

· 专题二:双清论坛“全维度数据与智能诊疗的前沿与挑战”·

# 基于多尺度多模态数据的脑疾病研究进展\*

——遗传影像大数据研究

程 炜<sup>†</sup> 王守岩<sup>†</sup> 谢小华<sup>†</sup> 冯建峰<sup>\*\*</sup>

复旦大学类脑人工智能科学与技术研究院,上海 200433

**[摘要]** 脑疾病是当前世界各国面临的共同挑战,给个人、家庭及社会都带来沉重的负担。随着各种高通量技术和无线远程获取技术的迅速发展,为脑科学与脑疾病的研究提供了前所未有的契机。借助多尺度多模态的脑科学大数据,我们能够从不同尺度、不同角度对人脑功能及脑疾病展开研究。本文将遗传影像大数据为例,首先介绍国内外遗传影像大数据的数据库建设进展与挑战。进一步我们将讨论当前遗传影像数据挖掘分析的主要数学理论和方法,以及遗传影像大数据与智能算法在揭示脑疾病病理机制和辅助临床诊断治疗中的应用。最后,我们就多尺度多模态大数据与脑疾病未来发展方向——疾病分型、计算神经科学、计算精神病学进行总结与展望。

**[关键词]** 脑疾病;磁共振成像;遗传影像;大数据;人工智能

## 1 全球脑疾病现状及趋势

随着全球人口老龄化和各种社会压力的加剧,精神疾病、神经系统疾病以及物质使用障碍等脑疾病正在成为世界各国面临的共同挑战;这些疾病的负担与所有疾病负担比达到14%,严重的精神疾病患者比普通人少活10~20年<sup>[1]</sup>。基于此,2007年Prince等研究者在*The Lancet*上发表题目为“没有精神健康就没有健康(No health without mental health)”的论文<sup>[2]</sup>。2008年世界卫生组织(WHO)也提出了旨在通过提高对这些疾病的关怀、预防与治疗来提高精神健康的“精神健康差距行动计划”<sup>[3]</sup>。

据最新数据统计,2017年世界范围内有9.7亿精神疾病患者<sup>[4]</sup>。常见精神疾病有抑郁症、双相障碍、精神分裂症和自闭症谱系障碍等,每5个儿童和青少年中就有1个人患有精神疾病;一半的精神疾病患者在14岁之前就已经出现症状;全球每年有80万的自杀患者,而这些患者中90%以上患有精神



**冯建峰** 教授,上海数学中心首席科学家,复旦大学类脑智能科学与技术研究院院长、大数据学院院长,上海脑科学与类脑研究中心副主任。长期致力于发展数学、统计与计算机的理论和方法,并原创性地将它们应用于解决神经科学和各类脑疾病中的具体问题。2011年获英国皇家学会沃夫森研究功勋奖。



**程炜** 复旦大学类脑人工智能科学与技术研究院青年研究员,博士生导师。主要从事多模态脑影像数据方法学及其临床应用研究,承担国家自然科学基金,上海自然科学基金等多项课题,至今发表学术论文30余篇。

疾病<sup>[5]</sup>。抑郁症是最常见的精神疾病之一,影响全世界2.6亿人口,世界范围内患病率为4.4%,其中我国为4.2%。此外,全球有4500万双向障碍患者、2000万精神分裂症患者、3000万自闭症谱系障碍患者,每160个儿童中就有一个儿童患有该障碍,

收稿日期:2020-06-21;修回日期:2021-01-28

\* 本文根据第232期“双清论坛”讨论的内容整理。

\*\* 通信作者: Email: jianfeng64@gmail.com

本文受到国家重点研发计划资助项目(2019YFA0709502)、高等学校学科创新引智计划资助项目(B18015)、上海市科学技术委员会基础研究领域第二批项目(16JC1420402)、上海市级重大科技专项(2018SHZDZX01)、科技部重大专项(2018YFC1312900)和国家自然科学基金重大研究计划项目(91630314)的资助。

<sup>†</sup> 共同第一作者

大多数患病儿童在 6 岁前就有症状<sup>[6]</sup>。最新的精神疾病流行病学研究表明我国精神病症的平均患病率达到 16.6%，其中焦虑症是患病率最高的精神疾病，达到 7.6%<sup>[7]</sup>。据统计，上海精神卫生中心 2019 年的门诊病人数达到 100 多万人次。2019 年 10 月第十三届 WHO 精神卫生论坛在《精神健康行动计划(2013—2020 年)》的基础上提出了 2019—2023 年精神健康的工作愿景为“加强各国在精神卫生上行动”且将计划延长到 2030 年<sup>[8, 9]</sup>。我国也于 2014 年底制定了《全国精神卫生工作规划(2015—2020 年)》。

除了精神疾病，神经系统疾病也是当前脑疾病中的重要挑战。2017 年全球有 5 000 万人患有痴呆，且每年有 1 000 万新增患者，每 20 年患痴呆人数会翻倍。这些患者中的 60%~70% 患有阿尔兹海默症<sup>[10, 11]</sup>。我国目前阿尔兹海默症患者已经超过 700 万，已成为继心血管疾病、恶性肿瘤、脑卒中之后危害老年人健康的重大问题<sup>[12, 13]</sup>。痴呆症不仅给患者带来了很大的身体、心理、经济负担，对家庭和社会也造成很大的负担。据估计，2015 年痴呆产生的社会损失多达 8 000 亿美元，相当于当年全球 GDP 的 1.1%，这个损失在 2030 年将会达到 20 000 亿<sup>[14]</sup>。基于此，2017 年 5 月 WHO 批准了针对痴呆的 2017—2025 全球行动计划，旨在提高全球对痴呆的意识，建立相关计划来减少痴呆风险、加强痴呆诊断、治疗、关爱以及研究<sup>[11]</sup>。

下面我们将多模态脑科学数据库、数据分析的数学理论方法、大数据在脑疾病中机制研究与临床诊疗中应用以及未来研究方向这五个方面展开讨论(图 1)。

## 2 国内外多尺度多模态脑科学数据库建设进展和挑战

面对脑疾病防治的严峻形势，全球各国加大脑疾病研究的步伐，标准化、大样本、多尺度、多模态数据采集平台的建设对于脑疾病的机制解析、早期预警、精准诊疗意义重大，已成为前沿研究的基础和共识。在基因层面，由于脑疾病脑认知普遍具有遗传性的特点，因此，遗传数据是我们研究人脑尤其是脑疾病的重要手段，经过几十年的发展，世界各地已积累了海量的脑疾病遗传数据，这些遗传大数据对于我们揭示脑疾病的遗传机制具有重要意义<sup>[15]</sup>。在分子层面，艾伦脑科学研究所(Allen Institute for Brain Science)发布了首个人脑基因表达数据，详细刻画了每个基因在全脑表达的数据(The Allen

Human Brain Atlas)，使得我们能够探索人脑各区域脑功能实现的分子遗传基础<sup>[16]</sup>；在细胞层面，骆清铭等人提出全脑显微光学切片断层成像技术，能够实现亚细胞分辨的小鼠全脑三维神经元连接图谱，使得全脑神经元连接数据库建设成为可能<sup>[17]</sup>。在脑区层面，借助磁共振等成像技术，我们能够无创地实现微米级别脑结构数据的采集，同时能够长时间对脑功能活动数据进行采集。这些不同尺度数据的不断积累成为我们揭开人脑奥秘必不可少的工具。鉴于遗传影像数据采集的无创性、易于开展性以及技术设备的日趋成熟，基于遗传影像数据的人脑数据库建设已经成为当前脑科学与脑疾病的热点，也成为全球大规模脑科学数据库建设的重点。因此，本综述主要聚焦基于多尺度多模态基因脑影像的脑疾病研究前沿与进展。

