

医学信息学的新机遇和挑战：从参与性医学到智能健康管理^{*}

白晋伟

沈百荣

(四川大学华西医院科技部图书信息中心 成都 610041) (四川大学华西医院系统遗传研究院 成都 610041)

〔摘要〕 从数据、技术、算法和应用 4 个方面探讨现代医学信息学所面临的新机遇与挑战。从学科生态角度分析现代医学信息学与本体构建、参与性医学、数据科学、人工智能、互联网医学等的关系并对其新的研究方向进行展望。

〔关键词〕 医学信息学；本体；参与性医学；智能健康管理

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2019.04.001

New Opportunities and Challenges for Medical Informatics: from participatory medicine to smart healthcare management

BAI Jinwei, Library Information Center of Science and Technology Department, West - China Hospital, Sichuan University, Chengdu 610041, China; SHEN Bairong, Institutes for Systems Genetics, West - China Hospital, Sichuan University, Chengdu 610041, China

〔Abstract〕 The paper discusses new opportunities and challenges confronting modern medical informatics in four aspects of data, technologies, algorithms and application. From the perspective of discipline system, it analyzes relationships between modern medical informatics and ontology building, participatory medicine, data science, Artificial Intelligence (AI) and Internet - based medicine. At last, it forecasts the new study orientation in this field.

〔Keywords〕 medical informatics; ontology; participatory medicine; smart healthcare management

1 引言

简单而言，医学信息学是关于医学数据的存取、分析和应用、医学与信息学交叉的学科，随着时代的演变，医学数据本身以及存取、分析、应用等各方面都发生巨大变化，如从原来的小数据演变

成现代的大数据；从群体病历发展到包含个人基因组信息、智能穿戴设备测定的人体生理等多维数据；从静态演变成复杂多样、动态的时空数据；从单一医院数据演变成患者及相关行业人员参与的医学数据；从以疾病信息为主演变成包含健康数据在内的人体全息数据等。在数据存取方面有互联网和云计算，在分析方面有可视化、深度学习、知识库和知识图谱等新方法，在应用方面则包括治病到治将病方向的疾病早期预警、治未病方面的健康管理干预等。在此背景下的现代医学信息学迎来前所未有的应用和机遇，其应用实例，见表 1。本文将从数据、技术、算法和应用 4 个方面讨论现代医学信息学的机会和挑战。

〔收稿日期〕 2019-04-15

〔作者简介〕 白晋伟，馆员，发表论文 8 篇；通讯作者：沈百荣，教授，博士生导师，发表论文 120 篇。

〔基金项目〕 国家重点研发计划重点专项“帕金森病生物数据整合和分析”（项目编号：2016YFC1306605）。

表 1 现代医学信息学新应用举例

应用类型	具体应用和机遇	挑战	参考文献
治病	辅助精准诊疗	精准基因型 - 表型数据，尤其是临床深度表型的收集、疾病数据隐私问题等	[1]
	疾病分子机制	个性化的基因 - 环境 - 生活习惯的相互作用，构成个性化的疾病分子机理和治疗需求	[2]
治将病	疾病演化规律	疾病与免疫系统的平衡制约；疾病的自愈现象与过度诊断、过度治疗	[3]
	早期疾病预警	早期信号的获取；早期疾病症状、个性化大数据建模、预警系统及其急性病的救治等	[4]
治未病	疾病风险预测	遗传和表观遗传的动态的变化和高危人群的寻找；复杂疾病风险预测的系统遗传学	[5]
	智能健康管理	健康生活习惯与疾病的个性化关系；个性化的健康管理和预防方案。	[6]

