·流行病与统计方法·

对话式人工智能对大学生抑郁焦虑症状疗效的 Meta 分析

唐玉萍[1],杨栋[2],杨璇[1],王福桂[1],蔡瑞[1],罗尧岳[1]
1.湖南中医药大学护理学院,湖南 长沙 410208;2.湖南省脑科医院,湖南 长沙 410000

摘要:目的　系统评价对话式人工智能(conversational artificial intelligence, CAI)干预大学生抑郁、焦虑症状的效果及其影响因素。**方法**　系统检索 PubMed、Embase、Cochrane Library、Web of Science、CNKI、WanFang、VIP 及 Sinomed 数据库中关于 CAI 干预大学生抑郁、焦虑症状的随机对照试验(randomized controlled trials,RCT),采用 Stata 18.0 进行 meta 分析。**结果**　共纳入 15 项 RCT, 涉及 1 773 名大学生。结果显示,CAI 干预组抑郁症状评分较对照组降低幅度更大(*SMD*=−0.40,95% *CI*: −0.63 ~ −0.16,*P*<0.001),焦虑结局亦呈现相似结果(*SMD* =−0.32,95% *CI*: −0.51 ~ −0.13,*P*<0.001)。亚组分析显示, 基于规则型的 CAI 在抑郁结局中效应量具有统计学意义 (*SMD*=−0.48,95*CI*:−0.77 ~ −0.18,*P*<0.001),大语言模型的 CAI 在焦虑结局中显示一定关联(*SMD*=−0.36, 95% *CI*: −0.72 ~ −0.01,*P*=0.04),但相关证据仍较有限;干预超过 4 周时合并效应量相对较大(抑郁 *SMD*=−0.86,*P*<0.001);从地域看,国内研究的合并效应量(抑郁 *SMD*=−0.65;焦虑 *SMD*=−0.50)相对大;脱落率 meta 分析显示脱落风险差异无统计学意义(*RR*=0.93,95%*CI*:0.74 ~ 1.18,*P*=0.57)。**结论**　现有 RCT 证据提示,CAI 干预与大学生抑郁或焦虑症状评分下降相关,有望作为高校心理健康服务的数字化辅助手段,但受异质性和大语言模型研究数量有限等限制,仍需更多高质量研究验证。

关键词:对话式人工智能;大学生;抑郁;焦虑;Meta 分析
中图分类号:R749.4;TP181　　**文献标志码**:A　　**文章编号**:1003−8507(2026)19−3505−07
DOI:10.20043/j.cnki.MPM.202604088

Efficacy of conversational artificial intelligence for depressive and anxiety symptoms in college students: a meta−analysis

TANG Yu−ping*, YANG Dong, YANG Xuan, WANG Fu−gui, CAI Rui, LUO Yao−yue
School of Nursing, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China

Abstract: Objective To systematically evaluate the effectiveness of conversational artificial intelligence (CAI) in alleviating depressive and anxiety symptoms among college students and to explore potential influencing factors. **Methods** Randomized controlled trials (RCTs) on CAI interventions for depressive and anxiety symptoms in college students were systematically retrieved from PubMed, Embase, Cochrane Library, Web of Science, CNKI, WanFang, VIP, and Sinomed databases. Meta−analysis was performed using Stata 18.0. **Results** A total of 15 RCTs involving 1 773 college students were included. The pooled results showed that the CAI intervention group had a greater reduction in depressive symptom scores than the control group (*SMD*=−0.40, 95% *CI*: −0.63~−0.16, *P*<0.001), with similar results observed for anxiety symptom scores (*SMD*=−0.32, 95% *CI*: −0.51~−0.13, *P*<0.001). Subgroup analyses showed that rule−based CAI had a statistically significant pooled effect on depressive outcomes (*SMD*=−0.48, 95% *CI*: −0.77~−0.18, *P*<0.001), while large language model (LLM)−based CAI was associated with reduced anxiety symptom scores (*SMD* =−0.36, 95% *CI*: −0.72 ~−0.01, *P*=0.04), although the evidence remained limited. Interventions lasting more than 4 weeks showed relatively larger pooled effect sizes (depression: *SMD*=−0.86, *P*<0.001). In terms of geographic region, studies conducted in China showed relatively larger pooled effect sizes (depression: *SMD*=−0.65; anxiety: *SMD*=−0.50). Meta−analysis of dropout rates showed no significant difference in dropout risk (*RR*=0.93, 95% *CI*: 0.74−1.18, *P*=0.57). **Conclusion** Current RCT evidence suggests that CAI interventions are associated with reductions in depressive and anxiety symptom scores among college students. CAI may serve as a digital adjunct to college mental health services. However, given the heterogeneity of included studies and the limited number of LLM−based studies, further high−quality research is needed for verification.

keywords: Conversational artificial intelligence; College students; Depression; Anxiety; Meta−analysis

