

基于主题 - 情感融合的帕罗西汀治疗孤独症评论数据研究*

吕艳华 崔云龙 王康龙 钟小云 陈家珠

(山西医科大学管理学院 太原 030600)

〔摘要〕 目的/意义 挖掘网络用户关于帕罗西汀治疗孤独症的相关评论,揭示患者的关注焦点和情感态度。方法/过程 利用双词主题模型和情感分析方法,分析多个媒体平台关于帕罗西汀治疗孤独症的评论文本,挖掘文本的主题分布和情感倾向。结果/结论 从 5 个方面分析患者对药物的关注焦点和情感态度,为孤独症患者及相关医务工作者提供有价值的信息,为更好地满足患者需求、改善其生活质量提供参考依据。

〔关键词〕 双词主题模型;孤独症;帕罗西汀;评论数据;情感分析

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2025.01.005

Study on Comment Data of Paroxetine Treatment of Autism Based on Theme - affective Fusion

LYU Yanhua, CUI Yunlong, WANG Kanglong, ZHONG Xiaoyun, CHEN Jiazhu

School of Management, Shanxi Medical University, Taiyuan 030600, China

〔Abstract〕 **Purpose/Significance** The comments of internet users on paroxetine treatment of autism are mined to reveal the patients' focus of attention and emotional attitude. **Method/Process** By using the biterm topic model (BTM) and sentiment analysis method, the paper analyzes the comments about paroxetine treatment of autism on multiple media platforms, and explores the topic distribution and affective tendency of the texts. **Result/Conclusion** The focus of attention and emotional attitude of patients to drugs are analyzed from 5 aspects, which provides valuable information for autistic patients and related medical staffs, and provides references for better meeting the needs of patients and improving their quality of life.

〔Keywords〕 biterm topic model (BTM); autism; paroxetine; comment data; sentiment analysis

1 引言

孤独症谱系障碍 (autism spectrum disorder, ASD), 简称孤独症, 是一种由生物性神经系统发育障碍引发的疾病, 其主要症状包括社会沟通能力障碍、兴趣狭窄及重复行为^[1]。目前, 关于孤独症的

发病机制尚未达成共识, 治疗方法仍处于研究阶段, 专门治疗孤独症的药物相对偏少^[2]。在缓解孤独症症状方面, 常用的治疗药物包括抗抑郁药、抗精神病药、镇静药等^[3]。帕罗西汀作为一种 5-羟色胺 (5-HT) 再摄取抑制剂, 主要通过影响突触前膜, 选择性抑制 5-HT 的再摄取, 从而提高血清及脑脊液中 5-HT 的浓度, 增强神经冲动传导, 从

〔修回日期〕 2024-08-18

〔作者简介〕 吕艳华, 副教授, 博士生导师, 发表论文 50 余篇。

〔基金项目〕 国家社会科学基金一般项目 (项目编号: 20BTQ064)。

而改善患者的抑郁和焦虑状态^[4]。帕罗西汀已在国内获得正式批准，广泛用于抑郁症、强迫症等心理疾病的治疗^[5]。同时，帕罗西汀在缓解孤独症患者的焦虑、抑郁以及强迫症等方面也有突出表现，成为医生在开具药物时的一种常见选择。

本文通过分析多个媒体平台用户关于帕罗西汀治疗孤独症的评论文本，利用双词主题模型（biterm topic model, BTM）和情感分析技术，探讨孤独症患者在接受帕罗西汀治疗时所关注的焦点及情感倾向。期望为孤独症患者及相关医务工作者在帕罗西汀药物使用方面提供更多有价值的参考。

2 数据与方法

2.1 数据收集与处理

2.1.1 数据来源 以“帕罗西汀治疗孤独症（或自闭症、广泛性发展障碍、孤独症谱系障碍）”“帕罗西汀的功效与作用”“帕罗西汀副作用和危害”为搜索句，在抖音、微博、小红书、知乎 4 个媒体平台进行检索。采用爬虫程序抓取各平台用户评论文本，分别获取 5 157 条、4 760 条、2 409 条、584 条数据，总计 12 910 条。

