

# 精准医学对医学信息学提出的挑战与任务

弓孟春

(InterSystems 公司 北京 100022)

〔摘要〕 介绍精准医学的定义及其新的医疗模式, 指出临床决策支持系统是实现精准医学的核心步骤。医学信息学在精准医学的发展中应该发挥引导研究方向、保证知识的准确性和可重复性、改变知识传递的时效性以及构建知识的审查复核机制等关键作用。阐述在推进精准医学发展的过程中, 医学信息学需要完成的关键任务。

〔关键词〕 精准医学; 医疗信息技术; 临床决策支持系统; 知识库

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2016.01.001

**Challenges and Missions for Medical Informatics Brought by Precision Medicine** GONG Meng - chun, Innersystems, Beijing100022, China

〔Abstract〕 The paper introduces the definition of precision medicine and its new medical models, points out that the Clinical Decision Support System (CDSS) is the core step for the realization of precision medicine. In the development of precision medicine, medical informatics should play a key role in guiding the research directions, guaranteeing the accuracy and repeatability of knowledge, changing the timeliness of knowledge transmission and building the knowledge review mechanism, etc. Besides, it explains key missions for medical informatics in promoting the development of precision medicine.

〔Keywords〕 Precision medicine; Medical information technology; Clinical Decision Support System (CDSS); Knowledge base

## 1 精准医学的定义及新的医疗模式

### 1.1 定义

精准医学即充分考量患者在基因、环境及生活方式中存在的个体差异以达成最有效的疾病治疗和预防的医学模式<sup>[1]</sup>。精准医学核心理念是将多个与人体健康及疾病预防相关的度量模块的结果进行整合。这不仅仅涵盖临床数据和基因组数据, 更包括环境暴露、日常生活习惯及其他多种多样的数据。藉此可以对人体的疾病状态和发展过程进行更相近

的描绘和更为透彻的理解。在精准理解经过整合的来自各个模块的数据的基础上, 可以明确患者所在的疾病的特殊亚组, 从而制定更好的个体化治疗方案。

### 1.2 新型医疗模式

以哮喘为例, 大量的证据证实哮喘患者的临床表现存在显著的异质性。这种个体差异不仅体现在发病年龄、性别、与肥胖的关系、气道高反应性的严重程度等, 更重要的是对于不同药物的治疗反应也存在显著差异。血液、痰液的传统生化指标、基因组学及转录组学研究可以对同样诊断为哮喘的患者进行进一步的亚群分组, 从而选择出最佳的治疗方案。耶鲁大学、哈佛大学等机构于 2015 年发表

〔修回日期〕 2016-01-04

〔作者简介〕 弓孟春, 博士, 临床顾问。

的联合研究<sup>[2]</sup>中, 研究人员对哮喘患者的痰液进行基因芯片的检测, 寻找对临床表型进行鉴别的依据。这一结果明确地显示, 不同的临床表型集群 (Cluster 1, 2, 3) 之间, 其基因指纹存在显著的差异, 见图 1。对吸入糖皮质激素的反应存在明显的差异。在哮喘的精准医学治疗模式中, 组学数据必将成为重要的决定因素。而从图中可看出, 其中大量基因位点的分析及计算必然需要依靠 IT 技术才能实现。依据目前 IT 技术发展的趋势, 可以对哮喘的精准医学诊疗模式做进一步的设想。结合患者的人

口学数据、诊断、基线肺功能评估结果、既往用药、基因组分析及痰液转录组分析制定初步方案; 利用可穿戴设备 (便携式峰流速仪), 收集患者每日的峰流速, 结合当日用药剂量及种类、环境中花粉监测数据、PM2.5 数据、流感病毒流行数据等, 使用人工神经网络构建机器学习模型, 逐步修正参数, 最终优选出最重要的指标及参数, 实现自动计算当日用药的功能, 目标是最大程度地控制急性哮喘发作, 减少急诊入院并在长期改善患者心肺功能。

