

精准医学大数据应用安全分类分级研究^{*}

刘 莉

陈先来

李忠民

(中南大学生命科学学院
长沙 410000)

(中南大学信息安全与大数据研究院
长沙 410000)

(中南大学生命科学学院
长沙 410000)

安 莹

朱勋梅 谢丽敏 杨 潇 邱 珂 于 冰

(中南大学信息安全与大数据研究院 长沙 410000) (中南大学生命科学学院 长沙 410000)

〔摘要〕 应用文献研究法、线面混合分类法划分精准医学大数据用户角色、应用场景，详细阐述划分方法及结果，指出精准医学背景下基于不同角色、不同应用场景的数据分类分级思路，为精准医学大数据标准化体系构建和安全隐私保护相关研究提供新方向。

〔关键词〕 精准医学；大数据分类分级；数据安全；隐私保护

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2021.01.002

Study on Security Classification of Big Data Application in Precision Medicine LIU Li, School of Life Sciences, Central South University, Changsha 410000, China; CHEN Xianlai, Information Security and Big Data Research Institute, Central South University, Changsha 410000, China; LI Zhongmin, School of Life Sciences, Central South University, Changsha 410000, China; AN Ying, Information Security and Big Data Research Institute, Central South University, Changsha 410000, China; ZHU Xunmei, XIE Limin, YANG Xiao, QIU Ke, YU Bing, School of Life Sciences, Central South University, Changsha 410000, China

〔Abstract〕 The user roles and application scenarios of precision medicine big data are classified by the method of literature research and mixed classification of line and surface. The paper expounds the classification methods and results in detail, and points out that the idea of data classification based on different roles and application scenarios in the context of precision medicine provides a new direction for the building of standardization system of precision medicine big data and related study on security and privacy protection.

〔Keywords〕 precision medicine; big data classification; data security; privacy protection

1 引言

〔收稿日期〕 2020-06-01

〔作者简介〕 刘莉，副教授，硕士生导师，发表论文 6 篇；
通讯作者：陈先来，教授，博士生导师。

〔基金项目〕 国家重点研发计划“精准医学研究”重点专项
基金项目“精准医学大数据体系的规范化应用
与评价”（项目编号：2016YFC0901705）。

人类社会步入大数据时代，大数据技术在各领域发挥重要作用，医学是一个典型应用领域。2016 年我国将精准医学研究列为重点研发计划，随着高通量测序、分子影像等技术的快速发展，精准医学

大数据时代已到来,伴随而来的还有数据有效开发利用及安全隐私保护等问题。精准医学集成多种生命组学数据,通过建模、机器学习等统计分析可以对大样本队列数据和疾病人群进行整合研究,明确基因、蛋白质等组学分子和环境、生活方式的个体差异,更好了解患者健康、疾病情况的复杂机理,进行疾病预防和治疗^[1-3]。而精准医学大数据使用过程中涉及病人隐私,具有特殊性和敏感性,存在泄露或暴露数据提供者个人信息和隐私的风险^[4]。数据分类分级是安全与隐私保护中必不可少的一部分,从泄露风险大小和范围、可识别程度等角度出发,对数据进行分类和等级划分,进而根据不同需要对关键数据进行重点防护。目前研究较少涉及详细精准的医学大数据分类分级标准,本文对精准医学研究背景下不同用户角色及应用场景进行划分,详细列出在精准药物研发、精准预防、精准诊断、精准治疗4个场景下可能涉及的数据范围,拟定标准对精准医学数据进行分级分类,旨在形成一套规范的精准医学数据分级分类体系,指导精准医学大数据应用的规范化、合理化发展,在安全前提下最大程度地提高精准医学大数据利用率^[5]。

2 资料与方法

2.1 资料来源

针对精准医学大数据应用安全分类分级进行研究,涉及数据主要为临床数据与组学数据。临床相关数据集来源于原卫生部发布标准《电子病历基本数据集》;组学数据相关数据集从北京基因组研究所组学原始数据库系统(Genome Sequence Archive, GSA)、北京蛋白质组研究中心(Beijing Proteome Research Center, BPRC)蛋白质组学数据库和快捷式医疗保健互操作资源(Fast Health Interoperable Resources, FHIR)中进行收集整理。