基金项目:湖南中医药大学交叉学科方向培育项目(2025JC0103)
作者简介:唐玉萍(1999—),女,硕士在读,研究方向:中西医结合护理
通信作者:罗尧岳,E−mail: 943629127@qq.com

抑郁与焦虑障碍是重要的公共卫生挑战,全球疾病负担严重[1-2]。《中国国民心理健康发展报告(2023—2024)》[3]显示,我国大学生焦虑和抑郁检出率分别为44.63%和 18.69%。长期情绪障碍不仅削弱认知效能与学业表现,更可能诱发自杀倾向[4]。探索适配大学生特质的心理健康支持模式是当前高等教育与公共卫生领域的重要课题。

传统心理治疗疗效确切[5],但高校仍面临专业资源不足、高成本及病耻感等限制。对话式人工智能(conversational artificial intelligence, CAI)是一种利用自然语言处理技术提供智能化心理支持的数字化方案,治疗的个性化与共情深度强,能有效打破传统高校心理干预的限制[6-7]。当代大学生作为数字原住民,对及时、匿名、非评判性的 CAI 接纳度更高[8]。这种高性价比、可大规模扩展的干预,为大学生抑郁或焦虑的早期干预提供了关键循证补充路径。

既往研究探讨了 CAI 对抑郁或焦虑症状的干预效果[9-10],但结论尚不统一,且多纳入成人、年轻人或混合人群,对大学生这一高校心理健康服务重点人群及其影响因素的整合相对不足。本文聚焦大学生开展 meta 分析, 以期为高校心理健康服务数字化转型提供循证依据。

1　资料与方法

1.1　文献检索策略　检索 PubMed、Embase、Cochrane Library、Web of Science、CNKI、WanFang、VIP 和 Sinomed 中关于 CAI 改善大学生抑郁或焦虑症状的随机对照试验 (randomized controlled trials,RCT),时限为建库至 2026 年 3 月 13 日。检索策略为主题词与自由词结合,中文数据库检索主题涵盖“大学生、高校学生”AND“对话式人工智能、心理机器人、大语言模型”AND“抑郁、焦虑”等;PubMed 检索式见图 1。本研究已在 PROSPERO 注册(CRD420261370283)。

Search Strategy

#1	"University Student*" OR "College Student*"
#2	"Artificial Intelligence"[Mesh] OR "Generative Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[Mesh] OR "Distance Counseling"[Mesh]
#3	"Artificial Intelligence" OR "Generative Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "Distance Counseling" OR Chatbot OR Chat GPT OR LLM OR NLP OR "Conversational Agent*" OR "Virtual Assistant*" OR "Generative AI" OR "Large Language Model*" OR "Rule-based" OR "Natural Language Processing" OR "online counseling" OR "digital psychotherapy"
#4	#2 OR #3
#5	"Depressive Disorder"[Mesh] OR "Major Depressive Disorder"[Mesh]) OR "Depression"[Mesh]) OR "Anxiety"[Mesh]) OR "Anxiety Disorders"[Mesh]
#6	"Depressive Disorder" OR "Major Depressive Disorder" OR Depression OR Anxiety OR "Anxiety Disorders" OR Depress* OR Depressed OR "Depressive Symptom" OR "Emotional Depression" OR "Major Depression" OR "Major Depressive Disorder" OR MDD OR Sadness
#7	#5 OR #6
#8	"Randomized Controlled Trial"[Publication Type] OR "Controlled Clinical Trial"[Publication Type]
#9	"Double-blind Procedure" OR RCT OR "Single-blind Procedure" OR Random* OR "Randomized Controlled Trial" OR "Controlled Clincial Trial"
#10	#8 OR #9
#11	#1 AND #4 AND #7 AND #10

图 1　PubMed 文献检索策略

Figure 1　Search strategy in PubMed

1.2　文献纳入与排除标准

纳入标准:(1)研究对象:在校大学生;基线有抑郁或焦虑症状,即相关量表(如 PHQ-9、GAD-7)得分超过临床筛查分界线或处于症状检出状态;(2) 干预措施:试验组采用 CAI 干预;包括纯文本、语音或虚拟形象的多模态互动; 内容基于认知行为疗法(cognitive behavioral therapy, CBT)、接纳与承诺疗法(acceptance and commitment therapy, ACT)等循证心理干预;(3)对照措施:对照组包括注意力控制(阅读相关电子书),标准对照(真人对话、同类聊天机器人及内部对照),等候名单;(4)结局指标:抑郁或焦虑量表评分及脱落率;(5)研究类型:RCT。

排除标准:CAI 与其他干预联用无法区分效应者、重复发表、综述、会议摘要、全文或数据不可得、非中英文文献。

1.3　文献筛选与数据提取　两名研究者独立筛选文献、提取资料并交叉核对,分歧由第三名研究者裁定。提取内容包括作者、年份、国家、基线程度、平均年龄、样本量、干预措施、干预 / 随访时长、结局指标及脱落情况。

