

医学大数据研究进展及应用前景

赵京瑞

东北大学医学与生物信息工程学院 辽宁 沈阳 110016

【摘要】：近几年来，医学大数据在不断发展之后，在临床上运用越来越广泛，它当前在群体上可以对一些疾病进行预防和治疗，在特殊疾病机制的阐述中提供有力支持，在个体患者的诊疗决策中发挥其关键作用，以及构建起临床决策支持系等方面。

【关键词】：医学大数据；临床决策支持；精准医学

1 引言

随着社会经济的快速进步，现代化医学技术不断进步，正在面临着重要的社会变革。此类社会变革的重要动力来源于靶向药物、计算技术、组学技术等等，对于各种类别的疾病防御以及诊疗确定颇为重要。由于临床数据多种多样，受到外在因素的综合影响，可为关系的预防工作以及治疗工作奠定基础，选择有效的治愈方案。目前，国内的医学数据可综合借助先进技术得到重要数据，取得比较显著的成绩。不论是对患者而言，还是对健康人群而言都是十分有益的。最近几年，医学大数据可以为人群中的疾病预防和治疗做出较为有效的评价、在一些特殊疾病的发病机制的叙述中也可以提供一些帮助以及帮助医生对个体患者的疾病进行治疗。

2 医学大数据的含义以及特征

大数据具有数据量多，数据结构复杂，依靠传统的工具难以进行处理的特点。处理的过程主要包括：数据的获取、储存、格式化以及分析等过程。当前的生命科学领域中涉及的大数据与我们熟知的社交媒体或是经济上的大数据有明显不同。它具有多样性，有较高的数据传输速度，其数据可靠性不稳定，对医疗工作流的各个环节构成挑战，对计算技术构成挑战、难以提取有意义的信息、临床及基础医学数据共享有困难性以及生物信息学专业人员不足等特征。

3 医学大数据对群体层面的疾病预防和诊疗的重要意义

众所周知，医院中的临床数据可以在一定程度上分析某一种疾病在不同人群中的情况，可以监测出该疾病的患病率好发病趋势，对医学大数据技术而言，它可以被用于人群中某种传染性疾病的监测。公共卫生界十分关注人群的疾病以及健康的情况，最近几年其最关注的议题之一是：如何更为有效的从谷歌的检索数据中得到流行病毒的监测，如果能从检索数据中得到高概率的预测，可以提高各种流感疫苗的研发效率，增加人群中高危人群的接种情况，可以提前做好各种

重症流感风险的准备工作。为了更好的应对埃博拉病毒的大流行，各国政府更加注重医学大数据的流行性疾病监测和分析工作，人们可以从各个发病地区的检索数据中进行病毒流行预警和各种流行趋势分析。服务器上的检索数据可以为公共卫生工作人员更好的对传染性疾病进行监测，并开展相关工作提供重要支持。

临床大数据可为传染性疾病的综合检测以及防御奠定基础，有效帮助研究工作者开展各项工作，给予研究者疾病分析数据支撑，致使其明确不同疾病之间的联系。就曾有研究者为了更好的了解风湿关节炎的发病机制，清楚其危险因素，经过层层分析，最后得到抗环瓜氨酸肽抗体对类风湿关节炎的评估可靠性更高。有研究者分析了来自日本三千多家医院经历了右半结肠切除术的电子病历，通过分析成功研发了一个风险预测模型。

对医学大数据进行有效分析，还可以在宏观层面上为各种重大决策提供帮助。研究者在对美国近些年的电子病历进行综合分析可知，患者在住院中所应用的超声心动图应用比较普遍，加之心血管重症入院患者经常在住院中应用超声心动图进行心率测量，这已经成为了重要的应用举措。与此同时，美国学术界长期存在着医院超声心动图会被滥用的情况，这已经是社会普遍认同的观点了。当这一研究结果公布之后，不仅修正了这一错误观念，而且还改进了临床诊疗质量、构建和完善各种慢性病管理指标体系以及对医保付费政策进行有效调整，这些都为患者在临床诊疗过程中避免了不必要的开支。

4 医学大数据为特殊疾病把机制阐述提供强有力支持

社会受众对于特殊疾病的认知并不完全，而且会受到样本数量以及内容的限制，加之存在外在因素的干扰。而在医学大数据技术全面应用背景中，特殊疾病机制完善存在一定的改善。在一项研究中，研究人员对能得到的现患有机例疾

