

基于实验室信息管理系统和帆软报表的 生物制品检测数据趋势分析研究

王坤*, 刘健, 徐成瑶

(吉林省药品检验研究院, 长春 130033)

摘要: 生物制品在生产和检验过程中会产生大量相关数据,但数据处理过程存在着低数字化管理、工作不便等问题。本文根据趋势分析在生物制品质量监控中的特点,建设了趋势分析系统。该系统基于实验室信息管理系统和帆软报表工具研发,融入仪器数据采集和检验数据清理相关技术,利用休哈特控制图,完成对检测数据的采集、清理、汇总,并提供统计、分析和可视化监控、报警等功能,提高生物制品趋势分析的质量和效率,追踪其在生产和检验中的稳定性和变异性,为生物制品的科学管理提供参考。

关键词: 实验室信息管理系统; 帆软报表; 生物制品; 趋势分析

Research on the trend analysis of biological products based on laboratory information management system and FineReport

WANG Kun*, LIU Jian, XU Cheng-Yao

(Jilin Institute for Drug Control, Changchun 130033, China)

ABSTRACT: In the process of production and inspection of biological products, a large amount of relevant data will be generated, but there are problems such as low digital management and inconvenient work in the data processing process. This paper constructed a trend analysis system based on the characteristics of trend analysis in the quality control of biological products. Based on the research and development of laboratory information management system (LIMS) and FineReport (FR), the system integrated the relevant technologies of instrument data acquisition and inspection data cleaning, used the shewhart control chart to complete the collection, cleaning and summary of the detection data, and provided statistics, analysis and visual monitoring, alarm and other functions to improve the quality and efficiency of the trend analysis of biological products. The stability and variability in production and testing were tracked to provide reference for the scientific management of biological products.

KEY WORDS: laboratory information management system; FineReport; biological products; trend analysis

基金项目: 吉林省科技发展计划项目(20220505023ZP)

Fund: Supported by Jilin Province Science and Technology Development Plan Item (20220505023ZP)

*通信作者: 王坤, 高级工程师, 主要研究方向为医药领域的信息化应用。E-mail: 412160344@qq.com

*Corresponding author: WANG Kun, Senior Engineer, Jilin Institute for Drug Control, No.657, Zhanjiang Road, Changchun 130033, China. E-mail: 412160344@qq.com

0 引言

在生物制品的生产、检验过程中存在大量相关的数据,由于低数字化管理给岗位人员和管理人员的工作带来极大的不便。为提高数据的分析能力和机构竞争力,同时提高工作效率和组织质量,行业内监控生物制品质量和一致性的手段之一是利用趋势分析的方法。趋势分析一方面有助于生产企业以及各实验室设定合理的控制警戒限度和纠偏限度,为常规的检测结果提供评价标准依据,另一方面也有助于各实验室对实验误差、仪器设备故障和异常或超标结果进行调查,了解产生问题的真正原因,有针对性地消除偏差和加强培训,在提高检验结果可靠性的同时避免造成更大的损失^[1]。

在趋势分析的过程中,应用计算机软件,如 SPSS、Minitab、Joinpoint、Excel,对产品质量数据进行统计和可视化展示,可以发现数据中的潜在规律和趋势^[2-5]。大部分用户使用 Excel 系统,利用插入图表方式实现数据的监测和分析,此种方法简单易用,但也存在一些问题:趋势分析会细化到不同的机构、品种、检验项目,大量数据的记录和复杂报表的编制给操作人员工作带来沉重的负担;趋势分析是随着时间的推移而发生变化的,即使同一个检验项目也会由不同的检验人员完成,虽然检验数据的产生是实时的,但不同的人手动录入到 Excel 系统中却是不定时的,一方面准确率得不到保障,另一方面分析图根据权限共享也很难做到。部分用户使用 SPSS 等专业数据统计软件,这类工具功能强大,具有独立的门户机制,但很难将生成的 HTML 与其他软件做到无缝嵌入,必须带出其产品本身的报表管理界面,使用户操作烦琐。