2 现代医学信息学面临的机遇与挑战

2.1 数据：从群体数据到个性化数据

过去 10 年来医学信息学所分析的数据最关键的演变是从群体数据到个性化数据，无论是精准医学还是精准健康管理，首先是要有个性化的数据作为前提并且是多维信息的个性化数据。目前在基因层面，随着测序技术的发展和降价，个人基因组测定的普及将逐渐成为可能，然而与个人基因组相对应的个性化表型数据的缺乏依然是基因型 - 表型关系研究的瓶颈。收集精准的个性化表型，尤其是临床相关的表型数据，建立相关的收集标准，即本体，是多维复杂数据收集的核心问题。目前较高层次的本体如疾病本体（Disease Ontology，DO）已经开发，但是针对具体疾病的个性化、特有的疾病本体仍然有限，亟需对多种疾病进行本体开发研究，才能有利于收集个性化和结构化的数据，以便进一步个性化建模分析。只有收集大量的基于本体的数据才有可能共享、进行个性化的建模分析（包括深度学习和人工智能的实现）。本体研究的根本挑战在于临床医生、医学信息学家和计算机专家在交叉融合基础上建设可以为临床应用的疾病术语及关联结构，构建医学数据收集平台，用于收集结构化、多维、成对的个性化大数据。疾病本体、人工智能及应用的关系，见图 1。本体研究基于临床、医学信息学和计算科学的合作，在本体支持下可以构建参与性网络平台，患者、医生和相关人员协同提供数据形成结构化的生物医学大数据，从而促进人工

智能的发展、知识发现和基于患者相似性的大数据比对应应用，这是实现医学人工智能的基础和必要条件。

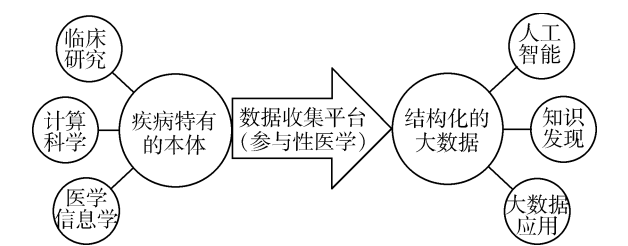


图 1 疾病本体、人工智能及应用的关系

2.2 技术：从互联网到社会网络

互联网的发展为解决远程数据空间孤立与分散的问题提供有利条件，本体建设可以为领域内提供一个共享的标准。与医学信息学相关联的不仅是互联网的发展，社会网络如 Facebook、Linkedin、We-Chat 使人们的生活通过网络虚拟地连接起来，患者、医生、亲友和相关人员社会网络的建立使优质、深度信息与资源的获取变得更加便利，如医生、治疗和药物信息、患者经验等。当然这种社会网络的信息传递也可能带来负面的心理干扰。互联网上可以通过交流建立起患者 - 医生 - 社会网络的信息交换平台，典型的平台如 PatientsLikeme 是患者或相关人员在网络上注册、交流并积累数据与信息分析的平台。基于本体的疾病信息平台可以为患者入院前后进行有指导的信息添加、分析与诊断，从而减少患者与医生的交流障碍与时间，同时有利于数据和时空信息的积累。参与性医学是在互联网和社会网络的技术下实现的一种以患者为主体的新

型医学模式，在互联网的支持下患者、医生和其他相关人员（如家人、好友等）形成的社会网络中患者、医生和相关的医学信息、资源等网络距离变得很近，从而让更多人参与到医学的实践、研究和传播，见图 2。