2.2 研究方法

2.2.1 文献研究法 通过查阅精准医学、数据分

类分级等相关论文、政策、标准规范及法律法规,对精准医学大数据应用的用户角色、应用场景进行划分。针对精准医学研究相关文献进行分析,首先确定研究所属场景,然后对文献中涉及数据进行归纳,并与已有数据集中数据项进行合并,基于此确定每个场景下可能涉及的数据项。接下来对当前庞大繁杂的精准医学数据进行合理有效分类,形成精准医学数据分类标准。在精准医学大数据分类标准基础上对分类之后的精准医学数据采取合理的安全分级原则,即能否通过单一数据项或多数据项联合识别到个人,将其中具体数据项从数据安全和隐私保护角度进行分级。

2.2.2 线面混合分类法 线分类法将信息对象按照选定特性等分类基础逐层分成若干层级的类,形成一个有序、逐级开展的分类体系^[6];面分类法将信息对象若干特性看成若干个“面”,每个“面”中又可以分成彼此独立的若干个类目^[7]。线面混合分类法是指面分类法和线分类法一起组合使用,其中一种为主要分类法,另一种发挥补充信息等作用。精准医学大数据可以按数据来源、数据特征、应用场景等方式进行分类,但考虑到医学数据来源、种类多,不同数据安全隐私保护要求不同,采用面分类法与线分类法相结合方法开展精准医学大数据分类研究,使用面分类法将精准医学大数据划分为28个彼此独立的大类,再使用线分类法在每个大类下划分具有层级结构的小类。

3 研究结果

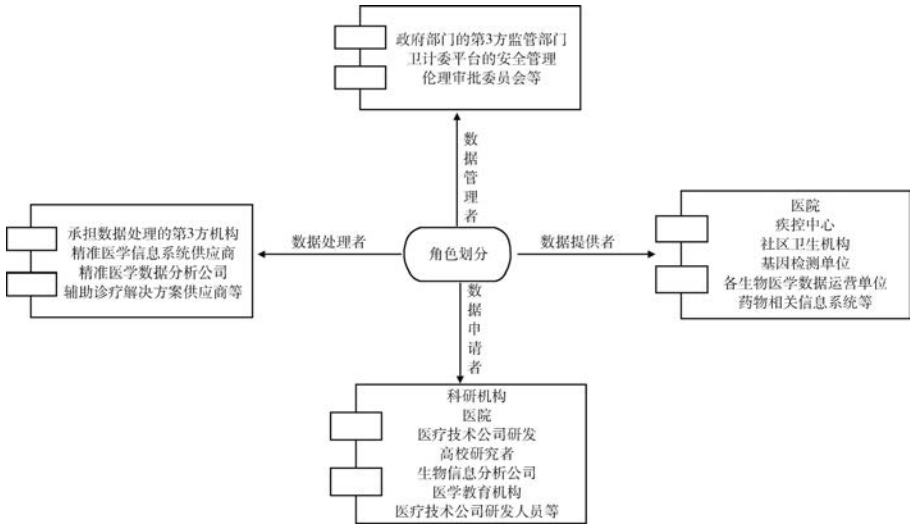
3.1 精准医学大数据用户角色划分

3.1.1 划分方法 由于精准医学大数据分类分级与用户行为紧密相关,因此可以考虑对用户角色进行划分,不同角色拥有不同权限,以限制用户在每个特定场景下可以获取的数据范围与级别。首先利用 Google Scholar、Web of Science、知网等学术搜索引擎检索与本研究相关的文献;其次以 Mendeley 建立电子文献索引,以云端同步 PDF 文献为主要储存

手段；最后利用文献记录表进行文献研究总结。

3.1.2 角色划分结果 将精准医学大数据用户角色分为 4 类：数据提供者、数据处理者、数据管理者、数据申请者，见图 1。数据提供者是在精准医学研究中向大数据平台提交基础数据或成果数据（2 次数据）等的组织或个人，主要在提供研究数据的场景下发挥其职能；数据处理者是对上传到精准医学大数据平台和平台共享的数据进行合并、变换、分组、脱敏、加密等工作的相关组织或个人，主要是在数据加工处理的场景下发挥职能，其在保证数据后续使用的安全与隐私、数据质量等方面发