本研究以某实验室为背景,以实验室信息管理系统(laboratory information management system, LIMS)和帆软报表工具(FineReport, FR)为基础,提出一套用趋势分析系统对检测数据进行采集、清理、汇总,并提供统计、分析和可视化监控、报警的方案,利用该方案追踪其在生产和检验中的稳定性和变异性,形成数据驱动的管理方式,对产品的质量控制在科学决策^[6],为生物制品的科学管理提供参考。

1 系统架构

趋势分析系统的设计,分为仪器设备、数据采集、数据库、应用程序4层。如图1所示。

1.1 仪器设备层

随着现代生产和通信技术的发展,各类仪器设备逐步走向集成化和智能化。在大型仪器的使用过程中,往往涉及各类传感器、监测器和控制单元的频繁交互。而常见的串行、并行通信在构建通信网络完成大规模集成时,存在反馈速度慢、线路拥堵等诸多问题。相比之下,拥有高速率、大数据量、远距离

传输的网络通信,是一种更适合于大规模集成时完成数据传输和整合的通信方式,所以将所有设备联入互联网成为当前发展的必然趋势^[7]。在本研究中,经过对仪器标准配置接口的分析后,按照其数据采集的类型以不同的方式联至以太网中。

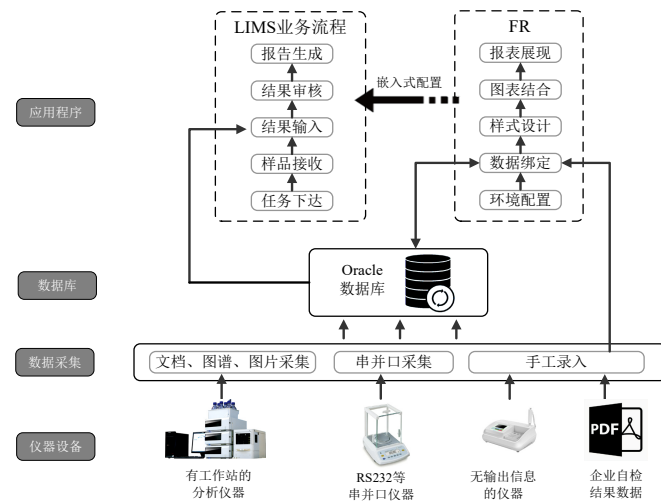


图1 生物制品趋势分析系统架构

Fig.1 Architecture of trend analysis system for biological products

1.2 数据采集层

面对实验步骤的复杂程度一再提升,样品量规模的持续增加,实验室智能化不再是片刻的需求,而是长远的规划。为了确保检验业务顺利高效开展,提高实验室合规化与公正性,各类实验室开始通过仪器的数据采集、数据自动分析和仪器参数智能控制等技术,实现实验室的可追溯、无纸化等全智能数据管理模式。通过研究生物制品涉及仪器的数据输出形式,基本确定了简单调用通信函数、编辑脚本调用、图谱解析、文件附件、CDS 接口对接等5种数据采集方式,最大化实现仪器设备联机的数据智能化处理。

1.3 数据库层

数据库中的架构作为整个系统的基础,对数据的存储、管理和系统运行起着至关重要的作用。本系统采用分布式架构,将数据存储在多台、多地服务器上,增加系统的可用性和可扩展性。同时管理人员持续做好索引设计、监控用户存取访问、制定备份计划等数据库的各项管理,提高整个系统的性能和稳定性。

1.4 应用程序层

LIMS 用于对实验室人、机、料、法、环、测各个环节进行全面量化和质量管理,实现检测工作的规范化、自动化和无纸化^[8]。目前在化工、医药、机械制造等多个行业得到广泛应用。LIMS 可以很好地实现查询功能,但报表能力较弱,一方面无论是从格式还是统计的内容,都无法从综合角度体现各类数据的相关性,另一方面需求的不断变化,促使报表要有灵活性、易修改性,因此报表开发工具 FR 极大弥补该方面需求。FR 乘

持零编码的理念^[9],可快速灵活地制作复杂报表,同时也可搭建统一的数据决策和管理平台,让更多的数据用于业务经营和决策^[10],也能做到与各类业务系统的无缝集成,加快项目的开发,节约开发的人力和成本。本系统利用 LIMS+FR 方式,数据从 LIMS 中获得,图表通过 FR 展示,能较好地实现趋势分析。