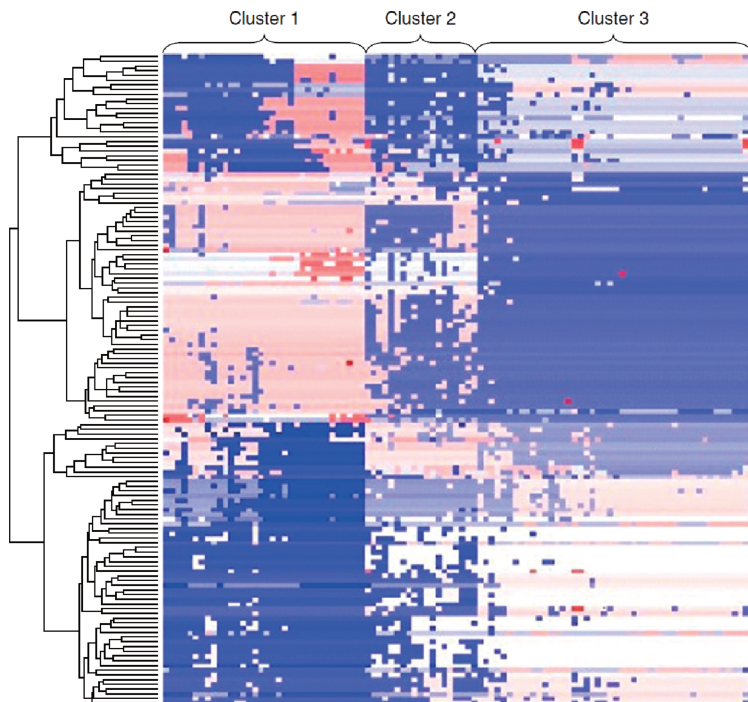


图 1 使用 MCLUST 根据京都基因与基因组百科全书通路进行的集群分析结果

## 2 实现精准医学的核心步骤: 临床决策支持系统 (CDSS)

### 2.1 基因组数据和临床数据

在一项针对新生儿糖尿病患者进行的为期 14 年 (2000 年 1 月至 2013 年 7 月) 的研究中<sup>[3]</sup>, 来自 79 个国家的 1 020 例新生儿糖尿病患儿的血液样本被转送位于英国埃克塞特的中心进行致病基因的测序。从最初的费用昂贵且耗时较长的 Sanger 法到

2012 年起开始使用成本和时间均大幅降低的二代测序技术, 研究组克服了种种方法学、经费等困难, 为儿科及内分泌科的临床医生绘制了由大量的“伞”和“篮”组成的图谱 (“伞篮图”), 见图 2。雨伞的顶端是左侧列举的 22 个与新生儿糖尿病发病相关的基因。篮子的底部是右侧 6 个对临床诊疗和预后判断具有显著影响的临床表型。一旦出现临床表现, 应立即采用基因测序的技术明确致病基因并制定相应的临床诊疗方案。未来, 大量的疾病诊治将遵循这样的基于海量基因组数据和临床数据的规范。

## 2.2 建立基因型-表型对应关系构建体系

现代医学的发展已经进入知识及数据呈指数级增长的时代<sup>[4]</sup>。在传统的临床数据之上，根据基因组、微生物组、环境暴露、生活习惯等不同维度的信息，对患者在人群中进行精准定位，制定精准的疾病预防、治疗方案并给予精准的疾病预后评估。而这些不同纬度的度量值之间的各种形式的组合所对应的不同的临床方案，就如上文提到的伞篮图一样，可能构建出极为复杂的对应关系。单纯依靠医生的记忆和手工文献检索是不可能应付所有病种的不断更新的知识体系的。精准医学的核心要务之一是建立基因型-表型对应关系的构建体系并以系统的形式呈现给医生，辅助临床决策。