随着信息获取技术如脑电图(Electroencephalography, EEG)、正电子发射断层扫描(Positron Emission Tomography, PET)、磁共振成像(Magnetic Resonance Imaging, MRI)等各种人脑功能与结构成像设备的不断发展，为脑疾病研究积累了前所未有的多维度数据。其中，磁共振成像技术的发展更为我们提供了观测和研究人脑的利器，该技术可通过无创方式对脑疾病患者的脑功能与结构数据进行采集。然而，早期的磁共振影像研究普遍基于小样本数据，大多数研究不到一百个被试，由于

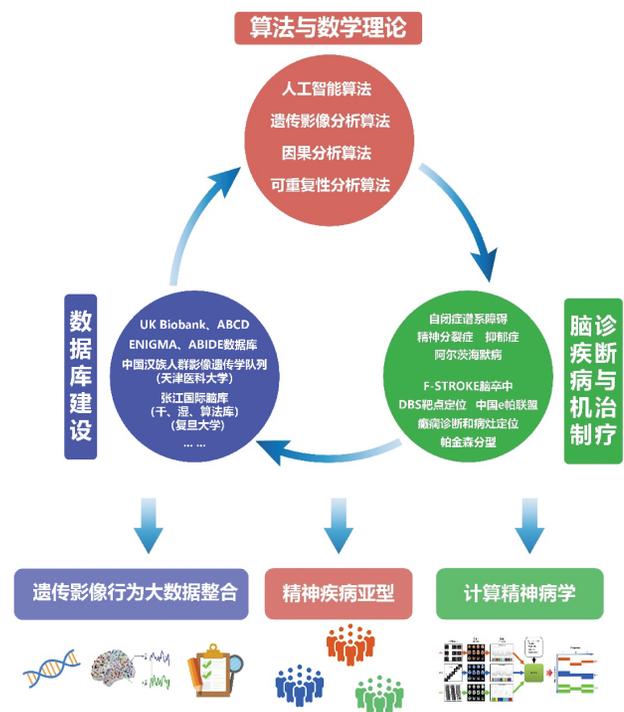


图 1 本文的整体框架

被试间的异质性、机器噪声等造成脑影像数据结果稳健性较差,同时,基于小样本脑影像数据结果的可重复性问题在脑疾病领域中尤为严重<sup>[18-20]</sup>。在此背景下,脑影像大数据应运而生并迅速发展。目前,世界范围内已建立起多个大规模的脑影像数据库,基本涵盖了所有常见的脑疾病。脑影像大数据的建设主要包括两种不同的思路,第一种是在统一的实验设计框架下采集大规模的脑影像数据,代表性数据库如英国生物样本(UK Biobank)数据库(<https://www.ukbiobank.ac.uk>)、美国青少年脑认知发育(Adolescent Brain Cognitive Development, ABCD)数据库(<https://abcdstudy.org>);第二种是基于世界各地已采集好的脑影像数据,通过数据合作共享的模式,对各站点的数据进行整合,形成一个大规模的脑影像数据库,代表性数据库如欧洲的 ENIGMA 数据库(<http://enigma.ini.usc.edu>)、美国的 ABIDE 数据库([http://fcon\\_1000.projects.nitrc.org/indi/abide](http://fcon_1000.projects.nitrc.org/indi/abide))。这些大规模的脑影像数据库对脑功能与脑疾病研究具有重要的推动作用,取得了一系列重要的研究成果<sup>[21-24]</sup>。

UK Biobank 数据库由英国政府于 2006 发起,纳入的被试超过 50 万人,被试年龄分布在 40~69 岁<sup>[25]</sup>。该数据库所采集的信息非常全面,包括血液、唾液、尿液、生活方式及血缘数据等,并将跟踪记录志愿者之后数十年的医疗健康档案信息,获得全基因组遗传数据、临床测量以及健康记录。2014 年,UK Biobank 开始大规模采集这些被试的脑影像数据。目前,该数据库已采集并开放了近 5 万名被试的多模态脑影像数据,并且数据采集仍在持续中,计划在未来几年采集近 10 万名被试,建立世界上最大的脑影像数据库<sup>[26]</sup>。该数据库的优势在于数据样本量大并且数据模态全面,涵盖了遗传、生理指标、环境、生活习惯以及疾病等各个方面,使得我们能够从遗传、环境等各个角度对脑疾病的发生发展展开研究,甚至能够建立脑疾病与其他疾病如高血压、糖尿病之间的联系<sup>[27]</sup>。

ABCD 数据库是美国发起的针对儿童发育的基因、影像及行为数据库。该数据包括近 11 000 多例 9~10 岁的儿童,该数据库的特点是纵向随访的实验设计,该项目计划从基线(9~10 岁)开始对孩子脑影像、行为等进行长达近 10 年的随访<sup>[28]</sup>。ABCD 研究的主要目的是追踪从童年到青春期的脑功能与结构发育过程,进而确定影响或改变脑发育轨迹的生理和环境因素。该数据库的优势在于其基线水

平的被试都是 10 岁左右的孩子,而这一阶段的孩子其实脑功能结构受环境等因素(如吸烟、喝酒等)影响很小,而通过对这些孩子行为及脑影像的纵向追踪,能够使得我们对脑功能结构与行为之间的复杂因果关系进行探究<sup>[29]</sup>。

国际多中心的全球合作研究计划 ENIGMA 是一个全世界来自 43 个国家 50 多个课题组的综合数据库,包括近五万被试的脑影像和基因数据,涵盖了所有常见的脑疾病<sup>[30]</sup>。ENIGMA 研究基本覆盖了所有常见的脑疾病,包括精神分裂症、双向障碍、抑郁症、过急性应激障碍、成瘾、强迫症、多动症、自闭症和癫痫等。ENIGMA 有完善的协同工作制度,各个课题组在 ENIGMA 综合调配下,采用统一标准来完成各自科研任务。这种合作共享的模式能够有效克服单一研究中心被试招募困难的问题,充分将世界各地积累的脑影像数据利用起来,快速建立各种脑疾病的影像数据库。目前,脑影像层面,ENIGMA 以结构磁共振影像为主,通过对结构脑影像数据进行建模,探究脑结构特征与脑疾病之间的关系<sup>[31-33]</sup>。此外,在基因层面,ENIGMA 能够将脑影像特征与基因数据进行关联分析,探究脑结构的遗传基础,并进一步综合分析基因、影像及脑疾病三者之间的复杂关系<sup>[34-36]</sup>。ENIGMA 计划建立到现在的十年期间,脑疾病的遗传影像机制方面取得了一系列重要的科研成果<sup>[21]</sup>。

