

普通高等教育“十三五”规划教材
全国高等院校规划教材

医学信息学

● 主编 赵越



清华大学出版社

普通高等教育“十三五”规划教材
全国高等院校规划教材

医学信息学

● 主编 赵越



清华大学出版社
北京

内 容 简 介

本教材结合了大量的实例,全面、系统地讲解了医学信息学的理论知识及其在医学领域的应用,主要内容包括生物医学信息学概论、医学信息标准化、数据库与信息系统、医院信息系统、医学图像信息系统、实验室信息系统、护理信息系统、社区卫生信息系统、区域卫生信息平台、电子病历、医学数据获取与医学信号处理、临床决策支持系统、远程医疗等。本教材力求理论与实例结合,深入浅出,循序渐进,使读者全面了解医学信息学技术,提高分析问题、解决问题的能力。

本教材可作为高等院校相关专业本科生和研究生的学习用书,也可作为广大科研人员、学者的参考用书。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

医学信息学/赵越主编. --北京:清华大学出版社,2016
(普通高等教育“十三五”规划教材·全国高等院校规划教材)
ISBN 978-7-302-37975-1

I. ①医… II. ①赵… III. ①医学信息—高等学校—教材 IV. ①R-0

中国版本图书馆CIP数据核字(2014)第209488号

责任编辑:罗 健
封面设计:戴国印
责任校对:赵丽敏
责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62770175-4134

印 装 者:北京密云胶印厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:18.25

字 数:493千字

版 次:2016年3月第1版

印 次:2016年3月第1次印刷

印 数:1~2500

定 价:39.80元

产品编号:050560-01

《医学信息学》编委会

主 编 赵 越
副主编 王之琼 徐礼胜
编 委 (以姓氏笔画为序)
王之琼 李 岭 李建华
张 陈 张静淑 赵 越
信俊昌 徐礼胜 崔笑宇
秘 书 (以姓氏笔画为序)
丁琳琳 曲璐渲

PREFACE 前言

医学信息学是一门基于信息科学、计算机科学、医学、生物学、管理学、工程学及其他应用科学的理念和方法的结合,以研究医学信息的搜集、处理、存储、检索、传输、管理和有效应用的原理与方法为基本内容的新型交叉学科。随着计算机技术的发展,数字化医疗已成为一种必然的发展趋势。云计算和大数据等新型数据库技术手段的出现,为处理医学海量数据提供了更加有效的方法,加快了数字化医疗的进程。

近年来,医学信息学在医疗研究、医学实践、医学教育、医学决策和管理方面起着越来越重要的作用,广泛应用于各种医学信息系统,如临床医疗系统、图像处理系统、公共卫生系统和决策支持系统等。目前,电子病历、电子健康档案、医学决策支持系统、生物信号采集与识别系统、医学影像信息系统等对各种医学信息资源进行整合和利用,为提高医疗卫生水平和诊断效率做出了重要贡献。医学信息学的广泛应用,在很大程度上改变了人们传统的医疗观念,还从来没有一门学科像信息科学这样广泛和深刻地影响和改变着医学的整体面貌。

这几年,我国一直在加大医疗产业的投入,努力推动数字化医疗的进程,提高全民的健康和医疗水平。随着信息新技术的发展,医学信息学领域也发生了一些新变化。国内专门介绍医学信息学的教材还比较少。作者根据医学信息学的发展情况及实际需要编写了本教材。本书内容新颖,系统性较强。书中部分内容来自于作者团队近几年的科研成果,也有一部分内容来自国内外同行的研究成果,书中还添加了大量的实际案例,对理论性较强的部分,作者绘制了大量的框图,有利于读者加深对理论、概念的理解。本书从医疗研究、医学实践、医学教育、医学决策和管理等领域的实际需求出发,突出专业特色,淡化理论推导,强化能力的培养,将医学信息学与专业实践相结合,注重培养学生解决工程中的医学信息学问题的能力。

本教材共分 15 章,详尽地介绍了医学信息学的理论以及相关应用知识:第 1 章介绍生物医学信息学概论。第 2 章主要介绍医学信息标准化及其分类的一些标准。第 3 章主要介绍数据库与信息系统:首先阐述了数据库技术,接着又介绍了基于数据库技术的信息系统。第 4 章主要介绍医院信息系统:首先概述医院信息系统;其次介绍医院信息系统的主要功能与内容;再次介绍医院信息系统的设计与实现;最后介绍医院信息系统的相关案例。第 5 章主要阐述医学图像信息系统,首先介绍医学成像的基本概念和医学影像的主要类型,然后介绍医学影像的 DICOM 标准,在此基础上阐述 PACS 及 RIS 的组成、应用与发展趋势等。第 6 章主要介绍实验室信息系统,先介绍

