2015-06-15