挥重要技术作用；数据管理者是能决定精准医学数据处理、共享目的、方式及范围等的组织或个人，在科研场景和精准医学大数据背景下主要指大数据中心对精准医学数据具有控制决策权限的管理人员；数据申请者是在科研场景下，对精准医学数据有使用需求而向平台提出申请数据进行利用的相关组织或个人^[8-9]。需要强调的是同一具体科研人员可能会被定义为不同角色。例如当其向平台提出数据申请时属于数据申请者；而当他向平台上传数据时则属于数据提供者。也就是说角色的划分不是由个人属性决定的，而是由正在进行的行为决定的。



3.2 精准医学大数据应用场景划分

3.2.1 必要性 本文主要研究精准医学大数据汇聚后数据开放共享过程中的安全与隐私保护问题，核心在于对精准医学研究相关各类数据按应用场景进行安全分级。在精准医学大数据应用环境中，与数据安全相关最主要主体是数据申请者，能否授权给相关用户使用是关键，因此本文着重强调数据的申请者及应用场景。数据申请者根据其研究内容对精准医学大数据平台中部分数据提出需求申请。不同项目的研究目的、应用场景不同，所需数据范围

也不相同。因此有必要对精准医学大数据应用场景进行划分，以明确不同场景下的数据需求范围，为数据使用授权提供支持。

3.2.2 场景划分 精准医学发展涉及多领域，其依据患者内在生物学信息为临床疾病亚型群体提供更精确的诊断和治疗，主要应用于疾病的预防、诊断、筛查和治疗等方面。根据研究目标的不同将精准医学大数据应用划分为 4 类典型场景：精准诊断、精准治疗、精准预防、精准药物研发，每个应用场景下的数据申请者能够获取的数据范围有所不同。4 个场景详细内容，见表 1。

表 1 精准医学大数据应用场景划分

场景名称	场景描述	研究举例
精准诊断	精准诊断是指通过基因组、蛋白质组等组学技术和医学前沿技术，基于病人个体生物学信息和临床数据来精确病因、病理形态和病理生理内容，从而对疾病进行精准判断 ^[10] 。与传统诊断相比精准诊断在形态学分析基础上更重视对功能信息尤其是遗传和代谢信息的获取，以明确患者遗传背景，揭示其疾病状态下细胞功能和代谢的动态改变，为复杂疾病的早期诊断与精准定位提供有效信息 ^[11]	循环肿瘤细胞（Circulating Tumor Cell, CTC）的无创基因组学研究代替传统 CTC 技术监测肿瘤进展和治疗效果 ^[12]
精准治疗	精准治疗是指以患者个人基因组信息为基础，通过分析蛋白质组学、代谢组学等内环境信息 ^[13] 寻找该疾病针对病人个体特点的致病原因或精确的基因治疗靶点，最终对特定患者实施个性化的最佳诊疗方案，以实现治疗效果最大化和副作用最小化 ^[14]	CT 引导精准微波消融术联合经肝动脉栓塞化术（Transcatheter Arterial Chemoembolization, TACE），协同治疗原发性肝癌患者可显著增强疗效 ^[15]
精准预防	精准预防是应用现代遗传、分子影像、生物信息技术结合患者生活环境和临床数据分析疾病表型特征，在了解疾病发生原因和分子机制的基础上 ^[16] 探讨疾病易感基因及其与环境因素的交互作用 ^[17] ，应用先进的分子遗传学技术进行致病突变检测，建立疾病基因突变谱 ^[18] ，制定具有个性化的疾病预防方案	安吉丽娜·朱莉根据 DNA 测序结果和家族病史切除乳腺以预防乳腺癌和卵巢癌
精准药物研发	精准药物研发是指通过基因组、蛋白质组、转录组等组学技术和医学前沿技术，识别、直接指向疾病的新靶点 ^[19] ，通过分析疾病相关基因和通路，进行高通路筛选和药物合理设计，针对特异性病人进行个性化药物研发 ^[20] ，主要包括：（1）药物结构及成分设计、改善。（2）发现已有潜在治疗药物。（3）已有治疗药物筛选。通过精准医学在药物研发中找到合适的靶点、适应证、人群，进而提高药物疗效，为解决新药研发困境和临床应用提供新方向	将高载药量的纳米晶技术、血液长循环的细胞膜包覆技术与病灶组织导航的靶向技术相结合，研发出一种可用于脑肿瘤治疗的新型智能纳米药物