实验室系统的基本概念, 然后介绍实验室信息系统的主要技术, 在此基础上介绍 HIS 的质量管理体系和应用等。第 7 章主要介绍护理信息系统, 主要阐述护理系统的相关概念以及移动护理信息系统的组成以及设计原理等。第 8 章主要介绍社区卫生信息系统及其结构模型和系统组成等。第 9 章介绍区域卫生信息平台, 主要介绍区域卫生信息平台的总体规划和实现策略等。第 10 章主要介绍电子病历: 首先概述电子病历及其功能与内容; 其次介绍电子病历的设计与实现以及电子病历面临的挑战; 再次介绍中医电子病历; 最后介绍电子病历的一些相关案例等。第 11 章介绍医学数据获取与医学信号处理, 主要介绍生物传感器的基本理论以及医学信号处理的基本方法。第 12 章主要介绍临床决策支持系统及其概念、理论基础、系统结构和特征等。第 13 章主要介绍远程医疗及其概念、组成与功能、组织管理、应用、发展趋势等。第 14 章主要介绍生物信息学的定义、研究内容、生物信息数据库及其在医学领域中的应用等。第 15 章主要展望了医学信息学的未来, 介绍了医学信息学现状和挑战, 并且介绍了目前医学信息学的一些研究热点。

教师可以根据本科生、研究生不同的教学要求灵活选讲本教材内容。另外, 可以根据自身的教学条件, 理论与实验结合, 通过多媒体教学可以使學生更加深刻地理解相关理论与技术。本书理论介绍通俗易懂, 也非常适合非专业人士自学使用。

本书课件可登录清华大学出版社网站 (<http://www.tup.com.cn>) 下载。

本书的出版与作者课题组多年的科研工作密不可分, 再次向一道工作的同事和研究生表示感谢。同时感谢清华大学出版社罗健编辑给予我们的帮助和支持, 对本书所列参考文献的作者也表示感谢。此外, 感谢赵歆卓、张维和魏永涛等同学在本书的编撰过程中给予的帮助。

限于编者的水平, 书中不妥和错误之处在所难免, 敬请广大读者及同行批评指正。

主编

2016年1月

第 1 章 生物医学信息学概论	1	2.4.5 国际社区医疗分类 ICPC	23
1.1 学科发展史	1	2.4.6 Read 临床代码	24
1.1.1 医学信息学发展史	1	2.4.7 检测报告逻辑命名与编码	
1.1.2 生物信息学发展史	2	系统 LOINC	24
1.2 学科的性质和任务	3	2.4.8 医学主题词表 MeSH	25
1.3 学科发展趋势	4	2.5 医学信息交换标准	25
1.4 学生学习本课程要注意的问题和		2.5.1 HL7	25
学习方法介绍	10	2.5.2 DICOM	27
本章小结	11	本章小结	27
课后习题	11	课后习题	28
第 2 章 医学信息标准化	12	第 3 章 数据库与信息系统	29
2.1 标准与标准化	12	3.1 数据库技术	29
2.1.1 标准	12	3.1.1 数据与数据管理	29
2.1.2 标准化	12	3.1.2 数据管理技术发展历程	30
2.1.3 标准化发展历程	14	3.1.3 数据库的数据模型	31
2.1.4 信息标准化	15	3.1.4 数据库的类型	32
2.2 分类与编码	16	3.1.5 数据库系统与数据库管理	
2.2.1 基本概念	16	系统	34
2.2.2 分类与编码的基本原则	16	3.1.6 数据库应用开发流程	36
2.2.3 分类与编码的方法	16	3.2 信息系统	38
2.3 医学信息标准	18	3.2.1 信息与信息系统	39
2.3.1 建立医学信息标准的艰巨性	18	3.2.2 信息系统的发展历程	41
2.3.2 医学信息标准的类型	18	3.2.3 信息系统的进化过程	41
2.4 医学信息学常用分类标准	19	3.2.4 信息系统的类型	42
2.4.1 国际疾病分类 ICD	19	3.2.5 信息系统的开发方法	42
2.4.2 人类与兽类医学系统术语		本章小结	43
SNOMED	20	课后习题	44
2.4.3 诊断相关组 DRG	22	第 4 章 医院信息系统	45
2.4.4 一体化医学语言系统		4.1 医院信息系统概述	45
UMLS	23	4.1.1 概论	45