1.4　文献质量评价　采用 Cochrane 2.0 工具评价随机化过程、偏离既定干预、缺失数据、结局测量和选择性报告五个方面的偏倚风险。每项研究在每个领域均被判定为低风险、高风险或不清楚并形成总体判断。

1.5　统计学方法　应用 Stata 18.0 进行分析,连续型结局以标准化均数差 (standardized mean difference, *SMD*) 及 95%置信区间 (confidence interval,95%*CI*)表示,脱落率以风险比(risk ratio,*RR*)及 95%*CI* 表示;采

用 χ^2 检验和 I^2 评价异质性，$I^2 \leqslant 50\%$ 且 $P \geqslant 0.1$ 时采用固定效应模型，否则选用随机效应模型，通过亚组分析、元回归以探究异质性来源。检验水准 α=0.05。用逐一剔除法进行敏感性分析，漏斗图及 Egger 检验评估发表偏倚。

2　结　果

2.1　文献检索结果　初步检索文献 1 758 篇，利用 Endnote 21.0 去重后剩 1 396 篇；经题录和摘要筛选排除 1 279 篇，阅读全文后排除 102 篇，最终纳入 15 篇 RCT，共计 1 773 名研究对象。筛选流程见图 2。

2.2　纳入文献的基本特征　共纳入 15 项 RCT[9–23]，发表时间为 2017—2026 年，中文 3 篇[21–23]，英文 12 篇[9–20]。基本特征表见表 1。

2.3　纳入文献的方法质量学评价　15 项[9–23]RCT 随机化过程均为低风险；偏离既定干预部分 12 项[9–13、16–17、19–23]低风险，3 项不清楚[14–15, 18]；缺失结局数据部分 12 项[9–12, 16–23]低风险，3 项[13–15]高风险；结局测量多为自评量表，1 项[16]低风险，14 项[9–15, 17–23]不清楚；选择性报告部分 7 项[10, 15–20]低风险，8 项[9, 11–14, 21–23]不清楚。整体质量尚可，见图 3。

2.4　Meta 分析结果

2.4.1　总体干预效果分析　12 项研究[9–13, 15–20, 22]报告抑郁结局，随机效应模型显示 CAI 缓解大学生抑郁症状达到中等偏小效应水平；13 项研究[9, 11–17, 19–23]报告焦虑结局，合并结果有统计学意义。见图 4。