2 应用程序设计

2.1 设计思路

在生物制品的检验中,有很多检验项目,其中数值型检验项目根据检验标准规定确定其指标的取值范围,检验结果高于上限或者低于下限均会被判定为不合格。基于可量化的检验数据服从正态分布,并且平均值符合或十分接近产品目标指标的原则,本系统采用统计学上质量控制 3 σ 原则对检验结果进行趋势分析。3 σ 原则:符合正态分布的概率密度函数,当 $\mu=0$, $\sigma=1$ 时,称随机变量 X 服从标准正态分布:数值分布在 $[\mu-\sigma, \mu+\sigma]$ 中的概率为 0.6826;数值分布在 $[\mu-2\sigma, \mu+2\sigma]$ 中的概率为 0.9544;数值分布在 $[\mu-3\sigma, \mu+3\sigma]$ 中的概率为 0.9947,可以认为服从正态分布的随机变量取值几乎全部集中在 $[\mu-3\sigma, \mu+3\sigma]$ 区间内,超出这个范围的可能性仅占不到 0.3%^[11]。检验数据中某个检验项目的结果,当数量少于 40 个,可做数据监测。可在至少 6 连续批定量检验结果的基础上计算平均值和标准差值,每个新的结果可以添加并重新计算平均值和标准偏差(standard deviation, SD)值;结果的数量等于或超过 40 个,以最初的 40 个结果确定平均值,新增的检验结果不再进行计算,避免因引入新的检测数据而掩盖检测结果发生偏移的情况。当连续批次检验结果明显出现偏移时,经分析调查,确定需要重新计算平均值的可重新计算,并将分析调查记录存档。

控制图是一种基于统计显著性原则进行过程控制的图形工具,有助于评估过程是否已经达到或持续处于“受控”状态,识别异常趋势,给出有价值的改进措施^[12]。本系统的趋势分析运用休哈特控制图的分析方法进行判断与确认,设定警告限和行动限,发现产品质量或是生产过程中发生的问题及偏差,反映生产和检验过程的稳定性,预见产品质量的趋势^[13]。以纵坐标为检验结果,横坐标为样本号绘制趋势分析图,根据试验的变异度和精确度确定每个产品的警戒限和行动限。通常平均值 $\pm 2SD$ 和平均值 $\pm 3SD$ 被设为警戒限和行动限。对于某些检测结果变异范围较大的指标,如:平均值 $\pm 2SD$ 、平均值 $\pm 3SD$ 已超出质量标准的情况,可依据制品检验实际情况分析制定警戒限和行动限。

2.2 源数据处理

生物制品可以进行趋势分析的常规项目有很多,如各个检验指标的检测数据、各种原料、半成品的检测数据、检验过程中各类仪器、设备或系统的故障记录数据、阴性对照或阳性对照不合格的结果数据等。因采集的数据其数据类型各不相同,需要在 LIMS 中设置相应程序对采集数据进行审查、甄别和校验,识别出“脏数据”,避免无效值、缺失值、格式异常值、重复值影响分析结果的准确可靠性^[14]。

采集的所有数据,其评价指标因为存在着不同的量纲和量纲单位,也会影响到数据分析的结果,运用数据的归一化处理可以消除量纲对最终结果的影响,使各指标处于同一数量级,适合进行综合对比评价。归一化不会改变数据分布,只是对数据进行线性无量纲处理,通过相应的缩放和平移使数据发生改变,但并没有改变原始数据的排列顺序^[15]。本研究通过文献调研在机器学习和数据分析领域内选取了几种数据处理的方法,最终选择 MinMaxScaler 函数法进行归一化数据处理。

3 应用效果及分析

根据应用程序的设计,本系统可实现趋势的可视化。

图 2 为某产品 FR 报表工具的开发界面。图 3 为企业和检验机构两个部门对某产品“蛋白质含量”指标的检验数据趋势控制图。当指标数值偏离常规数值范围,情况满足预先定义的异常情况时,图中数据可出现红色预警。本系统对需要引起重视的偏差进行了规定:1 个数据超过了 2SD 的范围;连续 5 个数据中有 4 个数据超过了 1SD 的范围;6 个连续的数据递增或递减;8 个连续的数据位于中线的一侧。当控制图出现上述情况时,应启动预防和纠偏程序,分析查找导致趋势偏移的原因,评估对产品质量的影响风险。