4.1.2 相关概念.....	46	第 6 章 实验室信息系统	99
4.2 医院信息系统的功能与内容	48	6.1 实验室信息系统概述	99
4.2.1 构造与系统结构.....	48	6.1.1 定义.....	99
4.2.2 功能.....	50	6.1.2 实验室信息系统的发展历史.....	99
4.2.3 组成.....	51	6.1.3 建立实验室信息系统的目的.....	100
4.3 医院信息系统的设计与实现	52	6.2 实验室信息系统的主要技术	100
4.3.1 模型的总体规划(设计、接口).....	52	6.2.1 模式识别技术.....	101
4.3.2 主要技术.....	57	6.2.2 条形码技术.....	101
4.3.3 实现(实施策略).....	59	6.2.3 自动控制技术.....	103
4.3.4 问题与对策.....	66	6.3 实验室信息系统的功能	104
4.3.5 应用与效益.....	68	6.3.1 LIS的用户权限管理.....	104
4.3.6 管理与发展.....	69	6.3.2 基础数据维护.....	105
4.4 案例	70	6.3.3 标本管理.....	105
本章小结	77	6.3.4 检验管理.....	106
课后习题	77	6.3.5 质量管理.....	107
第 5 章 医学图像信息系统	78	6.3.6 综合查询统计.....	108
5.1 医学图像信息系统的概述	78	6.3.7 实验室管理.....	108
5.1.1 数字图像的基本概念.....	78	6.4 实验室信息系统质量管理	
5.1.2 医学图像成像原理与设备.....	79	体系	108
5.1.3 医学图像的作用.....	83	6.4.1 ISO 15189 和 CAP 对 LIS 的要求.....	108
5.1.4 医学图像信息系统的发展历程.....	84	6.4.2 LIS 促进质量管理体系建立.....	109
5.2 医学图像的管理与显示	84	6.5 实验室基本业务流程	109
5.2.1 医学图像标准.....	85	6.5.1 门、急诊检验流程.....	109
5.2.2 医学图像采集.....	86	6.5.2 住院部检验流程.....	110
5.2.3 医学图像存储.....	86	本章小结	111
5.2.4 医学图像传输.....	90	课后习题	112
5.2.5 医学图像显示.....	90	第 7 章 护理信息系统	113
5.2.6 医学图像处理.....	92	7.1 护理管理学与护理信息学	113
5.3 医学图像信息系统的应用	94	7.1.1 护理学.....	113
5.3.1 医学图像存储与传输系统(PACS).....	94	7.1.2 护理管理学.....	114
5.3.2 放射信息系统.....	96	7.1.3 护理信息学.....	115
5.3.3 HIS-RIS-PACS 整合.....	97	7.2 护理信息系统	117
5.4 案例:沈阳某医院的 PACS 系统	97	7.2.1 护理信息系统的定义.....	117
本章小结	98	7.2.2 护理信息的特点.....	117
课后习题	98	7.2.3 传统护理信息的处理方式.....	118

7.3 移动护理信息系统	118	8.5 案例:北京市新社区卫生医疗信息 系统	139
7.3.1 国内外移动护理信息系统 的应用	118	本章小结	141
7.3.2 移动护理信息系统的组成 ..	119	课后习题	141
7.3.3 移动护士工作站对护理工作 的作用	121	第9章 区域卫生信息平台	142
7.3.4 移动护理信息系统存在的 不足	122	9.1 区域卫生信息平台概论	142
7.4 移动护理信息系统的设计	122	9.1.1 区域卫生信息平台的背景 ..	142
7.4.1 架构设计	123	9.1.2 区域卫生平台的作用	143
7.4.2 系统硬件设计	123	9.1.3 区域卫生平台的相关概念 ..	144
7.4.3 系统功能及操作过程	124	9.2 区域卫生信息平台的总体规划 ..	144
7.5 案例: H-ERP 系统的应用	125	9.2.1 系统体系结构	144
本章小结	127	9.2.2 数据分布体系结构	144
课后习题	127	9.2.3 功能与服务类型	146
第8章 社区卫生信息系统	128	9.3 区域卫生信息平台的设计 与实现	147
8.1 社区卫生信息系统概述	128	9.4 区域卫生信息平台的具体 实例	148
8.1.1 社区卫生信息系统的定义 ..	128	9.4.1 基于健康档案的区域卫生 信息化方案	148
8.1.2 社区卫生信息系统的目标 ..	129	9.4.2 基于社区卫生服务的区域 卫生信息交换平台	150
8.2 社区卫生信息系统的功能与 内容	129	本章小结	151
8.2.1 社区卫生信息系统的结构 模型	130	课后习题	152
8.2.2 社区卫生信息系统的数 据模型	130	第10章 电子病历	153
8.2.3 系统总体组成	130	10.1 电子病历概述	153
8.2.4 社区医疗管理子系统	131	10.1.1 病历的概念	153
8.2.5 社区医院行政管理子系统 ..	131	10.1.2 传统纸质病历的局限性 ..	154
8.2.6 社区卫生服务管理子系统 ..	131	10.1.3 电子病历的优势	155
8.3 社区卫生资源共享系统	134	10.2 电子病历的功能与内容	156
8.4 社区卫生信息系统的设计与 实现	135	10.3 电子病历的设计与实现	157
8.4.1 突发公共卫生事件应急指挥 信息系统概述	135	10.3.1 电子病历录入功能的设计 与实现	157
8.4.2 总体业务/信息流程——指挥 控制闭环系统	136	10.3.2 电子病历显示功能的设计 与实现	158
8.4.3 应用架构	136	10.3.3 电子病历存储与查询功能 的设计与实现	159
8.4.4 应用系统设计	137	10.3.4 电子病历安全管理功能 的设计与实现	161