病的病人进行外显子组数据分析,最终得出了该罕见病的多个已报到的突变为点的致病可能性。原先医学界往往认为人类一旦患上罕见疾病,基本上只能等死,甚至可能会先死于神经退行性疾病,但在本次研究过程中,已经证实突变位点所导致的患病概率较低,甚至会减轻疾病携带者对于死亡的恐惧,难以做到安乐死。国家科研机构建立的数据共享平台中有较多的数据样本,包含六万粉没有血缘关系的个体信息数据,基于此类数据,医学界的研究者可从对应的数据信息中进行样本分析,明确基因变异疾病出现的原因,为后续疾病诊断以及治疗奠定基础。

5 医学大数据在个体患者的治疗决策中提供有力支持

5.1 通过对现有的医学大数据进行挖掘分析,逐步完善精准医学知识体系

在对医学数据进行综合分析时,对于新得到的数据需要采取高效的措施,利用计算机技术对其进行验证,明确数据的真实性与准确性。由于医学大数据中的基因组、转录组、蛋白质组等类别属于研究人员创新性研究的重要资源,可通过对数据库类别的综合把关实现对人体不同类别疾病的有效分析。一般来说,信息共享平台包括 1000Genomes 项目、ENCODE 项目等,工作人员可利用数据库中存有的信息数据与其他数据之间进行比对,实现对数据的深度分析,运用主体因素分析方法找出不同类别数量下的生物学功能转录方法,明确基因表达谱中存在的差异。在综合探究中,工作人员可构建符合现阶段医学大数据特点模型,应用模型预测基因生物学功能,了解基因组因素与患者之间的关联性。在具体的实践探究中可知,用于预测肿瘤结构的基因数据与人体 DNA 数据存在联系,在其规律探究中或许能够寻求新的治疗方案。除此之外,需要根据具体情况构建完善的医学知识体系,明确可能出现的安全隐患,增强医学数据应用实效,

5.2 在大数据的基础上构建起临床决策支持系统

由于临床诊断的结果往往受到样本量以及环境等方面

的差异影响,若是仅从宏观层面对医学大数据个体进行临床治疗,很大概率会出现误诊,甚至可能发生意外事故。对此,要想实现在全方面诊断中应用医学大数据,最关键的是构建起临床决策支持系统,利用该系统将各个维度的数据进行有效整合,在该系统中出现一个具有专一化的治疗方法,为医生以及患者提供较为精细化以及个体化的治疗指导,最为直接的是可以表现在患者每天的用药量。对于哮喘这一疾病来说,每一个患者的临床表现都存在较为明显的异质性,这种差异会体现在年龄、性别、气道高反应的严重程度等方面,在治疗患者时可以结合患者的人口学数据、诊断数据、各项肺功能评估结果等方面提供初步的治疗方案;可以对于患者利用便携式设备收集患者每天相关数据反馈指标,最后结合每天的用药剂量种类,当天的环境花粉检测数据,所居住地区的 PM2.5 数据等,使用机器学习模型,对数据进行综合考量,修正相关参数,筛选出符合要求的指标和参数,并在信息系统中实现对患者每天用药量的自主计算。构建临床决策支持系统在于最大程度的控制患者出现急性哮喘发作概率,从源头上遏制社会受众由于哮喘出现住院的情况,提高哮喘患者的治愈效率,缩短发病周期。这种类别的临床决策支持系统一旦被应用,可实现对肿瘤、高血压、糖尿病等慢性疾病进行治疗,减轻医务工作者的工作量,让医务工作者有更大的精力去攻克当前的疑难重症,该系统也可减轻患者的疾病负担,延长患者的寿命。

6 结束语

综上所述,医学大数据在实际应用中效果颇好,已经取得了显著成绩。但依旧面临一系列阻碍,主要包括技术限制、资金应用量大、数据处理效果低以及技术要求高等,在我国的发展中还存在着政策法规,伦理内容以及安全技术上的不足之处,即使国内引发的系统对于应用基因测序技术以及计算科学等存在优势,但由于缺乏临床经验,很难对各种数据进行能动分析,因此我国在该道路上还面临着许多困难。但相信经过不断努力,必定积极推进医学大数据的全面应用,为人民带来各项福祉。

参考文献:

- [1] 洪娜,刘飞,张梦阳,张渝翔,刘超,张富博,王芳,唐燕平,刘欢妹,陈萌.OHDSI 通用数据模型在肿瘤大数据中的应用探索[J].中国数字医学,2021,16(11):24-28.
- [2] 王昕玥,渠鸿竹,方向东.组学大数据和医学人工智能[J].遗传,2021,43(10):930-937.DOI:10.16288/j.yczz.21-215.
- [3] 刘巧红,孙丽萍,李建华,凌晨.医学院校大数据专业课程体系建设的理论与实践——以上海健康医学院为例[J].中国医学教育技术,2021,35(05):549-553.