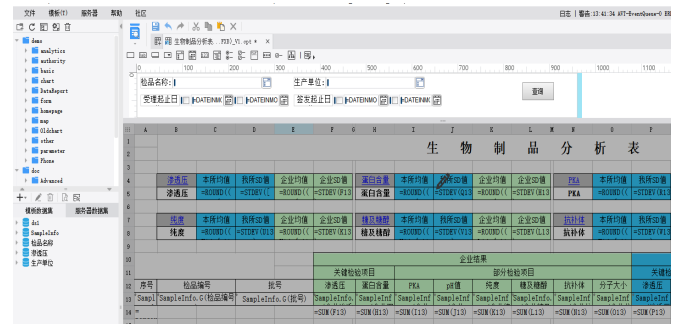


图 2 FR 报表开发界面

Fig.2 FR report development interface

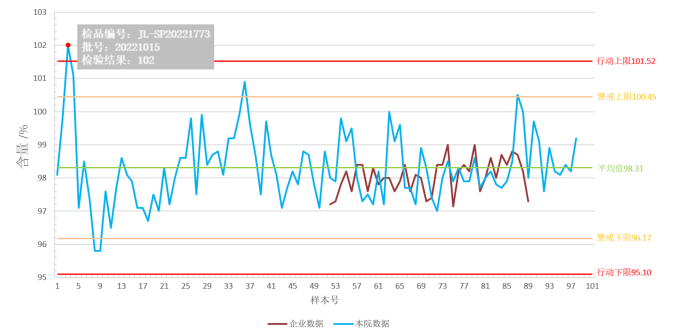


图 3 “蛋白质含量”趋势控制图

Fig.3 Trend control chart of “protein content”

图 4 为某产品“渗透压摩尔浓度”指标的原始检验数据趋势控制图,数值区间为 [260, 340];图 5 为对图 4 的数据集经过 MinMaxScaler 函数归一化处理后绘制的趋势控制图,数值区间为 [0, 1],从两个图中可以看出数据的可视化趋势一致。

图 6 为某产品“渗透压摩尔浓度”“蛋白质含量”“激肽释

放酶原激活剂含量”3个检验项目的归一化趋势控制图。将同一产品的多种关键检验项目数据绘制成一个趋势图的优势在于可以更加直观地看到该产品关键检验项目的整体趋势。从图6中可以看出,“渗透压摩尔浓度”与“蛋白质含量”两项之间整体趋势关联度较为密切;而这两项与“激肽释放酶原激活剂含量”密切度较低。经过资料查阅得知:白蛋白是血浆中含量最高的蛋白质,维持血浆胶体渗透压、保持血管内外液体平衡是白蛋白在人体内的生物学主要功能之一。因此该理论验证了上述对数据的趋势分析,具有一定的应用价值,为进一步判断和预测产品的质量提供更加可靠的依据。

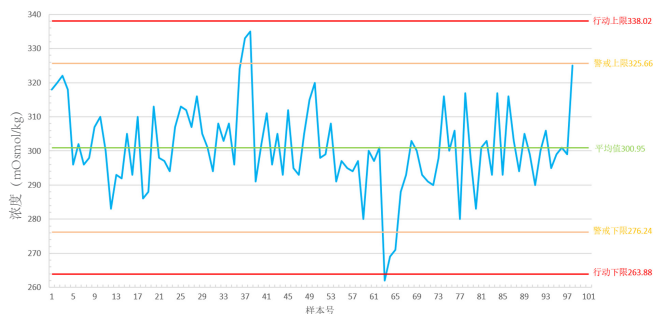


图4 “渗透压摩尔浓度”原始数据趋势图

Fig.4 Original data trend chart of “molarity of osmotic pressure”



图5 “渗透压摩尔浓度”数据归一化趋势图

Fig.5 Data normalization trend chart of “Molarity of osmotic pressure”