3.3 精准医学大数据隐私级别划分

在数据可识别程度方面，由强到弱可用直接标识符、准标识符、非标识符（一般数据）分别标识表示数据的可识别程度。借此思路将数据分为 3 级，由强到弱表示精准医学大数据中相应级别数据需要的安全保密程度。最强是 2 级，属直接标识符类数据，在特定环境下可以单独识别个人信息主体^[21]；其次是 1 级，属准标识符类数据，该数据项本身并不能识别特定个人，但可与其他数据信息聚

合和联系起来识别数据主体^[22]；0 级是一般数据，指除去准标识符和直接标识符特征之外的通用型数据项，0 级数据与数据主体的关联关系较弱，既不能像直接标识符一样直接识别单个主体身份数据，也不能通过同类数据项间聚合和联系识别数据主体。因此在对外发布数据集之前一般不需要对 0 级数据进行删除或屏蔽等处理，可直接发布或者采取适当技术对其进行简单处理^[23]。通常在数据发布前会设有“使用条款”等形式限制，0 级数据可作为准公开数据提供给数据需求者使用，见表 2。

表 2 精准医学大数据安全分级情况

安全等级（风险）	安全保密程度	代表的标识符	示例
2 级	强	直接标识符	姓名、地址、电子邮件地址、电话号码、传真号码、信用卡号码、车牌号码、车辆识别号码、社会保险号码、健康卡号码、病历号码、设备标识符、生物识别码、互联网协议（Internet Protocol, IP）地址号和网络通用资源定位符（Uniform Resource Locator, URL）等
1 级	需要加密	间接标识符	性别、出生日期或年龄、事件日期（例如入院、手术、出院、访问）、地点（例如邮政编码、工作单位地址、收件地址）、民族、国籍、语言、职业、婚姻状况、受教育水平等
0 级	弱	一般数据	西医诊断编码、手术及操作编码、主诉、药物名称、药物规格、检查类别、麻醉方法代码等

3. 4 精准医学大数据类别划分

依据原卫生部发布的《电子病历基本数据集》以及北京基因组研究所组学原始数据库系统、快捷式医疗保健互操作资源、北京蛋白质组研究中心中涉及的数据集，按照线面混合分类法，将精准医学大数据划分为 28 个类别，见表 3。

表 3 精准医学大数据类别划分表

类别	2 级项	安全等级
病例概要	传染病史、个人史等	0 级
	出生日期、婚育史等	1 级
门（急）诊病历	既往史、科室名称等	0 级
	就诊时间、性别代码等	1 级
	患者姓名等	2 级
门（急）诊处方	处方类别代码等	0 级
	处方开立日期等	1 级
	处方开立医师签名等	2 级
检查检验记录	标本类别、病区名称等	0 级
	检查报告日期、住院号等	1 级
	电话、报告医师签名等	2 级
一般治疗处置记录	操作次数、呼吸频率等	0 级
	病床号、随访日期等	1 级
	手术者签名、住院号等	2 级
助产记录	产后宫缩、腹围（cm）等	0 级
	待产日期时间	1 级
	产妇姓名、住院号等	2 级
护理操作记录	护理操作名称等	0 级
	出手术室日期、时间等	1 级
	交接护士签名等	2 级
护理评估与计划	安全护理代码等	0 级
	国籍代码、民族等	1 级
	联系人电话号码等	2 级
知情告知信息	患者基础疾病等	0 级
	病危（重）通知内容等	1 级
	法定代理人签名等	2 级
住院病案首页	病理诊断疾病编码等	0 级
	出院日期、时间等	1 级
	电话号码等	2 级
中医住院病案首页	辨证施护标志、病理号等	0 级
	出生地－县（区）等	1 级
	患者身份证件号码等	2 级
入院记录	出院医嘱、传染病史等	0 级
	初步诊断日期等	1 级
	接诊医师签名等	2 级

