

# 精准医学信息学对临床医疗和卫生领域的促进和影响\*

王丹 沙岩 胡俊峰 马凯

(徐州医科大学医学信息学院 徐州 221004)

**[摘要]** 介绍精准医学计划以及精准医学信息学为医疗卫生领域带来的变革,分析精准医学信息学发展存在的问题并指出相应对策,包括制定并完善精准医学法律规范、建立精准医学信息共享平台。

**[关键词]** 医学信息学;精准医学;人类基因组计划

**[中图分类号]** R-056 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2018.07.005

**Improvement and Impact Exerted on the Clinical Care and Health Domain by Precision Medical Informatics** WANG Dan, SHA Yan, HU Jun-feng, MA Kai, School of Medical Informatics, Xuzhou Medical University, Xuzhou 221004, China

**[Abstract]** The paper introduces the precision medicine program and the transformation brought by precision medical informatics to the medical health domain, analyzes problems existing in the development of precision medical informatics, and points out corresponding measures, including formulating and perfecting laws and regulations of precision medicine, establishing the information sharing platform for precision medicine.

**[Keywords]** Medical informatics; Precision medicine; Human Genome Project (HGP)

**[修回日期]** 2018-04-19

**[作者简介]** 王丹,博士,讲师,发表论文11篇;通讯作者:沙岩,硕士,发表论文8篇。

**[基金项目]** 江苏省博士后科研资助计划“胶质瘤细胞中特征性融合基因介导 GDNF 基因表观修饰的实验研究”(项目编号:1601080C);江苏省高校自然科学研究面上项目“胶质瘤细胞特征性融合基因介导 GDNF 表观修饰的研究”(项目编号:17KJB310015);江苏省高等学校大学生实践创新训练计划项目“肠杆菌科中抗生素抗性相关基因家族的进化与功能研究”(项目编号:201610313078H);徐州医科大学优秀人才科研启动基金“胶质瘤细胞中组蛋白乙酰化介导 GDNF 基因高转录的研究”(项目编号:D2015001)。

## 1 引言

与经典的物理学、数学等学科不同,精准医学信息学是一新兴边缘学科,主要由生物医学、信息科学、数理科学等众多相关学科相互交融所形成,是利用计算机技术、信息技术研究生物医学规律的学科<sup>[1]</sup>。精准医学信息学主要关注数据采集、处理、分析和解释,以期帮助医生、患者和科研人员更好地理解健康、疾病与生物医学大数据的关系。通过计算机技术、信息技术等手段对生物医学数据进行采集、处理、存储、传播、分析和解释,以期更好地理解健康、疾病与生物医学大数据的关系。精准医学信息学的起源可以追溯到人类基因组计划,现已成为当今生物医学领域的重大前沿之一,

在多个核心技术领域支撑精准医学的研究进展及临床实践。随着生物医学大数据采集和分析技术的发展, 广大人群的生理状态和疾病发展过程的数据可以被详尽地采集和分析, 帮助人们更好地理解健康与疾病的关系。

## 2 精准医学计划

为更深入地了解和认识自身制定宏伟的人类基因组计划 (Human Genome Project, HGP)<sup>[2]</sup>, 其与曼哈顿原子弹计划、阿波罗登月计划一起被列为人类自然科学史上的 3 大计划。随着人类基因组计划的完成, 美英中等国先后发动国际人类基因组单体型图 (HapMap) 计划<sup>[3]</sup>、癌症基因组图集 (The Cancer Genome Atlas, TCGA) 计划<sup>[4]</sup>、DNA 元件百科全书 (ENCODE) 计划<sup>[5]</sup>、千人基因组计划<sup>[6]</sup>以及精准医学 (Precision Medicine) 计划<sup>[7]</sup>。精准医学计划希望利用基因组、蛋白质组、转录组、代谢组、宏基因组等组学技术, 针对大样本群体与特定疾病类型进行生物医学大数据的分析、鉴定和应用, 从而精确找到疾病发生的原因和诊疗的生物标志, 对疾病不同状态和过程进行亚分类, 通过精准靶标对疾病和特定患者实现个性化精准治疗的目的。精准医学是生物技术和信息技术在医学中的交汇融合, 包括精准诊断、预防和治疗 3 个方面。我国在“十三五”发展规划中将精准医学列为国家重点研发项目并已正式启动。到 2030 年前我国将在精准医学领域投入 600 亿元, 以北京协和医院、华西医院等为代表的大型三甲医院正在积极筹建生物医学研究中心。如华西医院计划针对 100 万人口开展全基因组测序, 建立数据库和样本库, 研究疾病发生机制, 为精准医疗奠定基础。