2.1.2 数据处理 对评论文本去重，删除只含有“@ * * *”或文本长度小于 2 字符的无效评论，处理后各平台有效文本数量分别为 4 833 条、4 544 条、2 177 条、566 条，总计 12 120 条。为了提高主题模型精准度，去除对用户主题无实质性意义的符号和词语。采用正则表达式去除非中文字符、“@ * * *”以及表情包，借助自定义用户词典和停用词表，使用 jieba 工具进行分词。在情感分析部分，通过 Python 调用百度智能云的情感倾向分析应用程序编程接口（application programming interface, API），其内置预训练情感分析模型，能够自动识别评论文本的情感倾向，无须对文本进行情感分类或标注。

2.2 研究方法

2.2.1 BTM 目前文本主题模型中应用最广的是隐含狄利克雷分布主题模型^[6]。但互联网评论文本多以短文本形式呈现，如使用传统主题模型，同一

语句中的主题间缺乏关联性，单词在文档中的共现频次相对稀疏^[7]。Yan X 等^[8]针对传统主题模型存在的文本稀疏问题，提出 BTM。在使用 BTM 聚类时，将每个文档视为独立的上下文单元，并从中抽取任意两个词语构成词对^[9]。例如，“帕罗西汀治疗孤独症”这个短文本共有 3 个词对：“帕罗西汀”与“治疗”，“帕罗西汀”与“孤独症”，“治疗”与“孤独症”。

2.2.2 情感分析模型 情感分析技术主要分为基于词典与规则的多元分类方法、机器学习分类方法、基于词典和算法融合的方法以及深度学习方法 4 大类^[10]。百度智能云情感倾向分析 API 内置基于双向长短期记忆网络（bidirectional long short-term memory, BiLSTM）的情感分类模型，能够深入语义层面，精准分析语义情感的强弱程度、语言的复杂结构^[11]。所爬取数据均符合该工具参数要求，确保了使用的有效性。具体技术实现方法如下。（1）准备请求参数。将待分析的文本作为请求参数，设置 API Key 和 Secret Key。（2）发送 HTTP 请求。通过 HTTP POST 将请求参数发送到情感倾向分析 API 的 URL。（3）处理响应结果。API 将返回一个 JSON 格式的响应结果，包括情感倾向和置信度。

3 结果

3.1 评论数据的主题分析结果

3.1.1 主题提取（图 1）

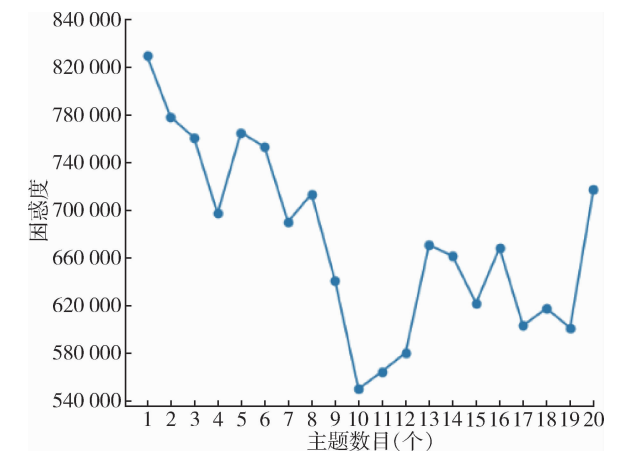


图 1 主题数 - 困惑度折线

BTM 基于困惑度确定最优主题数量 K 。计算 $K \in [1, 20]$ 不同主题数量下的困惑度。当主题数量为 10 时, 困惑度折线为最低点, 因此将主题数量 K 设为 10。使用 `dispTopics` 函数读取每个主题下单词的权重, 输出权重最高的前 10 个主题词 (权重值保留 4 位小数), 见表 1。