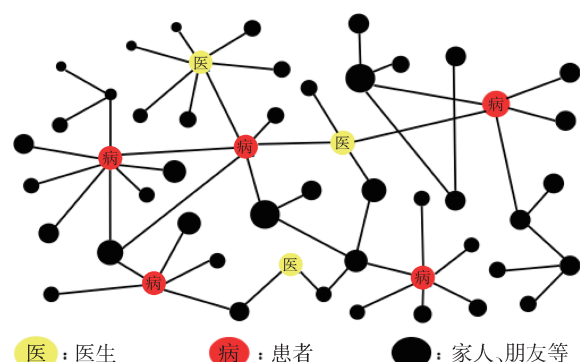


图 2 互联网与社会网络促进参与性医学的发展

2.3 算法：从统计分析到机理、融合分析

传统的个人医学信息有限、数据量小，数据分析方法通常是统计平均为主的群体分析。随着个性化数据量的增长，传统的方法往往不再有效，因为会将数据中的个性化模式平均掉。近年来提出不少对统计平均方法进行修正的算法，如 COPA、MOST 等都是在计算时更加注重隐含在特殊子群结构里的个性化统计算法^[7]。除统计方法外基于第一性原理的方法寻找医学大数据内在的基本规则也十分重要^[8]。大数据的积累为医学人工智能的发展提供机会，但基本规律的寻找和机制性的研究是人工智能所不能替代的。由于数据的多维性和复杂性，不同程度的数据融合是提高分析精准度的根本思路。最近发展起来的放射基因组学就是融合图像与组学两个层次的数据分析策略。随着智能穿戴设备的广泛使用，实时、动态、个性化数据的获取变得越来越便捷和重要，对实时动态信号的收集、处理也需要发展新的算法，如非线性信号变换方法：小波分析、各种熵的计算等。

2.4 应用：从治病到治未病

疾病生物医学数据反映人体疾病的演变特征和规律，人们可以根据其基本特征如动态演化性、系

统网络关系、文本数据、影像数据等开发相应的算法工具，理解疾病的复杂演化规律，从而提前预测疾病的发生发展。早在公元前 250 年，黄帝内经中就提出治病、治将病、治未病的概念，后两者是疾病早期诊断与预防的思想，与疾病演化密切相关，涉及基本的遗传风险预测、病变生理信号、分子变异早期发现与监控，从锻炼、营养和精神等多方面的网络系统演化。而这一复杂系统还有自组织、疾病自愈的特征，免疫系统、神经系统与疾病体系的复杂抗衡演变，即使有早期的疾病信号或遗传风险，在复杂的环境 - 基因 - 生活习惯的多系统耦合下疾病有可能自愈。过度诊断、治疗成为不可避免的现象，任何一个模型都难以精准预测这种系统的动态演变和自愈。目前治病模式更多的还是在生物医学模型下展开的研究或治疗，生物 - 心理 - 社会医学模型需要更大框架下的多维数据才能真正研究从治病到治未病的演变，促进智能健康管理 with 主动健康的实现。

3 现代医学信息学学科生态与新研究方向

由于现代医学信息学在内涵上发生巨大变化，该学科的发展与其他学科如生物信息学、数据科学、人工智能、健康管理、互联网支撑的参与性医学等形成一个交叉学科群。未来学科设置必须考虑医学数据的特征与学科群的协同发展。真正的学科交叉并非简单的加和，而是整合前提下的融合，传统的教学内容应随之发生变化，反映学科发展现状与未来趋势。医学信息学学科相关新研究方向，见表 2。根据医学数据 5 大特征，即演化性、隐私性、复杂性、多维性（文本、数字、图像、动态等）、跨层次性（分子、细胞、个体和群体等），随着数据技术的应用推广，将产生更多新的方向。现代医学信息学是一个跨生物、医学、计算的交叉学科，其发展制约着相关理论和应用科学的发展。从数据到算法、应用对应着不同新型学科的产生与交叉。参与性医学可以促进复杂数据的群体协作功能，而人工智能相关算法将促进精准的智能健康管理实施。

表 2 现代医学信息学新研究方向

新研究方向	研究内容	参考文献
参与性医学	患者和医疗专业人员作为健康的完全合作伙伴，积极合作和鼓励彼此的一种医学范式	[9 – 10]
特有疾病本体	描述某个疾病的词汇和这些词汇之间关系的标准化、计算机可读的形式化术语和关系的集合	[11 – 12]
达尔文医学	进化医学或达尔文医学是现代进化理论在认识健康和疾病方面的应用。目标是理解为什么人们会生病，而不仅仅是如何生病	[3, 13]
医学人工智能	在疾病个性化、精准化诊断和治疗中的应用	[14]
互联网医学	互联网技术在现代医学中的应用并改变传统的诊疗方式，包括网络远程诊疗、资源共享、参与性医学等	[15]
数据共享与隐私	医学数据的标准化、去隐私性政策和数据使用的隐私技术，如联盟数据分析模型等	[16 – 17]