近几年,国内大规模脑影像数据库建设突飞猛进。由天津医科大学等 30 个研究中心组成的中国影像遗传学团队近期发布了全球最大样本中国汉族人群影像遗传学队列,共收集了 7 000 多名 18~30 岁汉族健康人群的基因、神经影像、环境和行为学数据,旨在研究与影像、行为、基因以及环境四者之间的关系及交互作用<sup>[37]</sup>。由复旦大学类脑智能科学与技术研究院领衔的张江国际脑库目前也正在建设之中,按照“三库(干库、湿库、算法库)融合,多源共享”的建设思路和“全维度、多模态、跨时空”的数据获取模式,干库和算法库由研究院联合多家国内外医疗和科研机构共同建设,湿库由华山医院西院投入建设。到 2023 年,张江国际脑库将围绕精神分裂症、抑郁症、自闭症、脑卒中、精神退行性疾病以及大学生人群等六个队列,采集环境、行为、遗传、脑影像、神经等多尺度数据,在全国建立多个脑科学数据库临床合作点。目前张江国际脑库的六大队列共累计入组 1 900 余人,完成各模态数据采集 6 599 例次,采集工作仍在稳步推进之中。

虽然大规模人群队列的数据库建设取得长足进展,但也面临着巨大挑战。第一,目前大多数大规模人群队列客观上难以实现对被试的长期随访,缺乏纵向的实验设计,以横断面数据为主的研究难以深入刻画不同变量之间复杂的因果关系。第二,数据共享开放程度不足。囿于多数队列的建设归属于不同主体,受不同项目和基金支持,多自成体系,队列建设信息透明度较低,并且不同数据库之间所采集的信息、使用的参数都有较大差异,这给不同数据库之间的整合带来一定的困难。再次,脑疾病的大样本数据库仍比较匮乏。目前大多数针对脑疾病的队列样本量都相对较小,大部分数据库不到上千个被试,而脑疾病尤其是精神疾病具有很强的异质性,只有样本量足够大时,才能够找到可靠的生物学标记。最后,缺少对多模态多尺度数据的整合。目前脑相关的数据库仍以脑影像与行为学信息为主,而缺少对其他模态数据如基因组学、代谢组学、环境等数据的采集。虽然当前脑科学数据库的建设还面临诸多的困难与挑战,但随着脑科学和现代信息技术的迅速发展,基于队列研究建立脑科学样本库已成为脑科学发展的必然趋势。通过整合脑影像学、基因组学、表观组学、蛋白组学、代谢组学等多维度大数据,将有助于我们揭示人脑的工作原理及其脑疾病的发生发展的生理机制。随着大型队列研究在医学研究中的作用日益显现,大型人群队列的建设是人群疾病防治的重要资源,不仅为个体化治疗和精准预防提供科学证据,并且将极大地推动脑科学的新发展方向。

### 3 遗传影像数据挖掘分析的数学理论和方法

针对海量多尺度时空动态脑影像数据,如何发展和应用一套适合脑科学研究规律的非线性数学理论和方法,建立大规模数据挖掘的脑功能因果分析算法,阐明遗传、行为、生理与疾病相关的脑功能因果关系,解析脑功能机制和脑疾病致病机理是目前脑科学研究的新兴方向。

#### 3.1 人工智能算法

随着人工智能的兴起,人工智能算法在脑影像大数据分析中发挥着重要的作用。如脑影像数据分析中的配准问题,传统配准方法难以解决严重的头动问题,而基于人工智能算法的配准方法已经被证明可以用于畸变较大的图像中<sup>[38, 39]</sup>。未来,基于智能算法的配准或将更广泛地应用于脑影像数据的

处理中。此外,借助人工智能算法,我们可以对脑影像的特征进行深度挖掘,如通过非监督训练的智能算法(如 Deep Convolution Generative Adversarial Networks, DCGAN)提取脑影像深层次的信息,这使得对于影像特征更深层次的刻画成为可能,进一步这些特征可以被应用到疾病预测、病灶定位,甚至疾病亚型的发现等临床问题中<sup>[40-43]</sup>。此外,人工智能技术在脑出血判断<sup>[44]</sup>、脑瘤亚型分类<sup>[45]</sup>、神经退行性疾病诊断<sup>[46]</sup>等方面取得了重要进展,甚至能够用来生成各种脑图谱<sup>[47]</sup>。因此,结合行为、影像的大数据与现代人工智能算法,有可能为神经退行性疾病、脑卒中等脑疾病的诊断与康复提供量化、个体化策略,通过量化评估结果对病情进行更合理有效的解读,为临床治疗策略的调整控制和病人的居家管理提供有价值的参考。

#### 3.2 遗传影像分析算法

近几年,遗传影像数据分析成为脑影像领域的前沿与热点。通过整合基因、影像以及行为的数据,使得我们建立脑功能结构特征与行为的关系的同时,能够进一步探究它们的遗传基础及分子机制。由于遗传与脑影像的数据的特征维度都非常高(百万级),整合这两个模态的数据成为当前遗传影像数据分析的难点。复旦大学科研团队在非线性的遗传影像关联分析、随机场理论方面已经进行了一系列的前期工作。如 Tao 等人利用核函数方法,发展了非线性关联分析和非线性格兰杰因果关系的数学理论,并将其运用于研究功能磁共振数据的因果关系<sup>[48]</sup>。基于该方法 Jia 等人发现与纹状体相关的奖励预期行为受到 VPS4A 基因的调控,研究结果发表在 PNAS 杂志上<sup>[49]</sup>。Gong 等人利用半参数回归模型寻找中国精神分裂症人群高易感基因 DISC1 控制的脑结构,然后利用脑结构项上发现的信息作为种子点研究精神分裂症患者与正常人脑功能上的差异。结果发现精神分裂症患者大脑楔前叶的结构和功能发生了很大的改变<sup>[50]</sup>。此外,Ge 等人还率先提出了基于随机场理论的遗传影像数据关联分析算法,并将这一套的理论运用于 ADNI 数据库的阿尔兹海默症数据上<sup>[51]</sup>。遗传与影像数据的联合分析为研究脑功能与脑疾病提供更全面的视角,使得我们能够建立从微观分子到宏观行为表型之间的完整证据链,进而揭示脑疾病的分子影像机制。

#### 3.3 因果分析算法

从海量脑影像数据中对脑区间有向相互作用(因果关系)进行准确推断也是当前脑科学大数据方

法学的前沿与热点。目前探索因果关系的两种主要方法是动态因果模型(Dynamic Causal Modeling, DCM)和格兰杰因果模型(Granger Causality Modeling, GCM)。这两种模型被认为是完全不同的因果关系推断方法,并各自具有优缺点。Ge等人提出了一种新的因果推断方法,兼有格兰杰因果模型的性质和动态因果模型的双线性逼近的特征,从而具有两种传统方法的优点,通过结合DCM的双线性近似特征,他们进一步证明GCM可以扩展成为具有物理约束的模型<sup>[52, 53]</sup>。Luo等人利用时间序列分析方法,对脑功能信息的时变性进行建模,提出时变格兰杰因果关系算法并应用到脑影像数据的分析之中<sup>[54]</sup>。这些数据驱动的因果分析算法为脑影像数据的分析提供了新的思路,使得我们能够对脑区间的相互关系进行更准确细致的刻画。

### 3.4 脑影像数据可重复性问题方法学

随着生物学研究技术的迅速发展,我们经常需要从数量惊人的时空数据中找出变量之间的关联。但是,由于所研究问题的复杂性和数据中大量的噪声,变量之间的关联性通常比较弱。因此,部分研究对过去脑影像结果的可靠性提出了质疑,认为某些功能磁共振成像的结果可能被解释为夸大或虚假,这引起了科学界的激烈争论<sup>[18, 55-57]</sup>。Chen等人将Dice系数引入到脑功能网络指标的可重复性衡量上,并对不同指标研究所需要的样本量进行了计算<sup>[58]</sup>。同样的在基因数据的研究中,不同中心的相同临床类型研究样本的研究结果差异很大,并且只有很少的共同基因<sup>[4]</sup>。综上所述,平息研究可靠性争论的关键问题之一是需要提出一些简单有效的方法来获得给定研究的可重复性。为了克服这个问题,受微阵列数据分析和基因集富集分析的可重复性思想的启发<sup>[59-61]</sup>,复旦大学团队提出采用混合模型方法来定义可重复性评估指标,该方法可以避免传统方法中由阈值选取所带来的不准确性,从而更加精确地估计结果的可重复性以及研究设计中所需要的样本量。