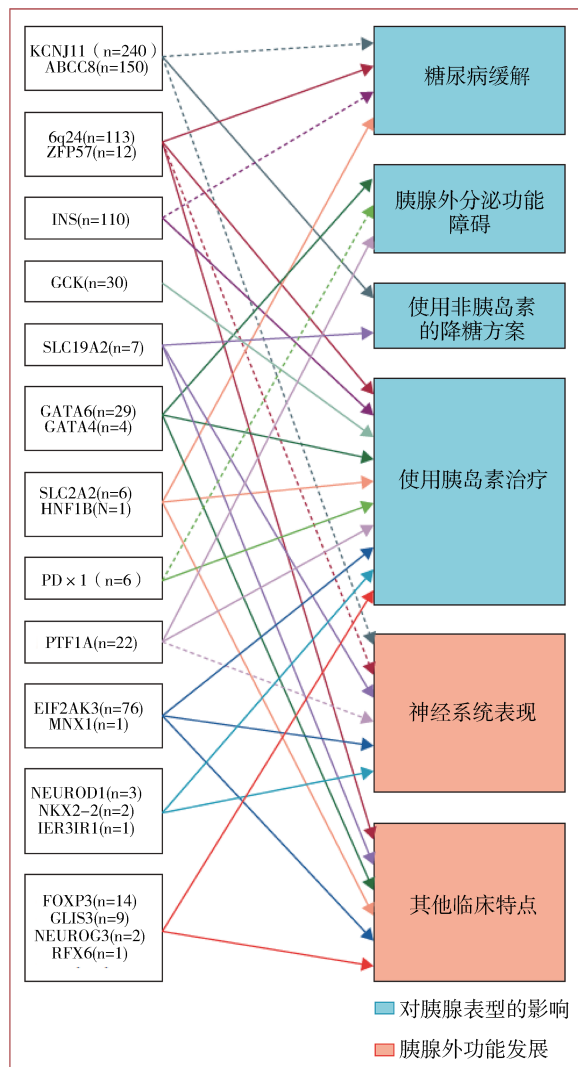


图2 新生儿糖尿病中基因突变与临床表现的对应关系<sup>[7]</sup>

## 2.3 应用临床决策支持系统十分重要

精准医学需借助 IT 技术的力量为临床服务的发  
生节点提供由计算机驱动的实时的临床决策支持  
(Clinical Decision Support System, CDSS)<sup>[5]</sup>。在数据  
量指数级增长、精准医学相关知识不断涌现的时  
代，即使是某一个很窄的学术领域的专家，也不可  
能掌握所有常见的及罕见的与治疗亚组相关的分子  
生物学类型。这些知识的迅速及时的传递需要借助  
于自动化的决策支持技术，从而将精准医学嵌入临  
床诊疗流程中。然而，市面上常见的电子健康档案  
系统呈现及计算组学数据 (Omics Study Data) 的能  
力目前极为有限<sup>[6]</sup>。相反，由基因测序公司提供的  
基因组量级的数据与目前的临床电子健康档案系统  
提供的高度相关性的数据之间缺乏有效的整合。为  
精准医学而设计的自动化的临床决策支持系统的创  
新性解决方案目前仍有广阔的市场。

## 3 医学信息学在精准医学的发展中发挥重要的作用

### 3.1 引导研究方向

近期在《新英格兰医学杂志》发表的“在转移  
性膀胱癌中 DNA 修复缺陷与 Olaparib”<sup>[7]</sup>、在《柳  
叶刀》杂志发表的“详尽的早期基因组检测对新生  
儿糖尿病临床诊疗的影响：国际队列研究”<sup>[3]</sup>以及  
在《美国呼吸及重症医学杂志》上发表的“痰液转  
录组的无创检测鉴别哮喘的临床表型”<sup>[2]</sup>等一系  
列文章将精准医学的临床应用研究推到了学术界的  
焦点中。而结合基因编辑技术在基础医学研究方面  
突飞猛进的发展，学术界及公众对于精准医学的  
认识水平也在逐步提升。这是医学信息学在精准  
医学的发展中可以也必须发挥的重要作用，即引导  
基础医学研究成果向临床应用的转化，推进精准  
医学的科研及临床实践。