10.4 电子病历面临的挑战	164	12.1.3 临床决策支持系统的 意义与种类	195
10.5 中医电子病历	165	12.2 临床决策支持系统的系统 结构与主要特征	195
10.5.1 中医病历的发展简史	165	12.2.1 临床决策支持系统的系统 结构	195
10.5.2 中医电子病历的作用	165	12.2.2 临床决策支持系统的主要 特征	196
10.5.3 中医电子病历的研究进展	166	12.3 临床决策支持方法	197
10.6 案例:某医院电子病历系统	166	12.3.1 医学实验结果分类	197
本章小结	167	12.3.2 基于贝叶斯定理的预测	199
课后习题	167	12.3.3 基于期望的预测	200
第11章 医学数据获取与医学 信号处理	168	12.4 临床决策支持系统的设计	202
11.1 生物医学传感器基础	168	12.4.1 患者数据的获取与验证	202
11.1.1 生物医学传感器的定义 与作用	168	12.4.2 医学知识的建模	202
11.1.2 生物医学传感器的分类	170	12.4.3 医学知识的导出	202
11.1.3 生物医学传感器的特殊性与 安全性	170	12.4.4 医学知识的表示和推理	203
11.1.4 生物医学传感器的标定 与校准	172	12.4.5 系统性能的验证	203
11.1.5 新型生物医学传感器及 系统	173	12.5 案例:MYCIN	203
11.2 生物医学信号的数据采集	174	本章小结	204
11.2.1 概述	175	课后习题	204
11.2.2 嵌入式数据采集系统	176	第13章 远程医疗	206
11.2.3 生物医学无线传感器 网络	177	13.1 远程医疗系统的概念	206
11.3 常用生物医学信号处理方法	181	13.1.1 远程医疗的定义	206
11.3.1 时域分析方法	181	13.1.2 远程医疗的内涵	208
11.3.2 频域分析方法	182	13.1.3 远程医疗的目的	208
11.3.3 时频联合分析方法	185	13.1.4 远程医疗的发展历程	208
11.3.4 非线性动力学方法	187	13.2 远程医疗系统的组成与功能	208
本章小结	188	13.2.1 远程医疗系统的组成	209
课后习题	188	13.2.2 远程医疗系统的功能	209
第12章 临床决策支持系统	189	13.3 远程医疗的组织管理	215
12.1 临床决策支持系统概述	189	13.3.1 远程医疗机构管理	215
12.1.1 专家系统与智能决策支持 系统	189	13.3.2 远程医疗专家管理	215
12.1.2 知识的表示与推理	191	13.3.3 远程医疗会诊流程管理	215
		13.3.4 远程医疗的隐私保护	216
		13.4 远程医疗的应用	216
		13.4.1 远程医学诊断	216
		13.4.2 远程医学治疗	217

13.4.3 远程监护	217	第 15 章 医学信息学的未来	230
13.5 远程医疗发展趋势	217	15.1 医学信息学的现状及挑战	230
13.6 案例：辽宁省甲型 H1N1 流感 远程医疗会诊系统	218	15.1.1 医学信息学的研究热点	231
本章小结	218	15.1.2 医学信息学发展面临的 挑战	232
课后习题	218	15.2 医学信息学研究的目标、机遇 与未来	235
第 14 章 生物信息学	219	15.2.1 医学信息学研究的目标	236
14.1 生物信息学概述	219	15.2.2 医学信息学研究面临的机遇 与未来	236
14.1.1 生物信息学的定义	219	本章小结	237
14.1.2 生物信息学的兴起与发展	220	课后习题	238
14.2 生物信息学的研究范畴	221	习题答案	239
14.2.1 生物信息数据库	221	第 1 章 生物医学信息学概论	239
14.2.2 序列比对(sequence alignment)	221	第 2 章 医学信息标准化	239
14.2.3 比较基因组(comparative genomics) 和系统发育 分析	222	第 3 章 数据库与信息系统	240
14.2.4 蛋白质结构预测	222	第 4 章 医院信息系统	242
14.2.5 生物芯片(biochip)	222	第 5 章 医学图像信息系统	244
14.3 生物信息数据库	223	第 6 章 实验室信息系统	245
14.3.1 核酸数据库	223	第 7 章 护理信息系统	246
14.3.2 蛋白质数据库	225	第 8 章 社区卫生信息系统	248
14.4 生物信息学在医学领域的 应用	227	第 9 章 区域卫生信息平台	249
14.4.1 新基因的发现	227	第 10 章 电子病历	250
14.4.2 疾病的诊断与预防	227	第 11 章 医学数据获取与医学信号 处理	252
14.4.3 疾病相关生物信息学数据库 的建立	227	第 12 章 临床决策支持系统	256
14.4.4 药物的研究与开发	228	第 13 章 远程医疗	258
本章小结	228	第 14 章 生物信息学	259
课后习题	228	第 15 章 医学信息学的未来	267
		参考文献	268
		中英文词汇索引	273