## 4 遗传影像大数据在脑疾病病理机制研究中的应用

随着遗传影像数据的不断积累和一系列方法学上的进步,我们对脑疾病的病理机制有了更深入的理解。近几年,遗传影像大数据在重大脑疾病研究中发挥了重要作用,为疾病的早期诊断、药物研发、临床诊断及个性化治疗等提供了有效的生物标志物

及致病机制的理论支持,取得了一系列重要的成果。

### 4.1 注意缺陷多动障碍

注意缺陷多动障碍(Attention Deficit Hyperactivity Disorder, ADHD)是最常见的一种儿童精神疾病之一,影响着大约10%的儿童<sup>[62]</sup>。基于ENIGMA基因影像大数据, Hoogman等人发现儿童ADHD患者的皮质表面积显著降低,但青少年和成人ADHD患者的皮质表面积与健康对照相比没有显著差异<sup>[63]</sup>, ADHD患者双侧杏仁核、纹状体和海马体积的显著减少,同样地,这种差异在儿童组尤其明显<sup>[64]</sup>。Shen等人基于ABCD青少年脑影像随访数据发现早期多动症症状容易引发后期的睡眠问题,这些睡眠问题所涉及的脑区,主要集中在注意力和觉醒系统,这些脑区的脑灰质萎缩程度与ADHD的严重程度相关,而更严重的ADHD症状会引发后期更多的睡眠问题<sup>[65]</sup>。另外, ADHD患者皮质下不同脑区的脆弱性存在显著差异,这种差异可能与氧化应激、神经发育和自噬途径相关基因在不同脑区的差异表达有关<sup>[66]</sup>。

### 4.2 自闭症谱系障碍

自闭症谱系障碍(Autism Spectrum Disorder, ASD)与其他大多数脑疾病不同,患者不同脑区与对照差异的方向不一致,且在整个年龄段内都是如此。在皮质区域, ASD患者有较大的侧脑室和颅内容积、较大的额叶皮质厚度和较低的颞叶皮质厚度。在皮质下区域, ASD患者苍白球、壳核、杏仁核和伏隔核体积较低。年龄效应分析表明, ASD患者皮质厚度的发育随年龄发生了改变,最大的差异发生在青春期前后,而在皮质下脑区未观察到年龄与ASD的交互作用。总体而言, ASD患者纹状体、额叶皮质和颞叶皮质等与认知和情感相关脑区的形态发生了改变<sup>[67]</sup>。在功能连接方面, Cheng等人基于ABIDE自闭症静息态脑影像数据,提出脑影像大数据的全脑关联分析算法,发现自闭症患者“人脸”、“自我”以及“社交功能”等相关脑区的功能连接显著减弱,揭示了自闭症核心症状社交功能下降的神经机制<sup>[68]</sup>。

### 4.3 抑郁症

ENIGMA-MDD工作组基于1728名重度抑郁(Major Depressive Disorder, MDD)患者和7199名健康对照的皮质下脑形态分析研究发现,复发性MDD的海马体积显著小于健康对照,而初发MDD与对照组相比皮质下各脑区体积无显著差异<sup>[69]</sup>。Schmaal等人基于ENIGMA抑郁症大数据发现成

人 MDD 患者的与情绪调节、奖赏功能、社会认知及执行功能有关的眶额叶皮层、前扣带回和后扣带回、岛叶和颞叶的皮质较薄<sup>[33]</sup>，这些差异在首发患者及成年起病的患者中更为显著，而青少年时期起病的成年抑郁症患者中未发现皮质变薄的改变。这些结果说明，MDD 患者的大脑发生了广泛改变，并表明 MDD 可能以动态的方式影响大脑结构，在不同的年龄段有不同的差异模式<sup>[33]</sup>。因此，有研究提出海马和内侧前额叶体积的减小可以作为抑郁症疾病进展的生物标志<sup>[70]</sup>。Cheng 等人基于来自三个站点上千例样本的脑影像数据发现抑郁症患者中人脑奖赏回路的内侧眶额叶皮层与记忆回路的海马旁、颞叶区功能连接减弱，而涉及人脑惩罚神经环路的外侧前额皮质与“自我”功能相关的楔前叶，语言相关的角回与颞中回连接增强<sup>[70]</sup>。Cheng 等人进一步基于对一万多名样本的影像大数据分析发现，在存在睡眠问题的患者中，外侧眶额皮层，楔叶以及背侧前额叶皮层等脑区的功能连接显著升高，并且这些连接的增强调制着抑郁与睡眠之间的关系<sup>[71]</sup>。

#### 4.4 精神分裂症

Van Erp 和 ENIGMA 的合作者首次报道，精神分裂症 (Schizophrenia, SCZ) 患者 (2 028 例) 与健康对照组 (2 540 例) 相比，海马、杏仁核、丘脑、伏核和颞内容积更小，而苍白球和侧脑室容积更大<sup>[72]</sup>。在扩大的样本 (4 474 名 SCZ vs. 5 098 名对照) 中，他们发现与健康对照组相比，SCZ 患者的皮质较薄，皮质表面积较小，这种差异在额叶和颞叶最为明显<sup>[73]</sup>。此外，SCZ 患者脑的结构改变还与疾病的临床症状相关，Walton 等人发现 SCZ 的阳性症状的严重程度与双侧颞上回的厚度负相关，而 SCZ 阴性症状的严重程度与前额叶的皮质厚度 (特别是左内侧眶额皮质) 呈负相关<sup>[74-76]</sup>。除灰质结构变异外，在分析了 1 963 名精神分裂症患者和 2 359 名健康人的大脑 DTI 图像后，Walton 等发现 SCZ 患者中还存在广泛的 FA 值降低，特别是在放射冠前部 (Anterior Corona Radiata) 和胼胝体 (Corpus Callosum)<sup>[77]</sup>。华西医院龚启勇等人基于 DTI 影像数据得到精神分裂的两个不同亚型，两组亚型在影像学特征及临床症状上都有显著的差别<sup>[78]</sup>。

#### 4.5 帕金森病

基于帕金森 (Parkinson's Disease, PD) 进展标志物研究 (Parkinson's Progression Markers Initiative, PPMI) 收集的脑影像数据，Zeighami 等发现新发帕金森患者的脑萎缩主要发生在中脑、基底

节、基底前脑、内侧颞叶和一些离散的皮质区域<sup>[79]</sup>。随着疾病的发展，帕金森患者出现皮质萎缩，并且皮质萎缩程度与其到基底节的结构连接强度成正比相关，提示 PD 患者脑萎缩是沿着结构连接从基底节传播到皮层脑区<sup>[80]</sup>。这一传播模式可能是错误折叠的  $\alpha$ -突触核蛋白从基底节沿着结构连接网络向外扩散所致<sup>[81]</sup>。此外，帕金森患者的基线时期脑萎缩程度可以预测 PD 患者疾病的发展速率<sup>[82-83]</sup>。Wang 等人基于 PPMI 数据的工作，发现新发 PD 患者皮质下脑区体积大的患者疾病发展速率慢，皮质下脑区体积小的患者疾病发展速率快。这一结果表明帕金森疾病发展速率与患者皮质下脑区体积有关，同时也提示帕金森患者疾病发展速率与患者的脑储备相关联<sup>[84]</sup>。