续表 3

住院病程记录	病例摘要、会诊类型等	0 级
	查房日期、时间等	1 级
	参加抢救人员名单等	2 级
住院医嘱	处方药品组号等	0 级
	体重（kg）、性别代码等	1 级
	医嘱开立者签名等	2 级
出院小结	切口愈合等级代码等	0 级
	病区名称、出院情况等	1 级
	居民健康卡号等	2 级
转诊（院）记录	辅助检查结果等	0 级
	转出医疗机构代码等	1 级
	转诊医师姓名等	2 级
医疗机构信息	医疗机构组织机构代码等	1 级
数据提交信息	文件数量	0 级
	发布日期、单位网址等	1 级
	联系邮件、名称等	2 级
实验信息	测序平台、MS 工具等	0 级
项目及机构信息	子项目状态、计划等	0 级
	先导项目、发布日期等	1 级
	项目标题、数据提供者等	2 级
样本测定信息	样本描述、细胞类型等	0 级
	取样时间、病变型等	1 级
	数据区块、样本标题等	2 级
分析和识别信息	图标分数、覆盖率等	0 级
	肽序列、蛋白质组等	1 级
	染色体、起始位置等	0 级
注释信息	转录本、剪切体等	1 级
	碱基序列组成等	2 级
链接	PubMed ID、参考文献等	0 级
FHIR 基因序列	类型、参考序列等	0 级
	参考序列变异起始点等	1 级
	标识符、样本等	2 级
FHIR 观测报告	观测报告，元数据等	0 级
	观测报告，染色体相集等	1 级
	观测报告，ID 等	2 级
FHIR 诊断报告	诊断报告，语言等	0 级
	诊断报告，隐式规则等	1 级
	诊断报告，编号等	2 级
FHIR 流程申请书	流程申请，修改拓展等	0 级
	流程申请，支撑信息等	1 级
	请求者，ID 等	2 级

3.5 不同应用场景下所需数据的类与级

根据对 4 个应用场景下典型研究内容的分析，每个应用场景下收集 5 ~ 10 个典型案例，结合其所需数据划分精准医学大数据平台不同应用场景所需数据范围，见表 4。可以看出，病历概要、门（急）诊病历、检查检验记录、实验信息、分析和识别信息、注释信息、链接、FHIR 基因序列数据、FHIR 观测报告信息、FHIR 诊断报告信息 10 个大类是精准诊断^[24-26]、精准治疗^[27-29]、精准预防^[30-32]、精准药物研发^[33-35] 4 个场景下科学研究需要的数据类别。助产记录、护理评估与计划、住院病程记录、住院医嘱、出院小结、转诊（院）记录、医疗机构信息、数据提交信息、FHIR 流程申请书 9 个类别是 4 个场景下科学研究目前不需要获取的数据，可暂时选择不开放。

表 4 不同研究场景下所需数据的类别情况

数据类别	场景需要情况			
	精准诊断	精准治疗	精准预防	精准药物研发
病历概要	√	√	√	√
门（急）诊病历	√	√	√	√
门（急）诊处方	○	√	○	√
检查检验记录	√	√	√	√
一般治疗处置记录	○	√	○	○
助产记录	○	○	○	○
护理操作记录	○	√	○	○
护理评估与计划	○	○	○	○
知情告知信息	○	○	○	○
住院病案首页	√	○	√	√
中医住院病案首页	√	○	√	○
入院记录	○	○	√	○
住院病程记录	○	○	○	○
住院医嘱	○	○	○	○
出院小结	√	√	√	○
转诊（院）记录	○	○	○	○
医疗机构信息	○	○	○	○
数据提交信息	○	○	○	○
实验信息	√	√	√	√
项目及机构信息	○	√	√	○
样本测定信息	√	√	√	○
分析和识别信息	√	√	√	√
注释信息	√	√	√	√
链接	√	√	√	√
FHIR 基因序列	√	√	√	√
FHIR 观测报告	√	√	√	√
FHIR 诊断报告	√	√	√	√
FHIR 流程申请书	○	○	○	○