表 1 主题 – 主题词分布结果

主题	主题词 (权重值)
主题 1	帕罗西汀 (0.040 3)、医生 (0.030 5)、治疗 (0.022 5)、症状 (0.019 1)、焦虑 (0.018 7)、感觉 (0.015 2)、药物 (0.014 4)、副作用 (0.013 1)、患者 (0.012 1)、失眠 (0.011 1)
主题 2	治疗 (0.078 7)、抑郁症 (0.056 1)、药物 (0.041 6)、患者 (0.036 6)、抗抑郁 (0.027 2)、抑郁 (0.022 6)、心理治疗 (0.016 8)、抑制剂 (0.011 2)、西酞普兰 (0.010 6)、氟西汀 (0.010 6)
主题 3	药物 (0.083 2)、治疗 (0.057 4)、抑制剂 (0.033 2)、抗抑郁 (0.030 8)、患者 (0.030 5)、服用 (0.027 1)、类药物 (0.016 7)、症状 (0.016 7)、抑郁症 (0.016 2)、有效 (0.015 7)
主题 4	焦虑 (0.088 3)、抑郁 (0.037 9)、情绪 (0.030 5)、生活 (0.024 1)、事情 (0.022 8)、感觉 (0.021 7)、工作 (0.019 1)、孩子 (0.017 3)、医生 (0.017 2)、焦虑症 (0.016 6)
主题 5	医生 (0.049 5)、帕罗西汀 (0.026 1)、感觉 (0.023 2)、吃药 (0.019 7)、医院 (0.019 4)、副作用 (0.017 9)、难受 (0.017 8)、焦虑症 (0.017 7)、发作 (0.017 1)、停了 (0.015 1)
主题 6	药物 (0.024 9)、服用 (0.018 5)、浓度 (0.018 1)、不良反应 (0.016 5)、帕罗西汀 (0.015 3)、用药 (0.013 7)、抗抑郁 (0.012 3)、系统 (0.012 1)、氯米帕明 (0.011 7)、本品 (0.011 2)
主题 7	一盒 (0.068 1)、帕罗西汀 (0.043 7)、抑郁 (0.021 5)、氟西汀 (0.019 2)、安眠药 (0.016 2)、医生 (0.012 3)、氟伏沙明 (0.012 1)、喜太乐 (0.011 8)、副作用 (0.011 8)、舍曲林 (0.011 4)
主题 8	治疗 (0.047 1)、生活 (0.028 4)、帕罗西汀 (0.019 2)、患者 (0.015 6)、抑郁症 (0.014 7)、抑郁 (0.014 4)、时间 (0.013 9)、药物 (0.012 6)、影响 (0.012 5)、情绪 (0.012 2)
主题 9	帕罗西汀 (0.036 7)、噩梦 (0.030 6)、进口 (0.016 1)、国产 (0.014 8)、治疗 (0.013 6)、医生 (0.012 5)、药物 (0.011 5)、说明书 (0.010 1)、感觉 (0.009 6)、舍曲林 (0.008 6)
主题 10	帕罗西汀 (0.048 3)、感觉 (0.048 1)、副作用 (0.039 9)、医生 (0.027 3)、难受 (0.025 1)、吃药 (0.022 4)、晚上 (0.022 3)、停药 (0.019 2)、恶心 (0.017 5)、头晕 (0.016 6)

3.1.2 主题名称及内容分析 通过人工识别与分析, 确定 10 个不同主题的名称并概括其主要内容。主题 1: 帕罗西汀的治疗作用与患者感受。帕罗西汀作为常用抗抑郁药物, 在缓解孤独症患者焦虑抑郁情绪方面表现出色。患者在服用帕罗西汀后, 出现副作用的概率较高, 因此应及时关注自身身体状况。主题 2: 抑郁症的治疗方法 with 药物。该主题主要包括药物治疗和心理治疗两个方面。常用的抗抑郁药物包括抑制剂类, 包括帕罗西汀、西酞普兰和舍曲林等。在心理治疗方面, 常用的方法有认知行为疗法、支持性心理治疗等。主题 3: 抗抑郁药物的治疗有效性。该主题包括药物对疾病的治疗效果。多数患者在服用药物后, 能够明显感觉到症状的缓解。少部分患者可能由于病情严重或服用剂量

不足, 效果较差。主题 4: 焦虑与抑郁情绪对生活的影响。该主题主要涉及焦虑与抑郁情绪对生活、工作、孩子教育等方面的影响。由于疾病的原因, 患者心情低落, 拒绝与家人朋友交流, 工作中注意力不集中, 效率低下, 且可能无暇顾及孩子的日常成长。主题 5: 帕罗西汀的副作用与患者体验。该主题展示了患者服用帕罗西汀后所产生的副作用以及停药反应。由于药物本身的作用和个人体质的差异, 服用药物后大概率会产生不同程度、类型的副作用。在药物产生不良反应时, 应避免突然停药, 因为停药所带来的戒断反应可能会进一步加重患者症状。主题 6: 抗抑郁药物的用药剂量与不良反应。该主题主要探讨不同药物浓度所带来的反应。孤独症患者的病情和严重程度不同, 会影响药物的服用