#### 4.6 阿尔茨海默病

基于阿尔茨海默病 (Alzheimer's Disease, AD) 遗传影像随访数据 (Alzheimer's Disease Neuroimaging Initiative, ADNI)，研究者发现 AD 患者  $\beta$ -淀粉样蛋白 ( $A\beta$ ) 沉积与默认网络功能连接性变化同时发生，随后  $A\beta$  沉积沿功能连接网络和结构连接网络特异性和渐进性地扩散到其他脑区<sup>[85]</sup>。随着疾病的进展和 tau 的积聚增加，内侧颞叶与后内侧区功能连接也随之不断减弱<sup>[86]</sup>。阿尔茨海默病临床前期的早期改变可能涉及 Anterior-Temporal 内连接的减弱，这些连接的减弱可能与记忆障碍有关，但与 AD 的结构变异无关。AD 基底前脑胆碱能投射系统的退化是内嗅和新皮质萎缩的上游事件，并且基底前脑体积可以预测阿尔茨海默病在皮质扩散<sup>[87]</sup>。AD 患者中认知抵抗力和大脑抵抗力较低的人随着疾病的发展认知能力下降快。女性和年轻患者的大脑对病理性 tau 的抵抗力更强，此外，tau 聚集和神经元丢失还与皮质铁积累存在相关关系<sup>[88]</sup>，而较高的教育水平和皮质厚度与认知抵抗力强相关<sup>[89]</sup>。

### 5 遗传影像大数据在脑疾病临床诊断治疗中的应用

近年来，遗传影像大数据分析在脑疾病临床转化和应用中，特别是在医学影像分析、临床决策辅助系统、疾病早期预警、疾病管理以及治疗预后预测等方面展现出重大的潜在应用价值，为推动重大脑疾病精准治疗、智能化健康管理、精准医学等方面的发展发挥重大作用。

遗传影像大数据分析在脑卒中救治、辅助决策及风险预测方面正在发挥重要的临床应用价值。复

旦大学类脑智能科学与技术研究院研发了 F-STROKE 脑卒中急诊期的患者风险分级自动智能化和预后风险预测方法,达到该研究领域国际一流性能,已在国内三十余家医院进行应用。

我国正在快速步入老龄化社会,患有帕金森病等神经退行性疾病的人群数目越来越大,基于影像、行为的大数据分析将有可能为神经退行性疾病的诊断与康复提供量化、个性化策略,通过量化评估结果对病情进行更合理有效的解读,为临床治疗策略的调整控制和病人的居家管理提供有价值的参考。帕金森病全球有超过 500 万患者,其具有高异质性临床病理学亚型以及不同的自然病程,为该疾病早期诊断以及临床药物使用带来极大困难。Fereshtehnejad 等通过融合帕金森病临床运动、认知以及精神状态等行为评分及脑影像、生化指标等,将帕金森疾病分成三个不同的亚型:弥漫性恶性,中度和轻度运动主导型;复旦研究团队从脑结构量化出发,发现帕金森病患者具有两个亚型,其不同的疾病发展进程与脑皮层体积密切相关<sup>[84]</sup>,这些研究为帕金森病的精细划分和精准理疗提供了理论支持<sup>[90]</sup>。另一方面,帕金森病患者分布在全国,而专家和专科门诊均集中在大中城市。为推动帕金森病远程管理新模式,复旦大学附属华山医院神经内科建立了“中国 e 帕联盟”,实现了帕金森病慢病管理和真实世界临床研究为一体,提高了慢病随访的依从性。同时,基于可穿戴设备的行为学定量评估为帕金森病精准管理提供了技术支撑<sup>[91, 92]</sup>。融合脑影像及行为定量评估,建立远程、智能、量化帕金森病慢病长程管理新模式,将提升帕金森病患者的用药管理、康复管理精准度和效率。

脑影像大数据分析在脑疾病功能神经外科治疗中发挥着重要的作用,在脑功能区精准定位、手术路径规划和导航、手术治疗效果预后等方面的临床应用越来越广泛。Wang 等采用影像技术与电生理技术融合的方法,用于癫痫诊断和病灶定位,引导脑深部病灶毁损治疗<sup>[93]</sup>;将术中神经影像与神经电生理结合,可以提升深部脑刺激电极植入手术中功能区特异性定位的准确性<sup>[94]</sup>。利用神经调控技术与脑影像技术的融合,DiMarzio 等研究帕金森病接受 DBS 刺激与否的状态下,慢性疼痛的缓解机制<sup>[95]</sup>。而更为重要的是,利用脑影像分析可以发现深部脑刺激更加有效的治疗靶点<sup>[96]</sup>,甚至发现新的治疗靶点<sup>[97]</sup>,为功能神经外科手术治疗开辟了基于脑网络调控的新方向。

## 6 未来研究方向展望

近二十年脑疾病研究与信息获取技术的发展相融合,在多尺度多模态大数据采集与分析等方面取得了快速发展。未来,大规模的标准化、大样本、多尺度、多模态国际共享数据平台的建设和发展,不仅可为脑疾病和脑科学研究提供更为丰富、高质量的数据基础战略资源,同时可以推动疾病分型、计算神经科学、计算精神病学等新兴研究方向和研究领域的发展。

### 6.1 遗传影像行为大数据的整合

过去十年,遗传影像数据在脑科学研究中扮演着重要的角色,通过整合遗传、影像及行为等数据,能够使得我们以更全面的视角研究人脑的功能以及脑疾病的发生发展机制。但目前遗传、影像以及行为三者高维数据的整合分析方法仍然面临着巨大的挑战。在传统脑影像分析中,随机场理论被广泛的应用于计算多重检验校正的阈值。与传统方法相比,随机场理论对脑影像的空间结构进行建模,有助于提升数据分析的统计功效。实际上,随机场理论对于任何存在结构的数据都是适用的,因此,我们可以对其他模态进行建模。比如,对于基因数据,众所周知,基因是分布在蛋白质上的,蛋白质的空间结构决定了它的功能,因此,在基因脑影像关联分析中,我们可以同时对脑影像和基因的空间结构进行建模,从而提升关联分析的功效。这就是我们所说的二场理论。更进一步,我们还可以对人体行为或表型数据进行建模,比如人脸等图像数据,实现对脑、基因、表型的三方面同时关联分析。这就是我们所说的三场理论,其对生命科学和神经科学的数据驱动发现具有重要意义。

### 6.2 精神疾病亚型

当前,精神疾病的诊断仍缺乏基于客观标志物的指标,临床诊断基本都是基于行为症状的打分<sup>[98-100]</sup>,当患者具有大于一定数目的疾病症状时,就会被诊断为某种精神疾病,因此,不同的病人往往具有不同的症状组合,在个体水平上存在广泛的异质性。如自闭症,由于其在症状<sup>[101]</sup>及致病因素<sup>[102]</sup>都具有广泛异质性,临床将其称为自闭症谱系病,而不是单一疾病。脑成像技术的快速发展为精神疾病的定义、诊断甚至治疗带来了新的契机。如 2017 年 *Nature Medicine* 发表首次利用静息态脑影像特征对抑郁症进行分型文章,该研究得到四类不同的生物亚型并且亚型间具有不同的治疗效果<sup>[97]</sup>。近几