### 3.2 保证知识的准确性及可重复性

精准医学的精华在于从无穷的数据收集整合的  
基础上，利用现有的研究结果和医学知识对数据进

行解读,对医生提供疾病诊断、治疗和预防的决策支持。体系化的知识是决定精准医学临床实践成果的决定性要素,需要持续地评估从数据源或知识源收集的信息的准确性和可重复性。例如,目前在全基因组序列信息中对于插入和缺失进行分析识别,因为所用技术不同得到的结果大相径庭<sup>[9]</sup>。需要在方法学上不断取得进步,以解决这一问题。在众多杂志出版方在推动下,权威杂志发布的同行评议论文中的知识能够得到重新评价,以确保学术知识的准确性和可重复性。

### 3.3 改善知识传递的时效性

相比于传统的医学模式,精准医学的知识传递对时效性的要求更高。例如,每年都有新的与基因突变的致病机理相关的数据发布。对于这些数据解读结果的重新计算或重新评价很快,因为基于信息技术的知识传递远快于传统模式。一个可信赖的致力于此的第 3 方机构对此类数据进行反复且频繁的再评价是至关重要的。传统的医学信息提供商,如威科集团、爱思唯尔等,对于知识库的维护已有成熟的技术、管理及商业机制。但为精准医学构建的知识库需要考虑到知识与医疗信息系统的整合、知识所涵盖的信息的维度及数据量等。在知识的更新方面,与卫生信息技术(Health Information Technology, HIT)深度结合的精准医学知识更新会更为快捷。某个基因突变的临床意义信息一旦在知识库内发布或更新,精准医学的信息平台即通过 CDSS 呈现于医疗端并对诊疗流程自动施加干预,从而大大增加知识传递的时效性。

### 3.4 构建知识审查复核体系

未来精准医学对于数据和知识的需求会越来越高,需要设立成熟的评估机构来持续评估数据的精准度。在我国,需要由国家支持的科研或临床机构完成这项工作并纳入我国精准医学的整体规划中。而在技术层面上,是使用半自动化的语义分析技术(如 IBM Watson)还是使用基于专家群体的人工审核机制,学术界尚在讨论中。解决这个问题也需要在精准医学的实践中逐步积累经验。

## 4 我国推进精准医学发展过程中医学信息学的关键任务

### 4.1 构建国家水平的精准医学知识库

目前在文献中已经大量存在使用组学数据指导临床诊疗的证据,可以借助 CDSS 在医生诊治相关患者时提供决策支持。然而,这样的知识库需要完善的疾病列表、高水平的编写团队、定期的知识更新及维护等。鉴于精准医学与 HIT 紧密结合的特点,编写的逻辑、格式、审阅及更新流程等都需要与 HIT 技术的特点相融合,如支持通用的数据传输格式、与逻辑引擎设置相匹配等等。目前我国缺乏基本的精准医学相关文献体系,而其中的知识筛选、整理、发布也只是出于远景规划阶段。这对医疗信息化领域的科研及产业界既是巨大的挑战,也是前景广阔的机会。与传统医学知识库体系相比,精准医学知识库具有明显的特点,必须在方法论、技术等层面进行大量的探索,而且要与传统的电子病历、医院信息系统、区域医疗信息共享平台、生物信息学数据处理平台等进行深度的结合,才能满足精准医学对于知识的需求。另一方面,一旦构建起这样内容全面、审核机制完善、及时高效、与各个医疗信息化技术平台无缝对接的精准医学知识库,则可以控制精准医学的临床服务质量、促进医生的继续教育、为 IT 产业界的解决方案创新提供平台及知识支撑,从而有效地推进精准医学的发展,造福广大患者。

### 4.2 在精准医学相关政策法规、伦理学等非医学议题方面加以引导

基于既往的经验,如果在临床诊疗及科学研究之间划分不清,就容易造成方法学及伦理学问题。精准医学借助于最新的信息技术来获取及共享个人相关信息。而精准医学的理念,又要求这样的数据范围广、时间跨度大且精度高。在这种前提下,如果对大范围数据分享及数据整合导致的隐私暴露风险重视不足,可能引发诸如“基因歧视”等严重后果<sup>[10]</sup>。因而精准医学规划必须涵盖相关的政策法规