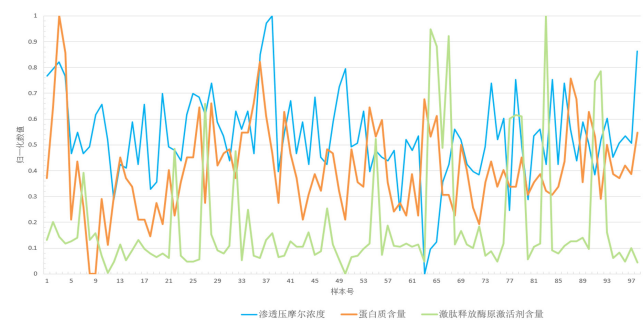


图6 3个指标数据归一化趋势图

Fig.6 Data normalization trend chart of 3 inspection items

4 结束语

LIMS+FR技术,本质上是对LIMS系统应用后产生的数据资源以及这一系统所具有的收集和传递信息的功能进行开发和利用的技术。基于这两者提出的方案,将数据实时采集与报表

实时监测结合起来,提高了数据记录的可靠性和可追溯性,杜绝人为捏造及修改数据的行为,也使数据共享更加准确、高效,利用信息化手段,强力助推实验室管理水平的提升。未来将结合生产企业实验室实时数据采集的功能进行深入探讨,挖掘趋势分析的技术,及时、客观地分析生物制品的质量和生产过程的一致性。

参考文献

- [1] 微生物检测数据的趋势分析和统计 [EB/OL]. [2017-09-12]. <https://max.book118.-com/html/2017/0912/133500455.shtml> [2023-11-30].
- [2] 王珊珊,李亚南,赵丹,等.我国A群C群脑膜炎球菌多糖疫苗有效组分质量趋势分析研究[J].药物评价研究,2021,44(09):1897-1901.
- [3] 余云龙,陶启果.基于Minitab的数据分析和可视化系统的设计与开发[J].印制电路信息,2023,31(S2):102-106.
- [4] 顾佳玲,高慧,沈一真,等.2009-2019年上海市长宁区病毒性肝炎流行病学特征及发病趋势分析[J].公共卫生与预防医学,2023,34(06):80-84.
- [5] 林雪倩,李嘉兴,兰洪彬,等.南充市2008-2022年丙型肝炎流行特征及趋势分析[J].现代预防医学,2023,50(21):4009-4014.
- [6] 魏凯峰,刘伟何,盛建村,等.基于FineReport的生产报表自动化系统设计与实现[J].科学技术创新,2023,(04):93-96.
- [7] 余毅.嵌入式多路串口服务器设计与实现[D].太原:中北大学,2023.
- [8] 李帅,廖彬,李文莉,等.药品检验行业LIMS系统ELN模板格式化解决方案[J].海峡药学,2021,33(07):59-65.
- [9] FineReport产品简介[EB/OL]. [2023-11-15]. <https://help.fanruan.com/fine-report/doc-view-1162.html> [2023-11-30].
- [10] 知乎.你用过Excel,却不知还有一款报表神器“FineReport”[EB/OL]. [2022-11-03]. <https://zhuanlan.zhihu.com/p/30404690> [2023-11-30].
- [11] 林伟强,郭晓敏,周宗梁,等.基于检验数据的药品质量风险预警模型研究[J].软件,2018,39(12):127-130.
- [12] 欧盟医疗器械上市后风险趋势分析方法研究[EB/OL]. [2021-11-11]. <http://www.anytesting.com/news/1927911.html> [2023-11-30].
- [13] 马霄,孟丽,谭亚军,等.趋势分析在疫苗产品批签发管理中的应用[J].中国药事,2014,28(07):750-753.
- [14] 林森,郭旭东,周雪晖,等.茶叶质量安全分析预警系统研究[J].质量与标准,2020,(01):60-67.
- [15] 知乎.数据处理中的标准化、归一化,究竟是什么?[EB/OL]. [2021-02-18]. <https://zhuan-lan.zhihu.com/p/351243330> [2023-11-30].

作者简介

王坤,硕士,高级工程师,主要研究方向为信息化在医药领域的应用。