年,越来越多的证据表明基于脑影像学特征对脑疾病进行细化分型的巨大潜力与临床价值<sup>[103, 104]</sup>。未来,随着脑影像数据的不断积累,结合现代的人工智能算法,基于客观生物学指标的疾病亚型将有可能迎来新的突破。这一突破将改变沿用了半个多世纪的 DSM 对精神疾病的定义与分类,对精神疾病的精确诊断以及个性化治疗都有重要意义。

### 6.3 计算精神病学

随着各种高通量技术和无线远程获取技术的迅速发展,为脑科学与脑疾病的研究提供了前所未有的契机。早在 10 年前,美国国家心理健康研究所(National Institute of Mental Health, NIMH)就提出针对的精神疾病 Research Domain Criteria (RDoC) 计划,旨在通过获取海量的多模态多尺度数据,整合基因、代谢、蛋白质、影像及行为等全维度的数据,从基因、分子、影像再到行为等多个水平对精神疾病展开研究,进而全面揭示这些疾病背后的病理机制,促进这些疾病的早期诊断与个性化治疗。除了遗传、影像、行为和环境这四个层次数据之外,目前通过其他现代化手段获取的数据也越来越受到人们的关注。如利用手机,可以无干涉地获取手机用户的夜间浏览情况,从而来预判人的情绪状态,应用于抑郁症的预警;通过手环设备,我们还可以获取用户每天的运动作息情况,从而得到运动量、昼夜节律、睡眠时长等信息,实现对个体生活习惯数据的量化。这些新型数据采集方式能够帮助我们获取海量的个体实验数据,如何利用这些数字化数据促进我们对精神疾病的早期预警与干预逐渐成为当前精神疾病的热点与前沿。因此, NIMH 新主任 Joshua Gordon 则提出了下一个目标是计算精神病学,旨在将之前的 RDoC 计算推向下一阶段,通过整合这些海量的生物数据,在精神疾病的诊断、分型以及个性化治疗方面取得突破,实现与临床医学更好的结合。

### 参 考 文 献

- [1] Liu NH, Daumit GL, Dua T, et al. Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. *World Psychiatry*, 2017, 16(1): 30—40.
- [2] Prince M, Patel V, Shekhar S, et al. Global mental health 1—No health without mental health. *The Lancet*, 2007, 370(9590): 859—877.
- [3] World Health Organization, Mental Health Gap Action Programme. mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings : version 1. 0. Geneva: World Health Organization, 2010.
- [4] James SL, Abate D, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990—2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 2018, 392(10159): 1789—1858.
- [5] World Health Organization. Mental health. (2019-10-02)/[2020-06-21]. <https://www.who.int/news-room/facts-in-pictures/detail/mental-health>.
- [6] World Health Organization. Mental disorders. (2019-11-28)/[2020-06-21]. <https://www.who.int/en/news-room/fact-sheets/detail/mental-disorders>.
- [7] Huang Y, Wang Y, Wang H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *The Lancet Psychiatry*, 2019, 6(3): 211—224.
- [8] World Health Organization. Mental health action plan 2013—2020. Geneva: World Health Organization, 2013.
- [9] World Health Organization. Comprehensive mental health action plan 2013—2030. (2013-06-06)/[2020-06-21]. [https://www.who.int/mental\\_health/action\\_plan\\_2013/en/](https://www.who.int/mental_health/action_plan_2013/en/).
- [10] Prince MJ. World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International, 2015.
- [11] World Health Organization. Global action plan on the public health response to dementia 2017—2025. Geneva: World Health Organization, 2017.
- [12] Jia LF, Quan MN, Fu Y, et al. Dementia in China: epidemiology, clinical management, and research advances. *The Lancet Neurology*, 2020, 19(1): 81—92.
- [13] Jia JP, Wei CB, Chen SQ, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. *Alzheimer's & Dementia*, 2018, 14(4): 483—491.
- [14] World Health Organization. Dementia. (2020-09-21)/[2021-01-28]. <https://www.who.int/news-room/facts-in-pictures/detail/dementia>.
- [15] Zhao BX, Luo TY, Li TF, et al. Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nature genetics*, 2019, 51(11): 1637—1644.
- [16] Shen EH, Overly CC, Jones AR. The Allen Human Brain Atlas: comprehensive gene expression mapping of the human brain. *Trends in neurosciences*, 2012, 35(12): 711—714.
- [17] Li A, Gong H, Zhang B, et al. Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain. *Science*, 2010, 330(6009): 1404—1408.

- [18] Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, 2016, 113 (33): E4929—E4929.
- [19] Poldrack RA, Baker CI, Durnez J, et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, 2017, 18 (2): 115—126.
- [20] Button KS, Loannidis JPA, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 2013, 14(6): 444—444.
- [21] Thompson PM, Jahanshad N, Ching CRK, et al. ENIGMA and global neuroscience: a decade of large-scale studies of the brain in health and disease across more than 40 countries. *Translational Psychiatry*, 2020, 10(1): 1—28.
- [22] Janiri D, Doucet GE, Pompili M, et al. Risk and protective factors for childhood suicidality: a US population-based study. *The Lancet Psychiatry*, 2020, 7(4): 317—326.
- [23] Cortes A, Albers PK, Dendrou CA, et al. Identifying cross-disease components of genetic risk across hospital data in the UK Biobank. *Nature Genetics*, 2020, 52(1): 126—134.
- [24] Martino DA, Yan CG, Li Q, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 2014, 19(6): 659—667.
- [25] Palmer LJ. UK Biobank: bank on it. *The Lancet*, 2007, 369(9578): 1980—1982.
- [26] Miller KL, Alfaro-Almagro F, Bangerter NK, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience*, 2016, 19(11): 1523—1536.
- [27] Collins R. What makes UK Biobank special?. *The Lancet*, 2012, 379(9822): 1173—1174.
- [28] Volkow ND, Koob GF, Croyle RT, et al. The conception of the ABCD study: from substance use to a broad NIH collaboration. *Developmental Cognitive Neuroscience*, 2018, 32: 4—7.
- [29] Casey BJ, Cannonier T, Conley MI, et al. The Adolescent Brain Cognitive Development (ABCD) study: imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, 2018, 32: 43—54.
- [30] Thompson PM, Stein JL, Medland SE, et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior*, 2014, 8(2): 153—182.
- [31] Kong XZ, Boedhoe PSW, Abe Y, et al. Mapping cortical and subcortical asymmetry in obsessive-compulsive disorder: findings from the enigma consortium. *Biological Psychiatry*, 2019, 87(12): 1022—1034.
- [32] Mackey S, Allgaier N, Chaarani B, et al. Mega-Analysis of gray matter volume in substance dependence: general and substance-specific regional effects. *American Journal of Psychiatry*, 2019, 176(5): 414—414.
- [33] Schmaal L, Hibar DP, Sämann PG, et al. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Molecular Psychiatry*, 2017, 22(6): 900—909.
- [34] Grasby KL, Jahanshad N, Painter JN, et al. The genetic architecture of the human cerebral cortex. *Science*, 2020, 367(6484): 6690.
- [35] Hibar DP, Stein JL, Renteria ME, et al. Common genetic variants influence human subcortical brain structures. *Nature*, 2015, 520(7546): 224—229.
- [36] Jia T, Chu CY, Liu Y, et al. Epigenome-wide meta-analysis of blood DNA methylation and its association with subcortical volumes: findings from the ENIGMA Epigenetics Working Group. *Molecular Psychiatry*, 2019: 1—12.
- [37] Xu Q, Guo LN, Cheng JL, et al. CHIMGEN: a Chinese imaging genetics cohort to enhance cross-ethnic and cross-geographic brain research. *Molecular Psychiatry*, 2020, 25 (3): 517—529.
- [38] Eppenhof K, Lafarge MW, Veta M, et al. Progressively trained convolutional neural networks for deformable image registration. *IEEE Transactions on Medical Imaging*, 2020, 39(5): 1594—1604.
- [39] Krebs J, Delingette H, Mailhé B, et al. Learning a probabilistic model for diffeomorphic registration. *IEEE Transactions on Medical Imaging*, 2019, 38 (9): 2165—2176.
- [40] Hsu CC, Lin CW. CNN-Based Joint Clustering and Representation Learning with Feature Drift Compensation for Large-Scale Image Data. *Ieee Transactions on Multimedia*, 2018, 20(2): 421—429.
- [41] Su WT, Hsu CC, Huang ZL, et al. Joint pairwise learning and image clustering based on a siamese cnn. 2018 25th IEEE International Conference on Image Processing. Greece: IEEE, 2018.
- [42] Hollon TC, Pandian B, Adapa AR, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nature Medicine*, 2020, 26(1): 52—58.
- [43] Gill RS, Caldairou B, Bernasconi N, et al. Uncertainty-Informed detection of epileptogenic brain malformations using bayesian neural networks. //International Conference on Medical Image Computing and Computer-Assisted Intervention. Switzerland; Springer, Cham, 2019.