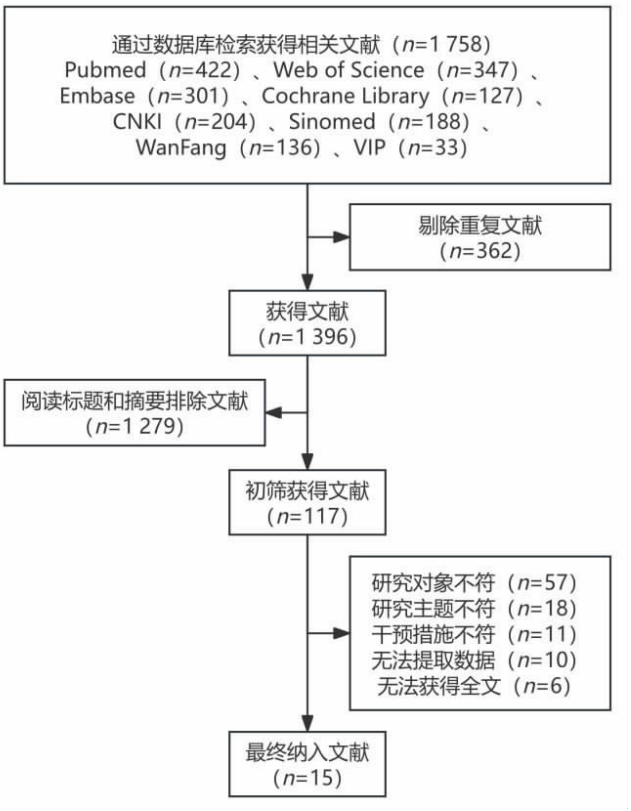

图 2　文献筛选流程图

Figure 2　Flowchart of literature screening

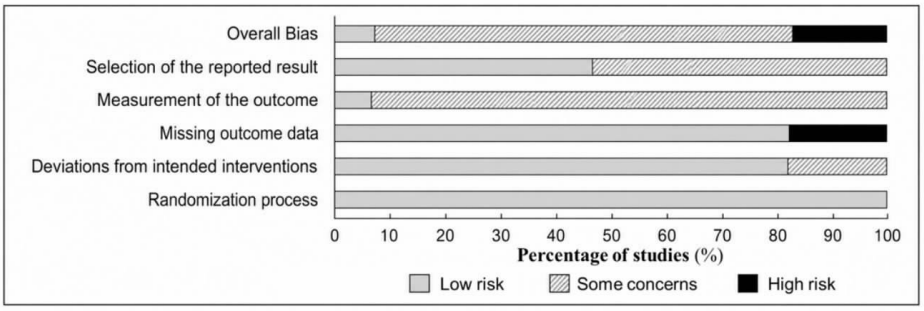

图 3　文献质量评价图

Figure 3　Quality assessment of included studies

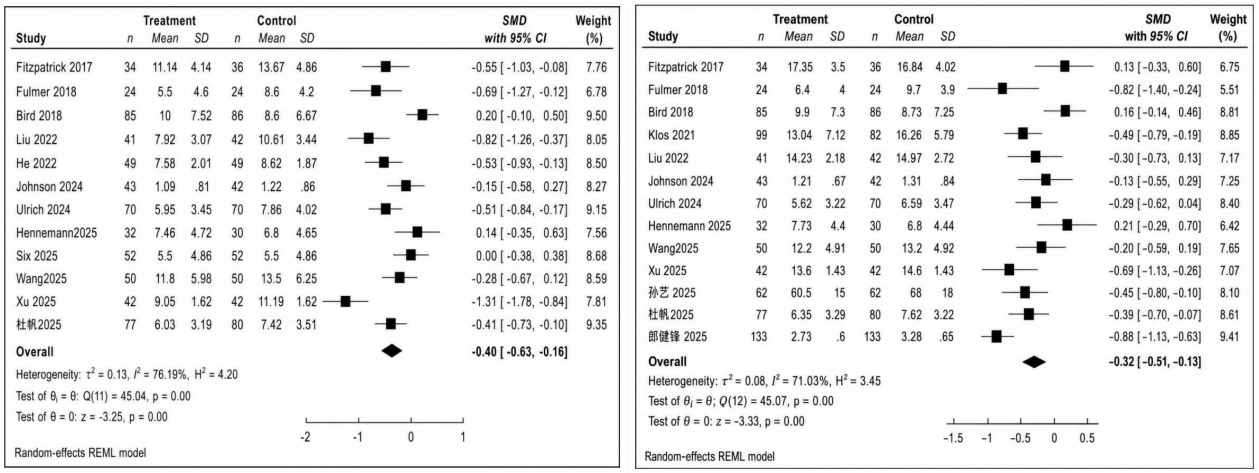

图 4　抑郁（左）及焦虑（右）症状干预效果森林图

Figure 4　Forest plot of intervention effects on depressive (left) and anxiety (right) symptoms

表 1　纳入研究基本特征表

Table 1　Basic characteristics of included studies

纳入研究及年份、国家	基线严重程度		女性占比%	平均年龄（岁）	样本量 I/C	干预措施	模型类型	对照措施	干预 / 随访时长	结局指标		脱落率（%）I/C
	抑郁	焦虑								抑郁	焦虑	
Fitzpatrick 2017[11]美国	中度	重度	67.0	22.2	34/36	Woebot,基于 CBT,每日对话、心情监测、认知重构、心理教育及放松练习等模块	Rule	①	2w/ 即刻	PHQ–9	GAD–7	8.82/30.56
Fulmer 2018[12]美国	轻度	轻度	70.0	22.9	24/24	Tess,整合多种心理治疗模式,经即时通讯软件提供情绪监测、心理教育及 CBT 干预	Rule	①	2w/ 即刻	PHQ–9	GAD–7	0.00/4.00
Bird 2018[13] 英国	轻度	中度	81.6	22.08	85/86	MYLO,基于 MOL,经开放式提问引导用户觉察思维过程并自主探索问题解决方式	Rule	②	单次 /2w	DASS–21–D	DASS–21–A	62.35/59.3
Klos 2021[14] 阿根廷	N/A	重度	87.2	22.90	99/82	Tess,同 Fulmer 2018	Rule	①	8w/8w	N/A	GAD–7	60.61/58.54
Liu 2022[9] 中国	中度	重度	55.4	23.08	41/42	小南,基于 CBT 和 RASA 框架,经微信公众号每日打卡、情绪评估、抑郁科普及认知重构训	Rule	①	16w/16w	PHQ–9	GAD–7	19.51/28.57
He 2022[10]中国	中度	N/A	37.2	18.78	49/49	小 E,基于 CBT 和 RASA 框架,经微信公众号认知挑战、放松训练、自尊提升、感恩日记等	Rule	①	1w/1m	PHQ–9	N/A	8.16/20.41
Johnson 2024[15]澳大利亚	轻度	轻度	84.0	20.65	43/42	GSH–AI,基于 CBT,引导使用 CAI 作为“虚拟教练”提供解释、案例和建议	LLM	①	4w/8w	PHQ–9	GAD–7	44.19/54.76
Ulrich 2024[16] 瑞士	轻度	轻度	73.6	26.71	70/70	MISHA App,整合 CBT、正念疗法及心理教育,提供压力管理会话、放松音频及情绪监测	Rule	③	4–7w/ 即刻	PHQ–9	GAD–7	40.00/20.00
Hennemann 2025[17]德国	轻度	轻度	93.5	23.60	32/30	ChatGPT–4,基于 CBT,围绕拖延相关心理教育、行为分析、自我管理及认知修正开展对话	LLM	②	17d/6m	PHQ–9	GAD–7	18.75/16.67
Six 2025[18]美国	轻度	N/A	18.7	19.97	52/52	AirHeart App,基于简版 CBT,包含虚拟教练、CBT、心情日志和情绪追踪	Rule	②	2w/2w	PHQ–8	N/A	24.82/24.82
Wang2025[19] 中国	轻度	中度	62.0	20.8	50/50	Psy–Bot,基于 CBT 并结合中国大学生情境,提供压力、焦虑管理及心理韧性训练	Rule	③	1w/1w	CES–D	GAD–7	6.00/16.00
Xu 2025[[20]中国	中度	重度	48.8	23.3	42/42	Neil,基于 CBT 和 RASA 框架,结合文本、语音、表情及动作提供心理教育与认知训练	Rule	②	16w/16w	PHQ–9	GAD–7	0.00/38.10
孙艺 2025[21] 中国	N/A	中度	79.0	20.44	62/62	ChatGPT–4,在模拟英语口语测试情境中提供即时反馈、语言验证及焦虑应对支持	LLM	②	单次 / 即刻	N/A	SAS	0.00/0.00
杜帆 2025[22] 中国	轻度	轻度	43.9	22.27	77/80	“心理陪伴 AI”,基于 CBT,围绕情绪识别、负性思维挑战等开展个性化干预	LLM	③	3w/6w	PHQ–9	GAD–7	14.29/7.50
郎健锋 2025[23]中国	N/A	中度	50.0	21.00	133/133	DeepSeek–AI,基于职业成熟度理论,提供自我评价、信息收集、目标筛选、职业规划及问题解决训练	LLM	③	5d/ 即刻	N/A	就业焦虑量表	2.26/2.26