注：√表示该场景下需要用到这一数据类别；○表示该场景下不需要用到这一数据类别。

4 结语

目前国内外精准医学大数据分级分类研究几乎空白，精准医学研究属于新兴领域，数据分类分级概念虽然已有研究提出^[36-38]，但是未涉及精准医学领域数据分类分级及安全指标体系构建^[39]。在此背景下本研究提出精准医学背景下基于不同角色、不同应用场景的数据分类分级思路，为精准医学大数据标准化体系构建和安全隐患保护相关研究提供新方向。但在研究中存在不足。首先，研究中涉及的精准医学数据范围不够全面，随着高通量、多组学、大数据技术的结合与快速发展，精准医学研究中涉及数据呈现出类型繁多、结构多样化、数量庞大^[40-41]等特点，本研究中只针对典型的精准医学研究数据进行了探讨。其次，研究中拟定的精准医学大数据静态分类分级标准及模型具有一定局限性，仅基于文献资料研究而未投入实际应用所以无法预知该模型是否能够契合精准医学研究的实际情况。需在后续研究中根据实际应用情况扩大补充数据范围、进行实时动态调整，形成更为完善的数据分类分级标准，构建能为精准医学研究提供精准服务的数据分类分级动态模型，从而促进精准医学发展。

参考文献

- 1 The White House Office of the Press Secretary. Fact Sheet: President Obama’s Precision Medicine Initiative [EB/OL]. [2019-10-25]. <https://obamawhitehouse.archives.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative>.
- 2 赵晓宇, 刁天喜, 高云华, 等. 美国“精准医学计划”解读与思考 [J]. 军事学, 2015, 39 (4): 241-244.
- 3 唐金陵, 李立明. 关于循证医学, 精准医学和大数据研究的几点看法 [J]. 中华流行病学杂志, 2018, 39 (1): 1-7.
- 4 Azencott C A. Machine Learning and Genomics: precision medicine versus patient privacy [J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2018, 376 (2128): 20170350.
- 5 娄培, 刘莉, 陈先来, 等. 基于问卷调查的医疗数据分