## 3 精准医学信息学为医疗卫生领域带来的变革

### 3.1 云计算助力精准医学

由北京贝瑞和康生物技术有限公司与阿里云共同打造的数据云以海量中国人群基因组数据为核心, 有望实现对中国人群个人基因组数据的精准解读, 目前已包含超过 40 万份基因组数据。通过对其数据资源的深入挖掘, 能够进一步揭示群遗传多

态性的分布情况, 显著提升遗传疾病诊断的效率和精准程度。基因组、蛋白质组、转录组、代谢组、宏基因组等海量组学数据对服务器计算性能提出很高要求, 因此云计算正成为生物医学界的关注焦点。基于阿里云强大的批量计算能力, 利用云计算的优势可以显著降低成本, 极大提高数据分析的速度。基于中国人群基因组数据的数据云将对我国临床医学的预防、精准诊断和治疗的发展产生巨大的推动作用。

### 3.2 生物公司抢占基因检测市场

来自波士顿的奕真生物 (Veritas Genetics) 宣布在杭州建立研发中心, 开发适用于亚洲患者群体的新一代测序技术, 开展适合亚洲人的基因检测服务, 如在亚洲人群中发病比例较高的  $\alpha$ -地中海贫血相关基因的研究<sup>[8]</sup>。该公司目前的 myBRCA 检测技术已经在欧洲获得 CE 认证, 可以用于检测与乳腺癌和卵巢癌发病相关的 BRCA 基因异常<sup>[9]</sup>。该检测方法也已登陆美国市场。除奕真生物公司外, 其他测序公司如阿金纳生物科学 (Agena Bioscience) 也纷纷在中国建立分公司以进一步开拓市场业务。

### 3.3 基因检测大众化

目前精准医学工作主要是由科研机构的实验室或几个国家性的项目在进行。世界领先的基因测序公司宜曼达 (Illumina) 瞄准普通消费者这一隐形的巨大市场, 通过创建庞大的 APP 应用商店, 即螺旋 (Helix) 公司, 以定制化收费的方式有效解决全基因组测序费用过高的问题, 消费者只需 99 美元就可以获得如家谱分析之类的特定基因信息。在英美等国家, 基因检测已经常态化。在英国政府每年投入 10 亿英镑用于国民的基因检测项目; 在美国每年有 800 万人选择基因检测。而在中国目前基因检测最为广泛的应用是产前无创产筛, 可以预防新生儿唐氏综合症。在医学上基因与癌症的紧密关系已被广泛发现和接受, 与此同时随着国内基因检测技术及市场机会的成熟, 针对大众的基因检测预示着极大的市场潜力。目前达安基因推出的大众基因产品主要包括天赋、女性两癌以及重大疾病易感基因检测, 极大推动基因检测技术在癌症等重大疾病预测和防治中的重要作用。特别是其中的 B+ 守护

主要是针对女性群体高发难治的乳腺癌、卵巢癌的相关发病基因检测,广大女性从中受益。基因检测的大众化应用时代已经到来。

### 3.4 开启无创胎儿检测新时代

在临床实践中胎儿染色体异常是控制和预防出生缺陷的重点。传统产前诊断依赖于羊膜穿刺术等有创取样手段,对母体和胎儿均存在一定风险,而且假阳性比较高<sup>[10]</sup>。随着 DNA 测序技术的突破,可以通过孕妇的血液来检测胎儿的 DNA 痕迹,基于高通量测序技术可以判断是否有染色体异常,在孕 5 周即可检出,与传统产前筛查相比无创产前筛查更灵敏,且假阳性和假阴性结果概率更低,具有良好的应用前景<sup>[11]</sup>。