第 1 章

生物医学信息学概论

学习目的

1. 熟悉医学信息学与生物信息学的概念，并了解其主要的研究内容。
2. 了解医学信息学的发展趋势及瓶颈。
3. 了解生物信息学的主要研究策略与热点问题。

引言

2013年3月出版的美国《大西洋月刊》以“机器人医师”为主题，介绍了以IBM公司新近推出的超级计算机“沃森（Watson）”为代表的IT技术即将给医疗行业带来的冲击。作者明确指出，多年来信息技术对各个行业均造成了巨大的冲击，然而医疗卫生行业却一直置身事外。而今，硅谷的精英们正在向着这个最后的堡垒发起冲击。在不久的将来，计算机将帮助医师诊断、治疗，而整个卫生体系乃至美国经济也会受到影响。

生物医学信息学（biomedical informatics）具体包括医学信息学（medical informatics）和生物信息学（bioinformatics），二者均涉及信息技术在生物医学研究以及临床医学各个方面的应用。随着计算机、互联网技术以及分子生物学，尤其是人类基因组研究的迅猛发展，生物医学信息学的重要性正日益凸显，其成果不仅为患者与医师之间的沟通提供了极大的便利，而且为医师从大量的文献和实验数据中找出规律和线索提供了强大的工具。

生物医学信息学是信息技术与生物医学深度融合的产物，具有明显的交叉学科性质。对相关专业的学生来说，有必要熟悉这一领域的理论、技术与研究前沿。

1.1 学科发展史

生物医学信息学诞生于20世纪中叶，最初被用于医疗数据的规范化管理。随着信息技术和分子生物学研究的不断突破，其形式和内容均发生了质的飞跃。

1.1.1 医学信息学发展史

医学信息学主要处理健康信息、医疗信息、护理信息、保健信息、实验信息、医学影像信息等，是涉及信息科学、计算机科学和医学的交叉学科，具体研究内容包括健康以及生物医学信息的采集、存储、检索、利用以及设备和方法的开发等，所使用的工具不仅包括计算机、互联网，还包括临床治疗、正规的医疗用语以及信息与通信系统等。

20世纪初，美国即出现了医院标准化的浪潮，最先体现在物资、设备和财务管理等方面。20世纪40年代，医学信息学随着数字计算机的发明而形成学科。1949年，德国的Gustav Wagner



图 1-1 Gustav Wagner(1918—2006)

(图 1-1) 注册了第一家专业的医学信息机构。随后, 法国、比利时、荷兰、波兰和美国亦先后建立了健康信息学的研究和教育机构。随着计算机技术的成熟和迅速普及, 医学信息学在全球迅速发展, 相继衍生出医疗计算、医疗电子数据处理、医疗自动化、临床信息管理、卫生信息管理、医疗软件工程等多个分支。20 世纪 60 年代, 计算机被首次用于解决生物医学问题, 并诞生了医学文献自动化检索系统 (medical literature automatic retrieval system)。20 世纪 70 年代, 开始出现医院信息系统。随着计算机技术的发展, 至 20 世纪 80 年代中期, 医学信息学逐渐从数据处理的阶段进入到以人工智能和专家系统为标志的知识处理阶段。成立于 1987 年的国际医学信息协会 (International Medical Informatics Association, IMIA) 已成为该领域最重要的学术机构。进入 21 世纪后, 医学信息学的发展主要体现在医学知识的表达与处理、临床决策支持、公共健康信息和网络以及医学信息学的教育与培训等方面。

医学信息处理涉及各类健康信息的有效组织、管理、分析和使用。在过去, 存储、集成和检索医疗和患者的信息均采用纸质文件。然而, 基于纸张的信息管理系统效率低下且花费巨大, 不利于患者的护理。具体的问题包括: ① 传统的医疗记录可能难以辨认, 因为它是手写的且内容简陋, 医师很难找到所需的信息, 如已做过的检查项目和结果; ② 一家医疗机构的病历往往不与其他医师、实验室和医院共享。患者的资料变得支离破碎, 这可能导致患者治疗的中断或延误。

在过去的几十年里, 医疗信息以惊人的速度增长。目前, 美国国家医学图书馆数据库一年所收集和维持的生物医学文献的规模即接近 50 万份。医师们可以通过大量的检索来理解临床试验以及其他研究的结果并辅助自己的治疗行为。另一方面, 患者也可以参考最新的证据来预防疾病和接受治疗。有研究表明, 患者了解其疾病并参与治疗决策, 将可能取得最佳的治疗效果。此外, 先进的计算机系统还可以收集、分析、协调并自动发布健康信息, 使医师和公共卫生工作者能迅速和有效地应对生物恐怖主义的威胁, 应付诸如炭疽、天花、禽流感等紧急疫情或突发自然灾害所导致的人员伤亡和混乱, 协调健康和医疗支持系统。