4 结语

新技术催生新的研究范式，现代基因组学是驱动医学信息学变化的根本动力，改变各层次（分子、细胞、个体、群体等）的研究范式以及这些不同层次交叉融合。这种融合也深化医学信息学的内涵，医学图像分析与基因组学融合产生放射基因组学，患者个体的电子病历数据与基因组学融合产生个体化治疗或精准诊疗，群体层次的卫生统计学也将受到基因组学发展的影响而研究群体基因图谱特征和演化等。对应于个体基因组学，个性化表型组学的研究将是医学信息学的重要研究方向，医学信息学的发展也将随着相关学科的生态变化而发生演变。现代医学信息学必将成为新时代重要和关键的交叉学科领域，也将会因时代的发展和需求迎来更多的机会和挑战。

参考文献

1 Chen J, Lin Y, Shen B. Informatics for Precision Medicine and Healthcare [J]. Adv Exp Med Biol, 2017 (1005): 1 – 20.

2 Lin Y, Chen J, Shen B. Interactions Between Genetics, Lifestyle, and Environmental Factors for Healthcare [J]. Adv Exp Med Biol, 2017 (1005): 167 – 91.

3 Stearns SC. Frontiers in Molecular Evolutionary Medicine [EB/OL]. [2019 – 04 – 01]. <https://link.springer.com/article/10.1007%2Fs00239-019-09893-5#citeas>.

4 Bai J, Shen L, Sun H, et al. Physiological Informatics: collection and analyses of data from wearable sensors and smart-phone for healthcare [J]. Adv Exp Med Biol, 2017 (1028): 17 – 37.

5 Jiang J, Cui W, Vongsangnak W, et al. Post Genome –

wide Association Studies Functional Characterization of Prostate Cancer Risk Loci [J]. BMC Genomics, 2013, 14 (Suppl8): S9.

6 Shen L, Ye B, Sun H, et al. Systems Health: a transition from disease management toward health promotion [J]. Adv Exp Med Biol, 2017 (1028): 149 – 164.

7 Tang Y, Yan W, Chen J, et al. Identification of Novel MicroRNA Regulatory Pathways Associated with Heterogeneous Prostate Cancer [J]. BMC Syst Biol, 2013, 7 (Suppl3): S6.

8 白晋伟, 沈百荣. 第二代互联网医学: 从数据到价值 [J]. 中华医学情报学杂志, 2018, 27 (11): 1 – 5.

9 Prainsack B. The Powers of Participatory Medicine [J]. PLoS Biol, 2014 (12): e1001837.

10 Hood L, Auffray C. Participatory medicine: a driving force for revolutionizing healthcare [J]. Genome Med, 2013 (5): 110.

11 Haendel MA, Chute CG, Robinson PN. Classification, Ontology, and Precision Medicine [J]. N Engl J Med, 2018 (379): 1452 – 1462.

12 Schriml LM, Mitraka E, Munro J, et al. Human Disease Ontology 2018 Update: classification, content and workflow expansion [J]. Nucleic Acids Res, 2019 (47): 955 – 962.

13 Nunn CL. A Roadmap for ‘Core Concepts’ in Evolutionary Medicine [J]. Evol Med Public Health, 2018 (2018): 24 – 25.

14 Yu KH, Kohane IS. Framing the Challenges of Artificial Intelligence in Medicine [J]. BMJ Qual Saf, 2019 (28): 238 – 241.

15 Vann Jones S, Breakey W. Informed Dissent: can we handle internet medicine [J]. Evid Based Med, 2015 (20): 190.

16 Raisaro JL, Tramer F, Ji Z, et al. Addressing Beacon Re – identification Attacks: quantification and mitigation of privacy risks [J]. J Am Med Inform Assoc, 2017 (24): 799 – 805.

17 Wang S, Jiang X, Tang H, et al. A Community Effort to Protect Genomic Data Sharing, Collaboration and Outsourcing [J]. NPJ Genom Med, 2017 (2): 33.