的制定,以明确在特殊的社会文化背景下,信息的采集、共享、深入分析等可能带来的大范围的社会影响及应对措施。美国的精准医学计划中包含了大量的关于政策、法规的构想及讨论,我国科技部的精准医学研究支持方案中,也对政策、法规、伦理等方面安排了专门的研究经费支持。此外,精准医学面临的伦理学问题有别于传统医学。尽管临床外显子组测序及全基因组测序是明确遗传疾病细节的符合成本效益分析的方案,但通常这些方法都是无靶点的过程,同时产生大量的医学及社会影响不明的数据。而解析技术不成熟、分析准确性判定通用标准的缺失、不同的检测平台、不完整的或能力不足的参比文库等都可能带来新的伦理难题。医学信息学领域需要对上述议题加以重视及引导,构建代表性的研讨机制及平台,将患者权益组织、医生行业协会、国家行政部门、法律及伦理学研究专家或机构等各方纳入讨论体系并以专题、研讨会、专刊、知识库、继续教育材料等多种形式推动各领域对上述议题的重视,为构建健康、可持续发展的精准医学生态环境发挥积极作用。

#### 4.3 医生及患者教育

精准医学对于医生和患者而言都是全新的事物。尽管“个体化治疗”与“精准医学”在核心理念上没有本质区别,然而二者在信息量级上的显著差别导致医生无法依靠传统的学习和工作方式来应对从各个渠道涌现出来的海量信息。精准医学可以借助 CDSS 及知识库等技术最大程度地辅助医生制定决策。然而,对于医疗工作者而言,也需要紧跟时代潮流,在量化分析、计算机辅助技术、新知识获取等各个方面进行能力的培养<sup>[11]</sup>。这对于当代的医学教育提出了新的要求,也是医学信息学的主要方向之一。而对于已经离开医学院的医生们,则需要继续教育中逐步传递精准医学的知识。反之,精准医学所使用的技术本身也可以对医生继续教育提供全新的模式。在循证医学高度发展的今天,各类临床指南在为医学标准化做出巨大贡献的同时,也为临床工作带来了诸多困扰。不断涌现的新的循证医学证据对于医生的业务更新水平以及医院的信

息系统提出了更高的要求。临床决策支持系统赋予医疗机构根据最新、最权威的临床诊疗指南制定相应规则的能力,从而保证医生可以在对特定的符合条件的患者进行诊治时,通过各种手段得到相关的临床指南信息并做出规范的诊疗决策。这是全新的知识学习的模式,会对医生继续教育这项工作产生深刻的影响。

#### 4.4 加强与 HIT 产业界的合作

医学信息学发展的趋势之一是与新技术的融合,其中重要的环节就是与医疗信息化技术融合,从而将医学信息学体系中获取的知识实时有效嵌入到临床诊疗流程中,对医生的临床决策进行支持。如 UptoDate 与电子病历的整合、威科集团提供的与临床情况相关的医嘱指导等都是这类合作的丰硕成果。在精准医学方面,这样的合作更为重要。最主要的原因是精准医学本身包含的信息量、知识量已经远远超出医生个体所能处理的能力范围,医学信息学提供的知识无法以传统的学习模式加以传递。而另一方面,电子病历与医院信息系统已成为医生工作必需的工具和环境,脱离医疗信息化技术谈精准医学及其知识体系是空中楼阁。医学信息学需要加快与医疗信息化技术的融合,在数据传输标准、临床应用实例、知识价值量化等方面增加投入并将其用于精准医学的知识体系构建。

### 5 结语

医学信息学处在精准医学发展的关键节点上。精准医学知识体系的初步构建依赖于目前已有的循证医学研究结果。然而,目前的研究体系中所涵盖的患者是极小一部分代表人群。97% 以上的肿瘤患者都未接受过任何临床研究<sup>[12]</sup>。如何将“真实世界”证据 (Real World Evidence) 纳入医学体系是未来医学信息学发展的重要方向。精准医学借助 CDSS 在规则构建的过程中引入人工神经网络,从而可以借助大数据技术处理社交媒体信息、病毒流行趋势、空气质量等,也可借助机器学习逐步优化