近年来, 医学信息学的成果已广泛应用于临床、护理、口腔、药剂、公共卫生、理疗以及生物医学研究等领域。随着新的诊断技术、药物、治疗方法和医学软件不断进入医师的日常工作, 医师们必须不断学习新的知识, 并将其用于患者的治疗。在临床工作中, 医师们需要详细记录每个患者的健康状况和既往病史, 并与患者及其家属、医院、服务商和供应商高效沟通。在过去, 满足这些要求较为困难, 因为系统化采集、存储、检索患者信息的标准尚未制定。而今, 计算机系统的发展可以帮助医师们应对这些挑战。

1.1.2 生物信息学发展史

除医疗机构所产生的大量临床数据外, 基础研究所产生的海量数据亦是生物医学信息学需要应对的难题之一。生物信息学是研究生物信息的采集、处理、存储、传播、分析和解释的一门学科, 是以计算机为工具对生物信息进行储存、检索和分析的科学, 也是当今生命科学和自然科学的重大前沿领域之一, 同时也是 21 世纪自然科学的核心领域之一。

生物信息学的起源可以追溯至英国学者 Francis Galton (图 1-2) 所创立的生物测量学 (biometrics)。

20世纪初,统计学和数学方法被用来分析生物学数据。20世纪80年代中期,计算机开始被用来分析各种生物特征,包括视网膜、虹膜、人脸、指纹等,生物信息学应运而生。1981年,美国的 Temple F. Smith 与 Michael S. Waterman 率先将数学和计算方法引入生物学研究,在生物信息学领域做出了开创性的贡献,被誉为“生物信息学之父”(图 1-3、图 1-4)。但“bioinformatics”一词则是华裔学者林华安(Hwa A. Lim)于1987年首创的(图 1-5)。20世纪90年代之后,随着DNA测序、蛋白质质谱分析、分子成像等技术的不断突破,生物信息学的研究重点逐渐转向生物组学(omics)数据,主要从核酸和蛋白质序列出发,分析序列中表达的结构和功能的生物学信息。

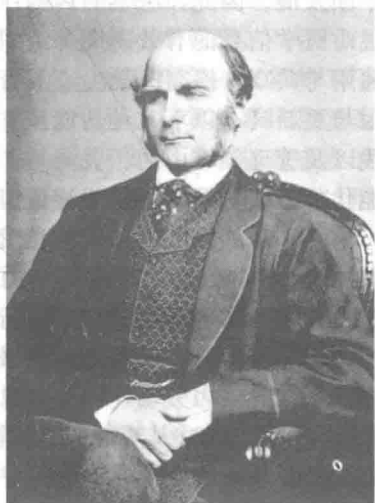


图 1-2 Francis Galton (1822—1911)



图 1-3 Temple F. Smith (1939—)



图 1-4 Michael S. Waterman (1942—)



图 1-5 林华安(Hwa A. Lim) (1955—)

1.2 学科的性质和任务

随着计算机在各个行业的应用普及,各个学科都在形成一个关于该学科信息问题的系统性知识体系。医学信息学与生物信息学均属于典型的交叉学科,既是生物医学信息化的产物,又是信息科学本身发展的必然趋势。构成医学信息学学科群的主体学科包括医学信息学、卫生信息学以

及健康信息学,其支撑学科包括信息科学、管理学、医学,其分支则包括临床信息学、临床研究信息学、牙科信息学、护理信息学、兽医信息学、药学信息学、影像信息学、公共卫生信息学、麻醉医学信息学、实验医学信息学等,相关学科则包括卫生信息管理学、卫生统计学、医学信息工程学、生物医学工程学、数字医学、医学影像学、医学技术学、转化医学等。

医学是研究人类生命过程以及同疾病作斗争的一门科学,其主要目的就是揭示疾病的本质,阐明疾病发生和发展的规律,并寻找有效的防治措施,进而达到消除疾病、保障人类生命健康的目的。医学信息学是伴随着计算机技术在医学领域中的应用以及服务于医学科学研究与实践而产生的,同时也在服务于医学及其相关领域的目标活动中不断发展。因此,医学的目的在很大程度上就是医学信息学学科的目的。医学信息学的任务是通过有效研究与组织、管理与控制、开发与应用,从而实现医学信息(知识)的充分利用与共享,提高医学决策与管理的效率和质量。从实践的角度看,医学信息学手段与方法的应用与更新具有以下几种功能:首先是整合功能,即按照医学信息产生与发展的基本规律对医疗实践以及管理领域中的相关流程、职能进行整合;其次是辅助功能,使临床诊疗技术向现代化、智能化的方向发展,不仅能够帮医师看,还能帮医师想;第三是拓展功能,即利用医学信息的规律,增强和扩展医务人员的信息/智力功能,主要包括加强医师的感觉功能(即医学信息的提取、检测、传递等功能),拓展医师的思维功能(即医学信息的转化、存储、识别、处理和决策等)以及执行功能(即利用医学信息进行调整、转换、控制和功能管理);第四是决策功能。各种信息管理系统在医学领域中的应用和发展,各类医学信息系统的广泛普及和使用,不仅可以帮助医务人员摆脱繁重的体力劳动,而且可以使他们从单调、重复的机械式脑力劳动中解放出来,使他们有更多的时间和精力去从事创造性的智力活动。更为重要的是,临床决策支持系统的应用将为临床工作者、患者以及健康个体提供更好的治疗和保健支持,极大地提高临床诊疗与健康管理的决策水平与质量。