注：PHQ–8/9 为 8 项 /9 项患者健康问卷;GAD–7 为 7 项广泛性焦虑障碍量表;DASS–21–D/A 为抑郁 – 焦虑 – 压力量表简版 (21 项)—抑郁 / 焦虑分量表;CES–D 为流调中心抑郁量表;SAS 为焦虑自评量表;Rule 为基于规则的技术;LLM 为大语言模型；MOL 为层级法;I/C 为干预组 / 对照组;①注意力控制；②标准对照；③等候名单；N/A 为无。

2.4.2　亚组分析与元回归　为进一步探究异质性来源,按照技术类型、地域背景、干预时长、对照组设置及基线严重程度进行亚组分析,对平均年龄和女性占比进行元回归。见表 2。

亚组分析显示,规则型在抑郁结局中具有统计学意义,大语言模型(large language model,LLM)型在焦虑结局中与症状评分降低相关;国内研究效应量较国外研究更大;干预时长>4 周时合并效应量相对较大;基线症状程度越高效应量越大。元回归显示,焦虑结局中女性占比与效应量存在弱相关(β=0.011,P=0.046,R^2=28.12%)。

2.4.3　脱落率 meta 分析　为评估 CAI 的长期依从性,对 15 项 RCT[9-23]进行脱落率 meta 分析。随机效应模型显示 CAI 组和对照组在脱落风险上无统计学差异(*RR*=0.93,95%*CI*:0.74 ~ 1.18,*P*=0.57)。