- 类分级研究 [J]. 中华医学图书情报杂志, 2018, 27 (6): 22–27.
- 6 周建勤, 张铎. 物流信息分类编码标准体系探讨 [J]. 物流技术, 2000 (2): 372–391.
- 7 宋晓, 韩璐遥, 梁建峰, 等. 海洋环境安全数据分类体系研究 [J]. 海洋信息, 2019, 34 (1): 1–5.
- 8 李英杰. 基于大数据分析的舰船通信网络访问权限控制方法 [J]. 舰船科学技术, 2018, 40 (18): 121–123.
- 9 马志程, 杨仕博, 邵诗韵, 等. 数据中心资源池用户身份识别与权限控制 [J]. 电子技术与软件工程, 2018 (14): 163–164.
- 10 魏洪泽, 李玉杰. 精准医疗与伴随诊断产业发展研究 [J]. 中国生物工程杂志, 2019, 39 (2): 13–21.
- 11 洪专. 精准时代检测先行, 生存获益基于精准的诊断、治疗和预后判断 [J]. 医学信息, 2018, 31 (17): 1–3.
- 12 Ghasemi M, Nabipour I, Omrani A, et al. Precision Medicine and Molecular Imaging: new targeted approaches toward cancer therapeutic and diagnosis [J]. Am J Nucl Med Mol Imaging, 2016, 6 (6): 310–327.
- 13 孙会杰. 精准医疗在肿瘤治疗中的应用 [J]. 现代商贸工业, 2019 (8): 93–94.
- 14 谭定勇, 邓青松, 龚建平. 肝脏外科精准治疗研究进展 [J]. 国际外科学杂志, 2019, 46 (2): 136–140.
- 15 沈珊珊, 江德鹏. CT 影像组学指导肺癌个体化治疗的研究进展 [J]. 影像研究与医学应用, 2019, 3 (10): 7–9.
- 16 杨瑞馥. 精准预防医学: 以大数据为基础精准医学的可预期目标 [J]. 中华预防医学杂志, 2015, 49 (12): 1025–1027.
- 17 Kenneth S Ramos, Emma C Bowers, Marco A Tavera – Garcia, et al. Precision Prevention: a focused response to shifting paradigms in healthcare [J]. Experimental Biology and Medicine, 2019 (244): 207–212.
- 18 Francis S Collins, Harold Varmus. A New Initiative on Precision Medicine [J]. The New England Journal of Medicine, 2015, 372 (9): 793–795.
- 19 展鹏, 王学顺, 刘新泳. “精准医疗”背景下的分子靶向药物研究——精准药物设计策略浅析 [J]. 化学进展, 2016, 28 (9): 1363–1386.
- 20 杨凌. 精准药物治疗的核心理念与科学内涵 [J]. 药学进展, 2017 (2): 19–21.
- 21 郑琳. 面向机构的科研人员唯一标识符应用与发展研究 [J]. 图书与情报, 2018, 38 (1): 134–140.
- 22 卞超轶, 朱少敏, 周涛. 一种基于 Spark 的大数据匿名化系统实现 [J]. 电信科学, 2018, 34 (4): 156–161.
- 23 周倩伊, 王亚民, 王闯. 基于互联网大数据的脱敏分析技术研究 [J]. 数据分析与知识发现, 2018, 2 (2): 58–63.
- 24 孙凤英, 于修义. 基于胸外科结构化电子病历的人工智能诊断系统构建实践 [J]. 中国数字医学, 2018, 13 (11): 29–31.
- 25 Liu Zhenyu, Wang Shuo, Dong Di, et al. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: opportunities and challenges [J]. Theranostics, 2019, 9 (5): 1303–1322.
- 26 何敬成, 李茹冰, 张碧青, 等. 基于临床决策支持技术实现高血压用药合理性分析的开发与应用研究 [J]. 海峡药学, 2019, 31 (2): 291–292.
- 27 赵妍, 王颖, 闫国涛, 等. 基于数据仓库的临床决策支持系统在我院的应用 [J]. 中国医疗设备, 2016, 31 (7): 95–97, 101.
- 28 Vishesh K Kapur, Dennis H Auckley, Susmita Chowdhuri. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: an American Academy of sleep medicine clinical practice guideline [J]. J Clin Sleep Med, 2017, 13 (3): 479–504.
- 29 Jonathan P Piccini, Christopher B Granger. Insulin Therapy and Stroke Risk in Patients with Diabetes and Atrial Fibrillation: guilty by association [J]. Journal of the American College of Cardiology, 2017, 69 (4): 420–422.
- 30 郑西川, 陈霆, 孙宇, 等. 基于临床数据中心的医疗管理和临床决策信息化应用实践 [J]. 中国医院, 2017, 21 (11): 64–66.
- 31 张慧娟. 基于电子住院病历的临床决策支持系统的设计与应用 [J]. 中国医疗管理科学, 2016, 6 (6): 10–14.
- 32 Matthew M Churpek, Ashley Snyder, Xuan Han, et al. Quick Sepsis –related Organ Failure Assessment; Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit [J]. American Journal of Respiratory and Critical Care Medicine, 2017 (195): 906–911.
- 33 王伟, 祁永飞, 包晗, 等. 基于医药大数据建立精准药理学临床评估体系的思考 [J]. 中国药事, 2017, 31 (12): 1478–1482.