生物信息学在本质上是分子生物学与信息技术的结合体,其研究材料为各种各样的生物学数据,研究工具为计算机,核心内容则是通过对DNA序列的统计和分析,更加深入地理解DNA的序列、结构、演化及其与生物功能之间的关系,其研究课题涉及分子生物学、分子进化、结构生物学、统计学和计算机科学等许多领域。生物信息学是内涵非常丰富的学科,其核心是基因组信息学,包括基因组信息的获取、处理、存储、分配和解释。基因组信息学的关键是“读懂”基因组的核苷酸顺序蕴含的意义,即全部基因在染色体上的确切位置以及各DNA片段的功能;在发现了新基因信息之后进行蛋白质空间结构模拟和预测,然后依据特定蛋白质的功能进行药物设计。此外,了解基因表达的调控机制也是生物信息学的重要内容,力图根据生物分子在基因调控中的作用,探索其在人类疾病的诊断、治疗中的作用。

在短短十几年间,生物信息学已经形成了基因序列比对、蛋白质比对、基因的识别与分析、分子进化、序列拼接、遗传密码研究、药物设计、生物系统、数据分析技术、生物图像分析、基因表达谱分析、代谢网络分析、基因芯片设计和蛋白质组学数据分析等多个研究方向。从技术上看,生物信息学主要涉及新算法和统计学方法研究、各类生物学数据的分析和解释,以及研制和有效利用管理数据的新工具。

1.3 学科发展趋势

在医学信息学领域,最重要的进展集中在以电子病历(electronic patient record, EPR)和医院信息系统(hospital information system, HIS)为核心的临床信息管理领域。医学信息软件及系

统的开发,初步满足了医师与其他医疗保健人员以及患者对于信息和知识的需求,有助于改善患者的护理,便于医院管理者参与定制、采购、实施评估和管理,不断提高医疗质量,以其安全、高效、有效、及时等特点,促进医疗公平的实现。

伴随着医疗信息化,目前的医疗概念已从过去的住院治疗向早期预测和家庭护理方面不断拓展。在过去,患有慢性疾病(如糖尿病或充血性心力衰竭)的患者往往需要定期监测血糖水平、体重、血压等,并在家中服药。很多时候,这类患者在下一次见医师时才能咨询相关的信息,而他们的医师则常常忘记了患者的病情。因此,患者血糖或血压的变化无法被及时发现,健康状况可能恶化,而这又使医疗保健更加昂贵。

随着计算速度的不断提高、互联网技术的不断完善和移动通信技术的创新,近年来又诞生了云储存、云计算、物联网等一批新的IT技术,加上传感器技术的不断创新,这将为上述领域带来根本性的改变。先进的计算机系统将显著提高护理质量,改善用药管理和慢性病患者的自我管理,在降低人工成本的同时,显著提高患者的安全。利用新的数据库技术、网络技术和信息管理技术,研究者可以采集大量的临床数据,通过信息分析确定最佳的治疗方案和临床路径,从而显著提高医疗质量。

近年来,医疗信息网络已逐渐从医院扩展至区域乃至国家范围。继新加坡之后,澳大利亚亦于2012年7月建成并投入使用了全国联网的医疗信息网络。美国的甲骨文(Oracle)公司最近推出了疑难病会诊信息平台,将专家、实验室以及各地的医疗机构联系在一起,打破了医疗机构的传统边界,实现了资源共享,最大限度地便利了患者。IBM公司亦推出了超级计算机“沃森”,通过集成病例记录、医学杂志上刊登的科研论文以及卫生部门掌握的数据等,帮助医师对大量的信息进行分析,并发现诊断和治疗过程中遗漏的细节。

截至2010年,我国90%以上的县级以上医院均建立了挂号收费、药品器材等医院管理信息系统,三分之一以上的医疗机构建立了电子病历系统。其应用显著改变了医院的管理风格,减少了医疗失误,为医护人员提供了实时更新的患者信息,减少了患者等待的时间,提高了护理质量。它不仅成为医院现代管理的象征,也成为医院核心竞争力之一。

20世纪80年代以来,根据其内容、风格和范围的不同,医院信息的管理和使用在我国经历了以下4个阶段:

- (1) 独立的门诊收费、住院收费和药房管理系统。
- (2) 医院的局域网络,包括住院管理、门诊收费和药物管理以及递送系统等。
- (3) 完整的医院信息系统。许多大医院利用网络构建了完整的信息系统,具体包括医院信息系统、医学图像存储与共享信息系统(picture archiving and communication system, PACS)、实验室信息系统(laboratory information system, LIS)、护理信息系统(nursing information system, NIS)等。
- (4) 远程医疗。随着IT和网络的发展,一些大型医疗机构开始研究实施数字图像(如CT、核磁共振和超声)的远程诊断。

从性质上看,前三个阶段集中在医院信息管理,而第四阶段则开始向以患者为中心的理念发展。

在生物信息学方面,得益于DNA克隆及测序技术的革命,人类基因组计划(human genome project)已于2006年正式完成,而针对大量其他物种的DNA测序工作亦进展迅速。随之而来的是各种生物组学数据的海量涌现。这些数据为从分子水平解读生命活动与疾病现象提供了素材。近年来,一些企业已经能够以可接受的价格为个人绘制基因图谱。尽管目前研究者尚不清楚基因

如何影响身体的各个方面，但他们已经清楚某些特定基因将引发心脏病和乳腺癌之类的疾病。可以预期，人类将很快进入个性化医疗的阶段。也就是说，医师将根据患者的基因特征来选择治疗方式。

以高通量 DNA 测序、生物芯片、质谱、双向电泳等为代表的组学极大地提高了实验数据的产量，并对实验设计产生了显著的影响。通过对各种细胞、组织乃至整个生物体的功能状态进行测量，组学技术能够产生从基因序列到蛋白质表达以及代谢模式等多个层次的大量数据，这也是近年来生物医学海量实验数据产生的原因。以我国的华大基因科技有限公司为例，在“无所不测”的思想指导下，这一全球最大的测序中心一天即可产生 3 Tb 以上的序列数据。目前，存储于各种公共数据库的基因、蛋白质、转录物、代谢物等生物数据种类繁多，数量巨大。

解读 DNA 序列数据、生物芯片数据、质谱数据等需要大规模的并行计算。由于这些数据来自不同的生物学层次，整合各种数据来回答生物学问题将涉及序列拼接和预测、模式识别、多尺度建模以及动态模拟等复杂计算问题，因此，计算生物学 (computational biology) 的概念应运而生。反过来，生物数据对于数据结构、算法、数据存储、计算量、计算结果可视化等研究的需求亦推动了计算机科学的发展。

曾几何时，数据是人类用于识别环境的计数工具，对其精确性的关注似乎仅限于科学研究领域。进入信息化时代后，每个人都明显地感受到与信息相关的数据无处不在。在人类不断地产生各种数据的同时，数据也在极大地影响着人类。在数据的发展历程中有过两次革命：第一次是在近代科学诞生之时，实现了数据与科学研究的融合，数据在科学研究中的基础地位得以确立。对研究过程和结果的精确化诉求已成为近代科学的基本特征之一。在以数据为依据的研究范式 (paradigm) 中，数据的可靠性和准确性代表了研究的精确性，人们甚至将以数据为依据的实证研究作为界定“科学”和“伪科学”的标准。随着科学和技术的发展，数据的形式和内涵也在不断地变化和发展。除观测数据、实验数据、理论数据、统计数据、模拟数据外，图、表、文字等也被纳入数据行列，形成了结构化数据和非结构化数据的多元化数据形式。信息技术的发展也从数据匮乏向“数据丰富，理论匮乏”的方向转变，数据产生的速度和规模急剧扩大；数据中蕴含的信息也远远超出了其工具性和依据性的特征，形成了可以从中挖掘出新知识的大数据 (big data)。与统计数据相比，大数据所强调的是全部样本；与科学数据的精确性相比，大数据则允许存在一定范围的不准确性；与科学范式的因果关系相比，大数据则通过关联性寻求自然和社会的变化规律。因此，大数据已引发了第二次数据革命，它不仅将改变科学研究的范式，实现研究的量化，也将促使相关的领域发生巨大的变革。

显然，人类在如何回答生物学问题方面已进入了一个全新的时代，即大生物学时代，对健康和疾病相关的生命现象进行更具系统性的整体测量和系统分析。在研究中，用计算机软件对通过生物芯片、大规模测序、蛋白质谱、生物成像等手段获取的密集型数据进行处理，所产生的信息/知识将存储在计算机中。研究者仅需在后台利用数据管理和统计的方法对这些数据进行处理、分析，获取知识，形成以大数据为基础的密集型科学。由数据驱动 (data-driven) 的研究范式已成为传统的假说驱动 (hypothesis-driven) 的研究方法的有益补充，科学研究的范式也从还原主义 (简化) 向复杂系统转变。

上述发展已为医学研究的许多领域带来了根本的变化。在过去，研究者通常采取假说驱动的策略，即事先提出某种经过推敲的问题或假说，之后再通过实验来验证。而组学研究的策略则无须事先提出问题，而是通过组学技术直接采集基因组或蛋白质组数据，之后通过分析来提出某种理论。这种研究策略的转变有利于发现未知的病理生理学机制以及明确生物体对药物或营养物的