表 2　抑郁或焦虑亚组分析结果

Table 2　Subgroup analysis results for depression or anxiety

亚组		抑郁						焦虑					
		k	*SMD*(95%*CI*)	*P* 异质性值	I^2(%)	*P* 效应值	*P* 组间	*k*	*SMD*(95%*CI*)	*P* 异质性值	I^2(%)	*P* 效应 值	*P* 组间
技术类型	Rule	9	−0.48 [−0.77, −0.18]	<0.001	79.1	<0.001	0.185	8	−0.29 [−0.52, −0.06]	0.005	64.9	0.012	0.732
	LLM	3	−0.19 [−0.50, 0.12]	0.168	44.1	0.242		5	−0.36 [−0.72, −0.01]	<0.001	79.6	0.044	
地域背景	国内	5	−0.65 [−0.99, −0.30]	0.008	73.3	<0.001	0.042	6	−0.50 [−0.73, −0.28]	0.027	57.8	<0.001	0.050
	国外	7	−0.20 [−0.46, 0.06]	0.007	64.8	0.134		7	−0.16 [−0.42, 0.09]	0.005	67.3	0.214	
干预时长(周)	短期(≤2)	5	−0.21 [−0.50, 0.09]	0.014	66.6	0.177	0.056	5	−0.26 [−0.66, 0.13]	<0.001	85.2	0.192	0.610
	中期(>2,≤4)	4	−0.27 [−0.57, 0.02]	0.123	46.5	0.073		4	−0.27 [−0.63, 0.11]	0.045	65.7	0.160	
	长期(>4)	3	−0.86 [−1.31, −0.40]	0.024	73.0	<0.001		4	−0.43 [−0.61, −0.25]	0.454	0	<0.001	
对照组	注意力控制	5	−0.53 [−0.76, −0.30]	0.291	23.4	<0.001	0.602	5	−0.31 [−0.59, −0.03]	0.072	52.9	0.029	0.592
	等候名单	3	−0.41 [−0.61, −0.21]	0.680	0	<0.001		4	−0.46 [−0.77, −0.15]	<0.001	74.2	<0.001	
	标准对照	4	−0.23 [−0.93, 0.46]	<0.001	91.6	0.510		4	−0.19 [−0.62, 0.24]	<0.001	79.9	0.382	
基线程度	轻	8	−0.20 [−0.41, 0.01]	0.014	58.7	0.062	<0.001	5	−0.27 [−0.53, −0.02]	0.090	47.4	0.029	0.912
	中	4	−0.79 [−1.01, −0.56]	0.063	58.8	<0.001		4	−0.35 [−0.79, 0.09]	<0.001	87.1	0.122	
	重	N/A						4	−0.35 [−0.67, −0.03]	0.062	60.9	0.034	

2.5　发表偏倚与敏感性分析　采用漏斗图结合 Egger 检验评估发表偏倚。结果显示,抑郁与焦虑结局漏斗图基本对称;Egger 检验均未提示显著发表偏倚($P_{抑郁}$=0.153、$P_{焦虑}$=0.289)。按国内外研究进一步分析,国内研究焦虑结局漏斗图略不对称,Egger 检验 *SMD*=−0.504(95% *CI*: −2.58 ~ 12.11,*P*=0.146),经剪补法补入 1 项研究后 *SMD*=−0.558 (95% *CI*: −0.774 ~ −0.343),与校正前 *SMD*=−0.504 方向及统计学意义未改变。逐一剔除任一研究,合并效应方向及显著性均未发生实质改变,结果较稳定。

3　讨　论

3.1　CAI 干预与大学生抑郁或焦虑症状评分下降相关　本研究显示,与对照组相比,CAI 干预组大学生抑郁或焦虑症状评分下降更明显,效应量为中等偏小范围,与既有研究结果基本一致[19,24],提示其在高校心理健康服务中具有一定辅助应用潜力,但由于纳入研究在干预形式、对照条件及结局测量等方面存在差异,仍应将其理解为统计学关联,而非对干预效果的确定性判断。纳入研究中的 CAI 干预多以 CBT、正念

或心理教育等结构化内容为基础,通过数字平台提供情绪监测、认知调整和应对技巧支持,其即时反馈、持续交互及低门槛使用特点可能有助于提高大学生接受度和自我调节参与度[25-26]。但相关机制仍需结合客观行为或生理指标进一步验证。

3.2 *疗效差异与异质性来源分析*　本研究抑郁和焦虑异质性较高,提示合并效应需谨慎解释。亚组分析显示,干预时长、基线症状程度和地域背景等可能与效应差异有关,但因亚组研究数量和研究设计差异限制,尚不能完全解释异质性。结合纳入研究特征,残余异质性还可能与干预载体和交互形式、干预内容、随访节点、对照组活性及结局测量工具差异有关。故 CAI 干预研究在方案设计、对照组设置和结局评价方面有待进一步标准化。

技术类型亚组显示,规则型 CAI 在抑郁结局中合并效应具有统计学意义,LLM 型 CAI 在焦虑结局中与症状评分降低相关,但 LLM 型研究数量有限,尤其抑郁结局仅 3 项 RCT,可能导致统计效能不足。故尚不足以判断规则型与 LLM 型 CAI 在大学生抑郁或焦虑症状干预中的优劣,相关结果应视为探索性发现。传统规则型多基于固定 CBT 流程,内容标准化程度较高;LLM 型在语言生成和交互灵活性方面具备特色,但其效果可能同时受提示设计、干预场景、对照组设置和用户参与度等多因素影响。未来仍需开展直接比较研究,以明确不同技术路径的适用范围。

干预时长达四周以上时合并效应量相对较大。这可能与心理调节技巧需要一定时间进行内容内化和行为实践有关,短期干预虽可能在情绪即时调节方面发挥作用,但较难形成稳定的认知和行为改变[27]。基线症状较重者合并效应量相对大,提示 CAI 对有一定症状基础的大学生可能更有干预价值。症状较重者通常具备更大改善空间,更容易在干预后显示出量表分值及症状变化[28];此外,症状较重学生的现实困扰更明确,与系统互动的主动性和持续性可能更高,故更易从心理支持中获益。