### 3.5 精准医学信息学助力肿瘤个体用药

精准医学信息学对制药行业也有巨大的推动作用。阿斯利康通过构建生物信息学系统,深入地分析临床数据并将其汇总成表格供医务人员实时查询,此外还会与肿瘤药物信息进行数据比对并相应调整临床治疗方案,为医务人员提供精准治疗信息,实现肿瘤患者个体化用药。英国卫生部下属的英国基因组学(Genomics England)联合世界多家生物医药公司将测定超过 10 万名英格兰人的基因组信息,以期找到开发肿瘤药物和罕见病药物的新方法,同时更清楚地了解一些疾病的致病机理。随着人类基因组学的发展药物基因组学领域得到迅猛发展,2004 年以来 40 多种肿瘤分子靶向药物已被批准并用于临床治疗,在提高肿瘤临床疗效方面取得巨大进展<sup>[12]</sup>。

## 4 存在的问题

### 4.1 缺乏统一的行业标准,医学信息数据共享难以实现

精准医学的实现需要建立生物医学信息资源库,以便能对大规模人群的海量生物医学大数据进行分析和高效整合。当前的科学研究已经深入到基因组、转录组、蛋白质组和宏基因组水平,对生物样本的质量提出更高要求。然而目前生物样本库的标准尚未建立,根据不同标准产生的临床实验结果

千差万别,难以实现数据共享。此外,目前我国在国家层面尚缺乏统一的专门机构负责医院和科研机构等生物医学信息的收集、分析和管理工作。很多大型医疗机构各自筹建医学信息研究平台。这种分散、重复的方式由于缺乏统一的数据标准,不但造成巨大的人力、物力、财力浪费,而且难以实现数据共享。

### 4.2 缺乏相关法律法规和监管机制

精准医学的发展目前尚处于起步阶段,缺乏完善的法律法规和监管机制。如果基因检测走向临床,首先面临的就费用结算等一系列问题,美国目前已有保险公司将多家基因检测产品纳入投保范围,美国医疗技术政策中心也表示会努力解决将外显子组合全基因组测序纳入医保范围的问题,而在中国尚未出台相应的政策。此外检测结果是否能与临床结果对接,是否人人都应该做基因检测,检测结果解读等方面也缺乏统一的技术标准与质量监管,不同机构、不同批次的检测结果存在质量参差不齐等问题。由于缺乏统一的行业标准以及质控、质检环节,现在的精准医学还并不精准。

## 5 相应对策

### 5.1 制定并完善精准医学法律规范

鉴于精准医学发展中的诸多问题,基因测序产业在我国经历无监管、被叫停及试点申报的过程。2014 年以前基因检测行业在我国处于无监管状态,基因检测设备基本完全依靠进口国外设备。2014 年国家食品药品监督管理总局、原国家卫生和计划生育委员会联合发布通知,首次在国家法律层面明确提出包括产前基因检测在内的基因测序相关产品和技术属于当代前沿技术研究范畴,涉及伦理、隐私和人类遗传资源保护、生物安全以及医疗机构开展基因诊断服务技术管理、价格、质量监管等问题。为保证公众使用基因测序诊断产品的安全、有效,国家食药总局、原国家卫生计生委积极组织相关领域专家调查论证<sup>[13]</sup>。经过 1 年的申报论证审批,2015 年原国家卫生计生委公布第 1 批医疗及临床试点单位,如华大基因和达安基因<sup>[14]</sup>。此外精准医学关系到全民健康乃至国家基因安全。我国在大力发展精准医学计划的同时应做好两件事,一是加强个

人隐私保护,二是严格规范临床试验。基因信息采集涉及大量的个人隐私,与社会伦理、医保、疾病诊疗、个性化健康等领域密切相关。在构建生物学大数据平台的同时应建立完善的隐私保护机制,这需要法律规范以及严格监管。因此为促进精准医学在我国健康有序的发展,政府还应从法律层面明确相关部门具体职责,监管部门及相关行业协会建立并逐步完善精准医学的行业标准,对于不符合国家行业标准的机构应按照法律和政策规范给予相应处罚。