(下转第 35 页)

- formation Extraction [C]. Washington DC: Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management, 2004: 581 – 588.
- 27 Suchanek F M, Ifrim G, Weikum G. Combining Linguistic and Statistical Analysis to Extract Relations from Web Documents [C]. Philadelphia: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006: 712 – 717.
 - 28 Rozenfeld B, Feldman R. High – performance Unsupervised Relation Extraction from Large Corpora [C]. Hong Kong: Sixth International Conference on Data Mining (ICDM06). IEEE, 2006: 1032 – 1037.
 - 29 Qingchen Zhang, Changchuan Bai, Zhikui Chen, et al. Smart Chinese Medicine for Hypertension Treatment with a Deep Learning Model [J]. Journal of Network and Computer Applications, 2019 (129): 1 – 8.
 - 30 Kim Y. Convolutional Neural Networks for Sentence Classification [C]. Doha: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014: 1746 – 1751.
 - 31 Liu S, Tang B, Chen Q, et al. Drug – drug Interaction Extraction via Convolutional Neural Networks [J]. Computational and Mathematical Methods in Medicine, 2016 (2016): 1 – 8.
 - 32 Hua L, Quan C. A Shortest Dependency Path Based Convolutional Neural Network for Protein – protein Relation Extraction [J]. BioMed Research International, 2016 (2016): 1 – 9.
 - 33 Pandey C, Ibrahim Z, Wu H, et al. Improving RNN with Attention and Embedding for Adverse Drug Reactions [C]. Beijing: Proceedings of the 2017 International Conference on Digital Health, 2017: 67 – 71.
 - 34 Xu B, Shi X, Zhao Z, et al. Leveraging Biomedical Resources in Bi – LSTM for Drug – drug Interaction Extraction [J]. IEEE Access, 2018 (6): 33432 – 33439.
 - 35 Qin Y, Zeng Y. Research of Clinical Named Entity Recognition Based on Bi – LSTM – CRF [J]. Journal of Shanghai Jiaotong University (Science), 2018 (23): 392 – 397.
 - 36 Lin C, Miller T, Dligach D, et al. A BERT – based Universal Model for Both within – and Cross – sentence Clinical Temporal Relation Extraction [C]. Minneapolis: Proceedings of the 2nd Clinical Natural Language Processing Workshop, 2019: 65 – 71.
 - 37 Tian C, Zhao Y, Ren L. A Chinese Event Relation Extraction Model Based on Bert [C]. Chengdu: 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD). IEEE, 2019: 271 – 276.
 - 38 Xiaohui Zhang, Yaoyun Zhang, Qin Zhang, et al. Extracting Comprehensive Clinical Information for Breast Cancer Using Deep Learning Methods [J]. International Journal of Medical Informatics, 2019 (132): 103985.

(上接第 15 页)

- 34 陈柯羽, 杜清, 张华, 等. 精准医疗背景下的药品研发策略 [J]. 转化医学电子杂志, 2018, 5 (10): 55 – 61.
- 35 龚兆龙, 林毅晖, 袁泰昌, 等. 精准医学时代的抗肿瘤药物研发 [J]. 药学进展, 2017, 41 (2): 97 – 100.
- 36 赵鹏, 马泽君, 乐嘉伟. 银行数据资产安全分级标准与安全管理建设方法 [C]. 北京: 软科学国际研讨会, 2012.
- 37 李松涛, 谢宗晓. 数据分类/分级及其相关标准解析 [J]. 中国质量与标准导报, 2019 (4): 14 – 16.
- 38 王敏. 大数据时代如何有效保护个人隐私? ——一种基于传播伦理的分级路径 [J]. 新闻与传播研究, 2018, 25 (11): 69 – 92, 127 – 128.
- 39 Haendel M A, Chute C G, Robinson P N. Classification, Ontology, and Precision Medicine [J]. New England Journal of Medicine, 2018, 379 (15): 1452 – 1462.
- 40 Lin E, Tsai S. Multi – omics and Machine Learning Applications in Precision Medicine [J]. Current Pharmacogenomics & Personalized Medicine, 2017, 15 (2): 97 – 104.
- 41 Tebani A, Afonso C, Marret S, et al. Omics – based Strategies in Precision Medicine: toward a paradigm shift in inborn errors of metabolism investigations [J]. International Journal of Molecular Sciences, 2016, 17 (9): 1555.