3.3 *本土化适配与高校应用价值*　国内研究效应量高于国外研究,可能与文化适配、服务平台、样本基线差异及潜在发表偏倚等多因素有关。与多国背景下的干预方案相比,国内研究多围绕中国大学生常见的学业压力、人际适应、就业压力等问题设计干预内容,并更多依托微信等高频使用平台,因而在便利性、接触频率和持续使用方面可能具备一定优势。但国内样本基线严重程度整体高于国外,且国内焦虑研究存在一定发表偏倚风险,可能在一定程度上放大干预效应。

对照组设置可能影响合并效应解释。本研究对照条件包括注意力控制、等候名单和标准对照,干预活性并不一致。等候名单通常缺乏主动干预,可能放大组间差异;注意力控制可部分控制时间和信息接触因素,但不一定具备等同心理干预成分;标准对照因包含真人对话、同类聊天机器人或内部对照等较高活性成分,可能缩小 CAI 相对效应。因此,总体合并效应可理解为不同对照条件下的平均估计,而非 CAI 相对于所有对照条件均具有一致优势。

从公共卫生视角看,CAI 的意义并不完全取决于其是否优于真人治疗,而在于能否以较低成本扩大有效干预的覆盖面。其一,既往研究显示,CAI 在部分心理健康场景中可获得与传统干预相近的疗效,可能为心理健康问题提供具有成本效益潜力的辅助支持方式[29-30];其二,与传统面对面咨询相比,CAI 可提供即时、重复、低门槛的心理支持,减少对专业人员数量、固定服务时间和服务地点的依赖;其三,当前高校仍面临专业人员不足、学生需求增长和求助污名感等现实问题。在此背景下,即使 CAI 未表现出明显优效,其在可及性、低成本、低污名感和规模化方面的优势,仍使其有望成为高校心理服务中的数字化辅助工具,具备一定的公共卫生价值。

3.4 *本研究存在的不足与局限性*　本研究仍存在一定局限。纳入研究异质性较高,合并效应可能受研究对象、干预方案、人机交互形式、对照组类型及结局测量工具等因素影响;结局多基于自评量表,缺乏客观行为、生理或数字表型指标,难以验证 CAI 干预的具体作用机制;部分研究脱落率较高且随访较短,长期效果仍不明确;当前 LLM 型 RCT 数量有限,抑郁结局仅 3 项,技术类型相关结果仍属探索性分析。未来需开展大样本、长期随访、设计更规范的高质量 RCT 研究。

利益冲突声明　本研究不存在任何利益冲突

参考文献

[1] Hannah K, Marie K, Olaf H, et al. The global economic burden of health anxiety/hypochondriasis– a systematic review[J]. BMC Public Health, 2023, 23(1): 2237.

[2] World Health Organization.World mental health report: Transforming mental health for all[M]. Geneva: World Health Organization, 2022.

[3] 孙向红,蒋毅,陈雪峰,等. 心理健康蓝皮书:中国国民心理健康发展报告(2023-2024)[M]. 北京:社会科学文献出版社,2025. Sun XH, Jiang Y, Chen XF, et al. Report on national health development in China (2023 - 2024) [M]. Beijing: China Social Science Documentation Publishing House, 2025.(In Chinese)

[4] Ramón-Arbués E, Gea-Caballero V, Granada-López JM, et al. The prevalence of depression, anxiety and stress and their associated factors in college students [J]. International Journal of Environmental Research and Public Health, 2020, 17(19): 7001.

[5] Fordham B, Sugavanam T, Edwards K, et al. The evidence for

cognitive behavioural therapy in any condition,population or context: a meta-review of systematic reviews and panoramic meta-analysis[J]. Psychological Medicine, 2021, 51(1): 21-29.

[6] Li Y, Liang SR, Zhu BQ, et al. Feasibility and effectiveness of artificial intelligence-driven conversational agents in healthcare interventions: A systematic review of randomized controlled trials[J]. International Journal of Nursing Studies, 2023, 143: 104494.

[7] Hua Y, Siddals S, Ma Z, et al. Charting the evolution of artificial intelligence mental health chatbots from rule - based systems to large language models: a systematic review[J]. World Psychiatry, 2025, 24 (3): 383-394.

[8] Zhang QY, Zhang RW, Xiong YN, et al. Generative AI mental health chatbots as therapeutic tools: systematic review and Meta-Analysis of their role in reducing mental health issues [J]. Journal of Medical Internet Research, 2025, 27: e78238.

[9] Liu H, Peng H, Song X, et al. Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness[J]. Internet Interv, 2022, 27: 100495.

[10] He YH, Yang L, Zhu XK, et al. Mental health chatbot for young adults with depressive symptoms during the COVID-19 pandemic: Single-Blind, Three-Arm randomized controlled trial [J]. Journal of Medical Internet Research, 2022, 24(11): e40719.

[11] Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial[J]. JMIR Ment Health, 2017, 4(2): e19.