- [44] Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *NPJ Digital Medicine*, 2018, 1: 9.
- [45] Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature*, 2018, 555(7697): 469—474.
- [46] Mutlu U, Colijn JM, Arfan I, et al. Association of retinal neurodegeneration on optical coherence tomography with dementia a population-based study. *JAMA Neurology*, 2018, 75(10): 1256—1263.
- [47] Robie AA, Hirokawa J, Edwards AW, et al. Mapping the Neural Substrates of Behavior. *Cell*, 2017, 170 ( 2 ): 393—406.
- [48] Tao CY, Feng JF. Nonlinear association criterion, nonlinear Granger causality and related issues with applications to neuroimage studies. *Journal of Neuroscience Methods*, 2016, 262: 110—132.
- [49] Jia TY, Macare C, Desrivières S, et al. Neural basis of reward anticipation and its genetic determinants. *Proceedings of the National Academy of Sciences of the United States of America*, 2016, 113(14): 3879—3884.
- [50] Gong XH, Lu WL, Kendrick KM, et al. A Brain-wide association study of DISC1 genetic variants reveals a relationship with the structure and functional connectivity of the precuneus in schizophrenia. *Human Brain Mapping*, 2014, 35(11): 5414—5430.
- [51] Ge T, Feng JF, Hibar DP, et al. Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures. *Neuroimage*, 2012, 63(2): 858—873.
- [52] Ge T, Kendrick KM, Feng JF. A novel extended granger causal model approach demonstrates brain hemispheric differences during face recognition learning. *PloS Computational Biology*, 2009, 5(11).
- [53] Ge T, Feng JF, Grabenhorst F, et al. Componential granger causality, and its application to identifying the source and mechanisms of the top-down biased activation that controls attention to affective vs sensory processing. *Neuroimage*, 2012, 59(2): 1846—1858.
- [54] Luo Q, Lu W, Cheng W, et al. Spatio-temporal granger causality: a new framework. *Neuroimage*, 2013, 79: 241—263.
- [55] Bennett CM, Miller MB. How reliable are the results from functional magnetic resonance imaging?. *The Year in Cognitive Neuroscience 2010*, 2010, 1191: 133—155.
- [56] Fostering reproducible fMRI research. *Nature Neuroscience*, 2017, 20(3): 298.
- [57] Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103 ( 15 ): 5923—5928.
- [58] Chen X, Lu B, Yan CG. Reproducibility of R-fMRI metrics on the impact of different strategies for multiple comparison correction and sample sizes. *Human Brain Mapping*, 2018, 39(1): 300—318.
- [59] Lai Y, Adam BL, Podolsky R, et al. A mixture model approach to the tests of concordance and discordance between two large-scale experiments with two-sample groups. *Bioinformatics*, 2007, 23(10): 1243—1250.
- [60] Lai YL, Zhang FN, Nayak TK, et al. Concordant integrative gene set enrichment analysis of multiple large-scale two-sample expression data sets. *BMC Genomics*, 2014, 1(15): S6.
- [61] Lai YL, Zhang FN, Nayak TK, et al. An efficient concordant integrative analysis of multiple large-scale two-sample expression data sets. *Bioinformatics*, 2017, 33(23): 3852—3860.
- [62] Xu GF, Strathearn L, Liu BY, et al. Twenty-year trends in diagnosed attention-deficit/hyperactivity disorder among US children and adolescents, 1997—2016. *JAMA Network Open*, 2018, 1(4): e181471.
- [63] Hoogman M, Muetzel RL, Guimarães JP, et al. Brain imaging of the cortex in ADHD: a coordinated analysis of large-scale clinical and population-based samples. *American Journal of Psychiatry*, 2019, 176(7): 531—542.
- [64] Hoogman DM, Bralten J, Hibar DP, et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults; a cross-sectional mega-analysis. *Lancet Psychiatry*, 2017, 4 ( 4 ): 310—319.
- [65] Shen C, Luo Q, Chamberlain SR, et al. What is the link between attention-deficit/hyperactivity disorder and sleep disturbance? A multimodal examination of longitudinal relationships and brain structure using large-scale population-based cohorts. *Biological Psychiatry*, 2020, 88 ( 6 ): 459—469.
- [66] Hess JL, Akutagawa-Martins GC, Patak JD, et al. Why is there selective subcortical vulnerability in ADHD? Clues from postmortem brain gene expression data. *Molecular Psychiatry*, 2018, 23(8): 1787—1793.
- [67] van Rooij D, Anagnostou E, Arango C, et al. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: results from the ENIGMA ASD working group. *American Journal of Psychiatry*, 2018, 175 ( 4 ): 359—369.