预测模型的指标及参数,构建针对单个个体的疾病管理模型。大数据技术及人工智能/机器学习等技术与精准医学的深度融合,必将为医学信息学提出更多的挑战。现有的循证医学评估体系可能发生重大的调整。如何在智能化精准医学的时代构建知识的审核、评价及呈递体系,是医学信息学在未来 10 年内要解决的核心问题。

## 参考文献

- 1 Collins FS, Varmus H. A New Initiative on Precision Medicine [J]. The New England Journal of Medicine, 2015, (372): 793–795.
- 2 Yan X, Chu JH, Gomez J, et al. Noninvasive Analysis of the Sputum Transcriptome Discriminates Clinical Phenotypes of Asthma [J]. American Journal of Respiratory and Critical Care Medicine, 2015, (191): 1116–1125.
- 3 De Franco E, Flanagan SE, Houghton JA, et al. The Effect of Early, Comprehensive Genomic Testing on Clinical Care in Neonatal Diabetes: an international cohort study [J]. Lancet, 2015, (386): 957–963.
- 4 Hawgood S, Hook – Barnard IG, O'Brien TC, et al. Precision Medicine: beyond the inflection point [J]. Science Translational Medicine, 2015, (7): 300, 317.
- 5 Masys DR, Jarvik GP, Abernethy NF, et al. Technical Desiderata for the Integration of Genomic Data into Electronic Health Records [J]. Journal of Biomedical Informatics, 2012, (45): 419–422.
- 6 Kohane IS, Isaac S. Health Care Policy. Ten Things We Have to Do to Achieve Precision Medicine [J]. Science, 2015, (349): 37–38.
- 7 De Franco E, Flanagan SE, Houghton JA, et al. The Effect of Early, Comprehensive Genomic Testing on Clinical Care in Neonatal Diabetes: an international cohort study [J]. Lancet, 2015, (386): 957–963.
- 8 Mateo J, Carreira S, Sandhu S, et al. DNA – Repair Defects and Olaparib in Metastatic Prostate Cancer [J]. The New England Journal of Medicine, 2015, (373): 1697–1708.
- 9 O'Rawe J, Jiang T, Sun G, et al. Low Concordance of Multiple Variant – calling Pipelines: practical implications for exome and genome sequencing [J]. Genome Medicine, 2013, (5): 28.
- 10 Green RC, Lautenbach D, McGuire AL. GINA, Genetic Discrimination, and Genomic Medicine [J]. The New England Journal of Medicine, 2015, (372): 397–399.
- 11 Manrai AK, Bhatia G, Strymish J, et al. Medicine's Uncomfortable Relationship with Math: calculating positive predictive value [J]. JAMA Internal Medicine, 2014, (174): 991–993.
- 12 American Society of Clinical Oncology. The State of Cancer Care in America, 2015: a report by the American society of clinical oncology [J]. Journal of Oncology Practice / American Society of Clinical Oncology, 2015, (11): 79–113.

## 2016 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的医学信息学方面的国家级期刊。主管:国家卫生和计划生育委员会;主办:中国医学科学院;承办:中国医学科学院医学信息研究所。中国科技核心期刊(中国科技论文统计源期刊),RCCSE 中国核心学术期刊(武汉大学中国科学评价研究中心, Research Center for Chinese Science Evaluation),美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引(WPRIM)收录,并收录于国内 3 大数据库。主要栏目:专论,医学信息技术,医学信息研究,医学信息组织与利用,医学信息教育,动态等。读者对象:医学信息领域专家学者、管理者、实践者,高等院校相关专业的师生及广大医教研人员。

2016 年《医学信息学杂志》国内外公开发行,每册定价:15 元(月刊),全年 180 元。邮发代号:2–664,全国各地邮局均可订阅。也可到编辑部订购:北京市朝阳区雅宝路 3 号(100020)医科院信息所《医学信息学杂志》编辑部;电话:010–52328673, 52328674, 52328671。

《医学信息学杂志》编辑部