浓度。医生应及时告知患者增大或减小药物的服用量,以确保达到较好的治疗效果。主题 7:多种抗抑郁药物及联合用药。该主题主要涵盖多种药物的联合使用情况,如氟西汀、氟伏沙明和舍曲林等药物,均是缓解抑郁焦虑情绪的药物。一些患者在单独服用帕罗西汀时效果不佳,应根据医生建议切换其他药物或联合用药。主题 8:帕罗西汀对抑郁症患者的影响。该主题聚焦于患者服用药物后对生活、情绪等方面的影响。在生活方面,帕罗西汀通过改善患者病情,扩大了患者的兴趣范围。在情绪方面,服用帕罗西汀能够对患者情绪产生积极影响。在时间方面,疾病的严重程度和药物剂量都会影响药物的起效时间。主题 9:进口与国产帕罗西汀的差异。帕罗西汀在市场上存在进口和国产两种版本。有研究^[12]对比了乐友牌国产药和赛乐特牌进口药,发现两种治疗方案的疗效接近,但国产药物在价格方面更为亲民。医生应根据患者意愿,在确保治疗效果的前提下,减轻患者的经济负担。主题 10:帕罗西汀的副作用与停药反应。帕罗西汀常见的不良反应包括胃肠道反应、食欲减退、体重增加等^[13],这些副作用的强度和持续时间因人而异,但通常会在服药后的几天到几周内逐渐减轻。患者切勿私自停药,避免突然停药反应对身体的损害。

3.2 评论文本的情感分析结果

- 3.2.1 总体情感分析 利用情感倾向分析 API,计算每个评论数据的情感类别以及置信系数,将评论数据归为 3 类,“2”为积极情感、“1”为中性情感、“0”为消极情感。其中消极情感 7 265 条 (59.9%),中性情感 504 条 (4.2%),积极情感 4 351 条 (35.9%)。
- 3.2.2 各主题情感分析 运用 BTM 并结合人工标注,确定每条评论数据所属主题,对各主题情感结果进行量化分析,见图 2。各主题情感分布呈现显著不平衡状态,其中积极情感占比为 25%~50%;中性情感占比为 3%~7%;消极情感占比高达 45%~70%。

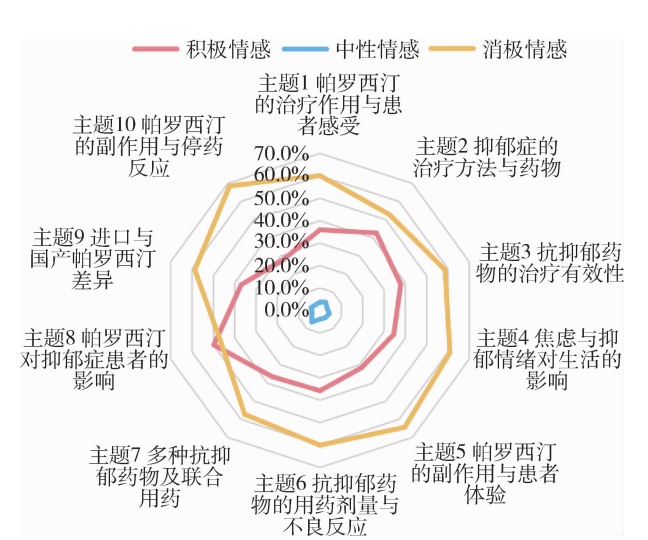


图 2 各主题情感类别分布

4 讨论

在上述 10 个主题基础上,从患者视角将主题整合为以下 5 类,讨论各类主题内容和情感倾向。

4.1 帕罗西汀的治疗效果

该类主题包括主题 1 和主题 3,平均积极情感占比为 36.95%,表明帕罗西汀对部分患者确实产生了积极的治疗效果,这些患者可能体验到了症状的缓解和情绪的改善。然而,平均消极情感占比 59.35%,表明患者出现了副作用或治疗效果不佳等问题,对患者的心理状态和日常生活产生了不利影响。

4.2 进口药物与国产药物

该类主题主要涉及主题 9。积极情感占比 37%,消极情感占比 58.7%。国产帕罗西汀与进口帕罗西汀在成分和功效上差异不大,主要区别在于价格、生产工艺和辅料。患者若购买价格昂贵的进口药物,但治疗效果并不优于国产药物,可能会产生消极情绪。同时,若患者购买价格低廉的国产药物,服用后产生了较大的不良反应,同样会导致情绪不佳。