## 5.2 建立精准医学信息共享平台

要解决目前分散、重复收集生物学数据的问题,需要政府、卫生部门在国家层面成立精准医学信息研究中心,构建生物学大数据平台的基本架构和技术体系,避免零散、重复建设。在精准医学信息数据共享的前提下科研人员可实现生物学数据的深入挖掘,为临床决策提供科学依据。基因与癌症的紧密关系已被医学界发现和认可。生物学大数据和信息技术是生物科学、精准医疗和健康产业的重要基础,将会引导个性化诊断和治疗。精准医疗大数据、个人基因测序的普及化正带来全球健康产业的大变革。有研究报告称未来 5 年基因测序市场年增速将超过 20%<sup>[15]</sup>。此外随着基因测序成本的显著降低,很多从未接触过基因测序的人能够有机会了解自身和家族的基因构成,有助于精准医疗的推广。

## 6 结语

当前科学家们已经意识到基因组学在个性化医疗领域的发展前景,而基因组学转化到医学应用主要依赖对大数据的分析。由谷歌、华大基因国际著名公司组成的全球基因组和健康联盟正在开发数据共享的标准形式,以促进临床及组学方面数据的共享<sup>[16]</sup>。这意味着未来将有较为统一的标准评估基因测序的检测方法,通过基因数据的共享,来自世界各地的研究人员将可以反复检验某种方法所提供的结果是否可靠,最终形成公认的“金标准”。

## 参考文献

1 Levine A G. An Explosion of Bioinformatics Careers [J].

Science, 2014, 344 (6189): 1303 - 1306.

2 DDelisi C. The Human Genome Project: the ambitious proposal to map and decipher the complete sequence of human DNA [J]. American Scientist, 1988, 76 (5): 488 - 493.

3 Wei Y L, Wei L, Zhao L, et al. A Single - tube 27 - plex SNP Assay for Estimating Individual Ancestry and Admixture from Three Continents [J]. International Journal of Legal Medicine, 2016, 130 (1): 1 - 11.

4 Dellinger T H, Smith D D, Ouyang C, et al. LICAM is an Independent Predictor of Poor Survival in Endometrial Cancer - An Analysis of The Cancer Genome Atlas (TCGA) [J]. Gynecologic Oncology, 2016, 141 (2): 336 - 340.

5 Consortium E P. The ENCODE (ENCyclopedia of DNA elements) Project [J]. Science, 2004, 306 (5696): 636 - 640.

6 赵辉,赵方庆. 基于千人基因组谱系数据的拷贝数变异识别与分析 [J]. 南方医科大学学报, 2015, 35 (6): 777 - 782.

7 王元火,贾旭明,张冬梅. 对话人类基因组计划领衔专家 J. Craig Venter [J]. 科学通报, 2017, 62 (8): 743 - 744.

8 杨阳,张杰. 中国南方地区地中海贫血研究进展 [J]. 中国实验血液学杂志, 2017, 25 (1): 276 - 280.

9 Oberguggenberger A, Sztankay M, Morscher R J, et al. Psychosocial Outcomes and Counselor Satisfaction Following Genetic Counseling for Hereditary Breast and Ovarian Cancer: a patient - reported outcome study [J]. Journal of Psychosomatic Research, 2016, 89 (8): 39 - 45.

10 魏贤达,吕卫刚,邬玲仟. 下一代测序技术应用于无创产前检测的现状与未来 [J]. 中华检验医学杂志, 2017, 40 (7): 489 - 491.

11 Barrett A N, Chitty L S. Developing Noninvasive Diagnosis for Single - Gene Disorders: the role of digital PCR [J]. Methods in Molecular Biology, 2014, 1160 (1160): 215.

12 Doroshow J H, Kummar S. Translational Research in Oncology - 10 yYears of Progress and Future Prospects [J]. Nature Reviews Clinical Oncology, 2014, 11 (11): 649 - 662.

13 周琴,郑芳秀,陈英苹,等. 高通量基因测序对血清学筛查临界风险孕妇的应用价值 [J]. 重庆医学, 2016, 45 (27): 3780 - 3782.

14 李晓明,杨莹,彭辉,等. 全基因组测序在医学应用进展 [J]. 基因组学与应用生物学, 2015, 34 (5): 1071 - 1075.

15 石铁流. 新技术能使 DNA 测序的成本降低多少? [J]. 科学通报, 2017, 62 (19): 2042 - 2046.

16 王文杰,侯艳,李康. 基因组学数据的网络构建与分析方法 [J]. 中国卫生统计, 2017, 34 (1): 177 - 180.