- [68] Cheng W, Rolls ET, Zhang J, et al. Autism: reduced connectivity between cortical areas involved in face expression, theory of mind, and the sense of self. *Brain*, 2015, 138: 1382—1393.
- [69] Schmaal L, Rose M, EMDDW Group. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. 2016, 21(6): 806—812.
- [70] Belleau EL, Treadway MT, Pizzagalli DA. The impact of stress and major depressive disorder on hippocampal and medial prefrontal cortex morphology. *Biological Psychiatry*, 2019, 85(6): 443—453.
- [71] Cheng W, Rolls ET, Ruan HT, et al. Functional connectivities in the brain that mediate the association between depressive problems and sleep quality. *JAMA Psychiatry*, 2018, 75(10): 1052—1061.
- [72] van Erp TGM, Hibar DP, Rasmussen JM, et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Molecular Psychiatry*, 2016, 21(4): 585—585.
- [73] van Erp TGM, Walton E, Hibar DP, et al. Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the enhancing neuro imaging genetics through meta analysis (ENIGMA) consortium. *Biological Psychiatry*, 2018, 84(9): 644—654.
- [74] Alnaes D, Kaufmann T, van der Meer D. Brain heterogeneity in schizophrenia and its association with polygenic risk. *JAMA Psychiatry*, 2019, 76(9): 986—986.
- [75] Walton E, Hibar DP, van Erp TGM, et al. Positive symptoms associate with cortical thinning in the superior temporal gyrus via the ENIGMA Schizophrenia consortium. *Acta Psychiatrica Scandinavica*, 2017, 135(5): 439—447.
- [76] Walton E, Hibar DP, van Erp TGM, et al. Prefrontal cortical thinning links to negative symptoms in schizophrenia via the ENIGMA consortium. *Psychological Medicine*, 2018, 48(1): 82—94.
- [77] Kelly S, Jahanshad N, Zalesky A, et al. Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the ENIGMA Schizophrenia DTI Working Group. *Molecular Psychiatry*, 2018, 23(5): 1261—1269.
- [78] Sun HQ, Liu S, Yao L, et al. Two patterns of white matter abnormalities in medication-naïve patients with First-episode schizophrenia revealed by diffusion tensor imaging and cluster analysis. *JAMA Psychiatry*, 2015, 72(7): 678—686.
- [79] Zeighami Y, Ulla M, Iturria-Medina Y, et al. Network structure of brain atrophy in de novo Parkinson's disease. *Elife*, 2015, 4.
- [80] Yau YHC, Zeighami Y, Baker T, et al. Network connectivity determines cortical thinning in early Parkinson's disease progression. *Nature Communications*, 2018, 9(1): 1—10.
- [81] Henderson MX, Cornblath EJ, Darwich A, et al. Spread of alpha-synuclein pathology through the brain connectome is modulated by selective vulnerability and predicted by network analysis. *Nature Neuroscience*, 2019, 22(8): 1248—1257.
- [82] Zeighami Y, Fereshtehnejad SM, Dadar M, et al. A clinical-anatomical signature of Parkinson's disease identified with partial least squares and magnetic resonance imaging. *Neuroimage*, 2019, 190: 69—78.
- [83] Zeighami Y, Fereshtehnejad SM, Dadar M, et al. Assessment of a prognostic MRI biomarker in early de novo Parkinson's disease. *Neuroimage: Clinical*, 2019, 24: 101986.
- [84] Wang L, et al. Two biotypes in Parkinson's disease with distinctive longitudinal progression. *Neurology*, 2020, 95(11).
- [85] Lee ES, Yoo K, Lee YB, et al. Default mode network functional connectivity in early and late mild cognitive impairment results from the Alzheimer's disease neuroimaging initiative. *Alzheimer Disease & Associated Disorders*, 2016, 30(4): 289—296.
- [86] Berron D, van Westen D, Ossenkoppele R, et al. Medial temporal lobe connectivity and its associations with cognition in early Alzheimer's disease. *Brain*, 2020, 143(4): 1233—1248.
- [87] Fernandez-Cabello S, Kronbichler M, van Dijk KRA, et al. Basal forebrain volume reliably predicts the cortical spread of Alzheimer's degeneration. *Brain*, 2020, 143(3): 993—1009.
- [88] Spotorno N, Acosta-Cabronero J, Stomrud E, et al. Relationship between cortical iron and tau aggregation in Alzheimer's disease. *Brain*, 2020, 143(5): 1341—1349.
- [89] Ossenkoppele R, Lyoo CH, Jester-Broms J, et al. Assessment of demographic, genetic, and imaging variables associated with brain resilience and cognitive resilience to pathological tau in patients with Alzheimer disease. *JAMA Neurology*, 2020, 77(5): 632—642.
- [90] Fereshtehnejad SM, Zeighami M, Dagher A, et al. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. *Brain*, 2017, 140: 1959—1976.
- [91] Wang JY, Gong DW, Luo HC, et al. Measurement of step angle for quantifying the gait impairment of Parkinson's disease by wearable sensors: Controlled study. *Jmir Mhealth and Uhealth*, 2020, 8(3).
- [92] 李亮, 俞乾, 徐宝腾, 等. 帕金森病患者多节点运动监测可穿戴设备. *生物医学工程学杂志*, 2016(6): 165—172.

- [93] Wang X, Hu WH, McGonigal A, et al. Electroclinical features of insulo—opercular epilepsy: an SEEG and PET study. *Annals of Clinical and Translational Neurology*, 2019, 6(7): 1165—1177.
- [94] Geng XY, Xu X, Horn A, et al. Intra-operative characterisation of subthalamic oscillations in Parkinson's disease. *Clinical Neurophysiology*, 2018, 129(5): 1001—1010.
- [95] DiMarzio M, Rashid T, Hancu I, et al. Functional MRI signature of chronic pain relief from deep brain stimulation in parkinson disease patients. *Neurosurgery*, 2019, 85(6): E1043—E1049.
- [96] Horn A, Neumann WJ, Degen K, et al. Toward an electrophysiological “sweet spot” for deep brain stimulation in the subthalamic nucleus. *Human Brain Mapping*, 2017, 38(7): 3377—3390.
- [97] Drysdale AT, Grosenick L, Downar J, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, 2017, 23(1): 28.
- [98] Hyman SE. The diagnosis of mental disorders: the problem of reification. *Annual Review of Clinical Psychology*, 2010, 6: 155—179.
- [99] Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, 2013, 11(1): 126.
- [100] Casey BJ, Craddock N, Cuthbert BN, et al. DSM-5 and RDoC: progress in psychiatry research?. *Nature Reviews Neuroscience*, 2013, 14(11): 810.
- [101] Lenroot RK, Yeung PK. Heterogeneity within autism spectrum disorders: What have we learned from neuroimaging studies?. *Frontiers in Human Neuroscience*, 2013, 7: 733.
- [102] Betancur C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Research*, 2011, 1380: 42—77.
- [103] Chand GB, Dwyer DB, Guray E, et al. Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain*, 2020, 143(3): 1027—1038.
- [104] Hong SJ, Valk SL, Martino AD, et al. Multidimensional neuroanatomical subtyping of autism spectrum disorder. *Cerebral Cortex*, 2018, 28(10): 3578—3588.

## Exploring Brain Diseases with Multi-scale, Multi-modality Data and Computational Approach

Cheng Wei    Wang Shouyan    Xie Xiaohua    Feng Jianfeng\*

*Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, Shanghai 200433*

**Abstract** It is challenging in diagnosis and treatment of brain diseases across the whole world, which places heavy burden on individuals, families and society. The rapid development of various high-throughput and remote healthcare technologies provides unprecedented opportunities for advancing the research of brain sciences and brain diseases. The multiscale and multimodality data of the brain structure and functions enables us to explore the mechanisms of brain diseases and develop novel approaches for precision diagnosis and treatment. In this paper, we firstly reviewed the progress and challenges in developing genetic-imaging database in China and worldwide. Furthermore, we discussed current mathematical theories and methods for analyzing genetic imaging data, as well as the translational research in application of genetic imaging big data analysis to revealing the pathological mechanisms of brain diseases and assisting clinical diagnosis and treatment. Finally, we looked into perspectives on future research directions in multiscale and multimodality big data, including the phenotypes of brain diseases, computational neuroscience, and computational psychiatry.

**Keywords** brain diseases; magnetic resonance imaging; genetic imaging; artificial intelligence

(责任编辑 姜钧译 仲斌演)

\* Corresponding Author, Email: jianfeng64@gmail.com