[12] Fulmer R, Joerin A, Gentile B, et al. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial[J]. JMIR Ment Health, 2018, 5(4): e64.

[13] Bird T, Mansell W, Wright J, et al. Manage your Life online: a web-based randomized controlled trial evaluating the effectiveness of a problem-solving intervention in a student sample[J]. Behavioural and Cognitive Psychotherapy, 2018, 46(5): 570-582.

[14] Klos MC, Escoredo M, Joerin A, et al. Artificial Intelligence-Based chatbot for anxiety and depression in university students: pilot randomized controlled trial[J]. JMIR Form Res, 2021, 5(8): e20678.

[15] Johnson C, Egan SJ, Carlbring P, et al. Artificial intelligence as a virtual coach in a cognitive behavioural intervention for perfectionism in young People: A randomised feasibility trial [J]. Internet Interventions-the Application of Information Technology in Mental and Behavioural Health, 2024, 38: 100795.

[16] Ulrich S, Lienhard N, K ü nzli H, et al. A Chatbot-Delivered stress management coaching for students (MISHA App): pilot randomized controlled trial[J]. JMIR mHealth and uHealth, 2024, 12: e54945.

[17] Hennemann S, Fähnrich JM, Tietze C, et al. Efficacy of a chatbot (Chatgpt-4)compared to a human conversational agent for reducing procrastination:a randomized controlled pilot-trial [EB/OL]. [2026-07-09]. https://www.epistemonikos.org/en/documents/f15c4596b5f00b6de3b9122b56dfa5e64ef95fc6.

[18] Six S, Schlesener E, Hill V, et al. Impact of conversational and animation features of a mental health App virtual agent on depressive symptoms and user experience among college students: randomized controlled trial[J]. JMIR Ment Health, 2025, 12: e67381.

[19] Wang Y, Li X, Zhang Q, et al. Effect of a cognitive behavioral therapy-based AI chatbot on depression and loneliness in Chinese university students: randomized controlled trial with financial stress moderation [J]. JMIR Mhealth Uhealth, 2025, 13: e63806. (In Chinese)

[20] Xu S, Ma TC. Depression intervention using AI chatbots with social cues: a randomized trial of effectiveness [J]. Journal of Affective Disorders, 2025, 389: 119760.

[21] 孙艺. ChatGPT 口语互动对高社交焦虑大学生焦虑状态的影响及其神经机制[D]. 成都:四川师范大学,2025.
Sun Y. The influence of ChatGPT oral interaction on anxiety of college students with high social anxiety and its neural mechanism [D]. Chengdu: Sichuan Normal University, 2025.(In Chinese)

[22] 杜帆. 心理陪伴 AI" 聊天机器人对大学生抑郁倾向的干预效果与用户体验评估[D]. 上海:华东师范大学,2025.
Du F. Evaluation of intervention effect and user experience of "Psychological Companion AI" chatbot on depression tendency of college students [D]. Shanghai: East China Normal University, 2025. (In Chinese)

[23] 郎健锋. 大学生就业压力对就业焦虑的影响:社会支持与择业效能感的链式中介作用及干预[D]. 重庆:西南大学,2025.
Lang JF. The influence of employment pressure on employment anxiety of college students: the chain mediating effect of social support and career self-efficacy and intervention [D]. Chongqing: Southwest University, 2025.(In Chinese)

[24] Li H, Zhang R, Lee YC, et al. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being[J]. NPJ Digit Med, 2023, 6(1): 236.

[25] Im CH, Woo M. Clinical efficacy, therapeutic mechanisms, and implementation features of cognitive behavioral therapy-based chatbots for depression and anxiety: narrative review[J]. JMIR Mental Health, 2025, 12: e78340.

[26] Tan Y, Lattie EG, Xiang H, et al. Acceptability and usability of a WeChat-based intervention for depression in China:a mixed-methods study[J]. Digit Health, 2024, 10: 2301.

[27] Lu SC, Xu M, Wang M, et al. Effectiveness and minimum effective dose of App-Based Mobile health interventions for anxiety and depression symptom reduction: systematic review and Meta-Analysis [J]. JMIR Mental Health, 2022, 9(9): e39454.

[28] Sextl-Pl?tz T, Steinhoff M, Baumeister H, et al. A systematic review of predictors and moderators of treatment outcomes in internet-and mobile-based interventions for depression [J]. Internet Interventions-the Application of Information Technology in Mental and Behavioural Health, 2024, 37: 100760.

[29] Feng Y, Hang YM, Wu WZ, et al. Effectiveness of AI-Driven conversational agents in improving mental health among young People: systematic review and Meta-Analysis [J]. Journal of Medical Internet Research, 2025, 27: e69639.

[30] He YH, Yang L, Qian CL, et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials[J]. Journal of Medical Internet Research, 2023, 25: e43862.

收稿日期:2026-04-08