4.3 抗抑郁药物的治疗与应用

该类主题包括主题 2 和主题 7, 平均积极情感占比 39.6%, 平均消极情感占比 55%。目前常用的抗抑郁药物包括劳拉西泮、奥沙西泮、帕罗西汀等; 镇静药物包括艾司唑仑、地西泮等; 中成药包括解郁安神颗粒、九味镇心颗粒、舒肝解郁胶囊等。由于患者体质或药物本身原因, 药物可能需一段时间才能起效, 或在起效后带来不同不良反应, 这些都可能增加患者的消极情绪。

4.4 帕罗西汀的不良反应与停药反应

该类主题主要包括主题 5、主题 6 和主题 10, 平均积极情感占比 31.8%, 平均消极情感占比 64.3%。消极情感在 5 类主题中占比最高, 反映患者在服用药物后, 所产生的不良反应较多, 且停药过程中也常遇到不适。由于药物作用机制, 患者在服用帕罗西汀后产生不良反应的概率较大。另外, 任何选择性 5-羟色胺再摄取抑制剂 (selective serotonin reuptake inhibitors, SSRI) 类药物在停药后都可能会出现戒断症状, 但帕罗西汀的症状更加频繁^[14]。因此, 逐步减少药物剂量是一种既合理又安全的方法^[15]。

4.5 孤独症对生活的影响

该类主题包括主题 4 和主题 8, 平均积极情感占比 42.2%, 平均消极情感占比 53.7%, 在主题 8 中, 积极情感和消极情感相对均衡, 但与其他主题相比积极情感略高。孤独症和抑郁症等情绪障碍显著影响患者生活, 帕罗西汀虽然可能带来副作用, 但也能改善患者的情绪和生活状态。当药物产生良好效果时, 患者体验到情绪的改善和生活兴趣的增加, 进而在评论区相互加油鼓励, 也导致积极情感占比相对较高。

5 结论

传统舆情分析大多基于某一商品、社会现象或突发事件进行研究, 医药方面研究相对较少。本文通过

爬虫程序获取患者关于“帕罗西汀治疗孤独症”的评论数据, 深入剖析公众对帕罗西汀药物的关注焦点和情感倾向。从帕罗西汀的治疗效果、进口药物与国产药物、抗抑郁药物的治疗与应用、帕罗西汀的不良反应与停药反应、孤独症对生活的影响 5 个方面进行分析。研究发现, 帕罗西汀在缓解孤独症患者的抑郁、焦虑等方面具有积极作用, 但其副作用也较大, 停药后不良反应较多。因此, 相关研究者应更加关注药物的不良反应, 在提高治疗效果的同时, 不断探索减少或降低帕罗西汀副作用的方法。本文打破传统方法, 从患者评论角度对药物疗效进行主题分析和情感分析, 为医药卫生领域提供了新的研究视角和方法体系, 也为孤独症患者和医疗工作者提供了有价值的信息, 为后续研究提供了数据支持和参考。本研究仍存在局限性, 如数据量不够充足, 未来将纳入更多媒体平台数据, 更精准、全面地获取患者对于帕罗西汀的态度; 此外, 还将关注患者对于其他治疗药物的关注焦点和情感倾向, 为孤独症治疗领域提供更多有价值的信息。

作者贡献: 吕艳华负责研究设计、论文修订; 崔云龙负责数据分析、论文撰写与修订; 王康龙、钟小云负责数据分析、论文修订; 陈家珠负责论文修订。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 YU X, QIU Z, ZHANG D. Recent research progress in autism spectrum disorder [J]. Neuroscience bulletin, 2017, 33 (2): 125-129.
- 2 武莉. 基于生物信息学的阿米替林对自闭症潜在治疗意义研究 [D]. 太原: 山西医科大学, 2022.
- 3 邓丽娟, 张铮, 杨倩, 等. 自闭症治疗药物的中国专利状况分析 [J]. 中国当代医药, 2018, 25 (24): 26-29.
- 4 罗珺钰, 刘芳, 罗耀辉, 等. 抑郁症的中西医药物治疗进展 [J]. 云南中医中药杂志, 2019, 40 (5): 84-87.
- 5 李凌江, 马辛. 中国抑郁障碍防治指南 [M]. 2 版. 北京: 中华医学电子音像出版社, 2015.
- 6 BLEI D M, NG A Y, JORDAN M I. Latent Dirichlet allocation [J]. Journal of machine learning research, 2003, 3 (1): 993-1022.

(下转第 38 页)

Transfusion, 2003, 43 (9): 1186 – 1189.

15 卢芷晴, 张文明. 加拿大电子健康档案建设经验及启示 [J]. 中国档案, 2021 (3): 82 – 83.

16 Alberta Government. Shared health record – Alberta netcare [EB/OL]. [2024 – 10 – 05]. <https://www.alberta-netcare.ca/SharedHealthRecord.htm>.

17 Alberta Government. Shared health record (SHR) data and messaging specification. Version 1.10 – open government [EB/OL]. [2024 – 10 – 05]. <https://open.alberta.ca/publications/shared-health-record-shr-data-and-messaging-specification-version-1-10#summary>.

18 Canadian Institute for Health Information. Better access to electronic health information key to improving health systems [EB/OL]. [2024 – 09 – 18]. <https://www.cihi.ca/en/taking-the-pulse-a-snapshot-of-canadian-health-care-2023/better-access-to-electronic-health>.

19 KARIOTIS T, PRICTOR M, CHANG S, et al. Evaluating the contextual integrity of Australia’s my health record [J]. Studies in health technology and informatics, 2019, 265: 213 – 218.

20 Australian Digital Health Agency. My health record [EB/OL]. [2023 – 08 – 03]. <https://www.digitalhealth.gov.au/initiatives-and-programs/my-health-record>.

21 Australian Privacy Foundation. A summary of my health record – Australian privacy foundation [EB/OL]. [2023 – 08 – 03]. <https://privacy.org.au/campaigns/myhr/a-summary-of-my-health-record/>.

22 Office of the Australian Information Commissioner. About my health record [EB/OL]. [2024 – 09 – 18]. <https://www.oaic.gov.au/privacy/your-privacy-rights/health-information/my-health-record/about-my-health-record>.

23 Dossier médical partagé (DMP) [EB/OL]. [2024 – 09 – 19]. <https://sante.gouv.fr/systeme-de-sante/e-sante/mon-espace-sante/article/dossier-medical-partage-dmp>.

24 DYB K, WARTH L L. The Norwegian national summary care record: a qualitative analysis of doctors’ use of and trust in shared patient information [J]. BMC health services research, 2018, 18 (1): 252.

25 董超. 每人都将有电子健康码 [N]. 保健时报, 2022 – 11 – 17 (2).

26 赵统堂. 国外电子健康档案管理模式分析及经验借鉴 [J]. 数字与缩微影像, 2022 (1): 11 – 13.

(上接第 32 页)

7 崔金栋, 杜文强, 关杨, 等. 微博用户信息个性化推荐主题模型 LDA 演化分析研究 [J]. 情报科学, 2017, 35 (8): 3 – 10.

8 YAN X, GUO J, LAN Y, et al. A biterm topic model for short texts [C]. New York: The 22nd International Conference on World Wide Web, 2013.

9 李敏, 项朝辉. 基于微博短文本的 ChatGPT 话题舆情分析 [J]. 电脑编程技巧与维护, 2024 (4): 88 – 91.

10 韩清云. 基于主题与情感融合的网络舆情动态分析与预测反馈方法研究 [D]. 广州: 华南理工大学, 2022.

11 张悦. 知识标签网络中话题流行度预测及情感演化 [D]. 石家庄: 河北地质大学, 2022.

12 黎小妍, 张平, 黄亚非, 等. 国产和进口帕罗西汀片治疗抑郁症的成本 – 效果分析 [J]. 今日药学, 2009, 19 (8): 20 – 21, 29.

13 江开达. 精神药理学 [M]. 2 版: 北京: 人民卫生出版社, 2011.

14 FORNARO M, CATTANEO C I, DE BERARDIS D, et al. Antidepressant discontinuation syndrome: a state-of-the-art clinical review [J]. European neuropsychopharmacology, 2023, 66 (1): 1 – 10.

15 FAVA G A, GATTI A, BELAISE C, et al. Withdrawal symptoms after selective serotonin reuptake inhibitor discontinuation: a systematic review [J]. Psychother sychosom, 2015, 84 (2): 72 – 81.