

《心理科学进展》审稿意见与作者回应

题目：精神病临床高危人群的症状和功能改变：一项系统综述和三水平元分析

作者：赵子卿，余锦婷，陈嘉彦，王芸茹，黄佳，陈楚侨

第一轮

审稿人 1 意见：

研究在前人元分析的基础上，更新了 CHR-P 人群临床症状纵向发展的定量整合分析结果；区分了 CHR-P 人群的社会功能和角色功能，并对二者的纵向改变进行了定量整合；探索了 CHR-P 人群临床症状和功能纵向发展的调节变量；阐述了 CHR-P 症状对功能的纵向影响，具有一定的理论与实践指导意义，但仍有以下问题需要做进一步考量。

意见 1. 元分析本质上属于系统综述(Systematic Review)，只不过是量化的系统综述。当然，通常我们使用系统综述来表示质性的系统综述，且一般将系统综述放在元分析前面，这样比较符合证据强度从弱到强的逻辑。

回应：感谢您的指正。我们修改了题目中“系统综述”和“元分析”的顺序，修改后的题目为“精神病临床高危人群的症状和功能改变：一项系统综述和三水平元分析”。在文稿正文中，“症状与功能的纵向关系”部分的系统综述是基于元分析纳入文献先进行的相关量分析，然后在讨论中系统阐述，故而正文的陈述顺序没有变化。

意见 2. 摘要中没有提及本研究所发现的调节变量。

回应：感谢您的建议。我们已经在摘要里补充了调节变量的结果。

见中文摘要：“调节效应结果表明，受教育程度显著调节轻微阳性症状($\beta = 0.456, p = 0.036$)和阴性症状($\beta = 0.317, p = 0.050$)的纵向变化；年龄($\beta = -0.147, p = 0.014$)、轻微阳性症状(attenuated positive symptoms, APS)亚组比例($\beta = -0.054, p = 0.001$)和短暂间歇性精神病性症状(brief limited intermittent psychotic symptoms, BLIPS)亚组比例($\beta = 0.185, p = 0.002$)显著调节情感症状的纵向变化。”

见英文摘要：“The moderation analysis revealed that level of education significantly moderated the longitudinal changes in attenuated positive symptoms ($\beta = 0.456, p = 0.036$) and negative symptoms ($\beta = 0.317, p = 0.050$); age ($\beta = -0.147, p = 0.014$), proportion of attenuated positive symptoms (APS) subgroup ($\beta = -0.054, p = 0.001$), and proportion of brief limited intermittent psychotic symptoms (BLIPS) subgroup ($\beta = 0.185, p = 0.002$) significantly moderated the longitudinal changes in affective symptoms.”

意见 3. 1.4 研究目的与研究问题，没有突出本研究区分社会功能和角色功能的创新点。

回应：感谢您宝贵的意见。我们已在原文添加了相应表述，以突出本研究的创新点。

见引言 1.4：“综上，本研究的目的是：(1)纳入 CHR-P 人群临床结局纵向发展的最新研究，量化整合 CHR-P 人群各维度的临床症状和整体功能水平变化，区分社会功能和角色功能，关注功能的纵向发展，并探讨潜在的调节变量对 CHR-P 人群临床纵向结局的影响；(2)系统综述 CHR-P 人群临床症状对功能的纵向影响。”

意见 4. 文献检索一般需要提供具体的检索截止日期以及最近更新检索的日期。

回应: 感谢您的指正。本研究首次的检索日期为 2023 年 9 月 8 日, 我们于 2024 年 6 月 5 日进行了更新检索, 并将更新检索后符合标准的研究纳入重新进行分析。

意见 5. 检索词中为何没有包括“瓦解症状”以及 CHR-P 人群各个亚组名称? 另外, 最好提供下瓦解症状的英文, 方便读者理解。

回应: 感谢您的指正。我们研究最初确定检索词参考 De Pablo 等人(2021)的研究, 因此没有将“瓦解症状”纳入检索词, 但是我们的研究结果最终包含瓦解症状, 确实需要将其纳入检索词的范围。我们重新将“瓦解症状”和 CHR-P 各亚组名称纳入检索词进行检索, 并且与 2024 年 6 月 5 日进行更新检索, 最终检索词如下:

英文检索词: (at-risk mental state (ARMS)) OR (clinical high risk (CHR)) OR (ultra-high risk (UHR)) OR (basic symptom) OR prodrom* OR CHR-P OR (attenuated psychosis syndrome) OR (psychosis risk) OR (psychotic risk) OR (attenuated psychosis syndrome (APS)) OR (brief limited intermittent psychotic symptoms (BLIPS)) OR (genetic risk and deterioration syndrome (GRD)) OR (cognitive-perceptive basic symptoms (COPER)) OR (cognitive disturbances (COGDIS)) AND (negative symptom) OR (attenuated positive symptom) OR (depressive symptom) OR (disorganized symptom) OR (disorganization symptom) OR (general symptom) AND (functioning) OR (global functioning) OR (social functioning) OR (role functioning) OR (neurocognitive functioning) AND longitudinal OR follow-up OR cohort;

中文检索词: 超高危 OR 超高危人群 AND 阴性症状 OR 缓解阳性症状 OR 阳性症状 OR 一般症状 OR 抑郁症状 OR 瓦解症状 AND 功能 OR 社会功能 OR 角色功能

另外, 我们也已将本研究出现的症状名称进行了英文注释, 方便读者理解。

见引言 1.1.1: “目前已有大量实证研究关注 CHR-P 人群的临床症状的纵向改变, 包括轻微阳性症状(attenuated positive symptom, Værnes et al., 2022)、阴性症状(negative symptom, Paetzold et al., 2021)、瓦解症状(disorganized symptom, Saito et al., 2017)、一般精神病理症状(general symptom 李佳 等, 2022)、情感症状(affect symptom, de Wit et al., 2014)等,”

意见 6. 文献的编码完成后是否计算了评分者一致性?

回应: 感谢您的指正。我们补充了计算了编码结果的评分者一致性, 采用 R 计算了 Kappa 系数。

见方法 2.2: “三位作者分别单独对纳入文献进行如下编码, 编码不一致的内容经过协商之后达成一致: (1)取样地区(亚洲、欧洲、北美洲、大洋洲) (Kappa = 0.92); (2)随访时间(Kappa = 0.73); (3)被试数量(Kappa = 0.76); (4)被试年龄(Kappa = 0.92); (5)被试性别(Kappa = 0.93); (6)受教育年限(Kappa = 0.99); (7)CHR-P 亚组(APS: Kappa = 0.64; BLIPS: Kappa = 0.91; GRD: Kappa = 0.88); (8)CHR-P 评估工具(Kappa = 0.78); (9)精神病转化率(Kappa = 0.63);”

见方法 2.2: “本研究采用针对队列研究的修改版纽卡斯尔-渥太华量表(Newcastle-Ottawa-Scale, NOS)评估纳入文献的质量。该量表根据纳入研究的代表性(representativeness)、暴露(exposure)、结果(outcomes)、随访时间(follow-up period)和随访脱落(losses to follow-up)进行 0-8 分的质量评估(De Pablo et al., 2021; De Pablo et al., 2022), 分数越高代表文献质量越好。编码一致性为 Kappa = 0.32。”

意见 7. 作者混用了基本症状(basic symptoms, BS)和认知障碍组(cognitive disturbances ,

COGDIS), 在引言部分介绍亚组时使用了 COGDIS, 在方法部分介绍亚组时使用了 BS, 在表格呈现数据时又用了 COGDIS。请作者澄清一下并做统一表述。

回应:感谢您仔细的审阅。我们重新阅读对于 CHR-P 人群确定标准的相关文献, 澄清了 BS 与 COGDIS 之间的关系, 并在引言和方法中加以修改。

见引言 1.3: “CHR-P 人群不同亚组也可能是调节变量。根据 UHR 标准, CHR-P 人群可以分为轻微阳性症状(APS)亚组、短暂间歇性精神病性症状(BLIPS)亚组、遗传风险和衰退综合征(GRD)亚组; 根据 BS 标准, CHR-P 人群可以分为认知-知觉基本症状(COPER)亚组和认知障碍(COGDIS)亚组(Fusar-Poli et al., 2013)。APS 亚组指过去一年内出现轻微的精神病性症状, 但未达到精神病诊断标准的人群; BLIPS 亚组指过去一年内经历了一两次短暂的、间歇性的精神病性症状的人群; GRD 亚组指有精神病家族史或分裂型人格障碍, 并且在过去一年内出现明显的功能损害的人群; COGDIS 亚组指根据 BS 诊断的纳入标准, 在过去一段时间内经历一系列认知功能异常的人群。以往元分析的结果显示, CHR-P 样本中更低的 BLIPS 亚型占比与更好的恢复率相关(De Pablo et al., 2022); 与 APS 亚组相比, BLIPS 亚组有更高的精神病转变风险; GRD 亚组不常见, 转变为精神病的风险较低; 而 COGDIS 亚组比 COPER 亚组更能预测精神病的转化(Schultze-Lutter et al., 2007), 因此, 不同亚型 CHR-P 人群可能会有着不同的临床症状和功能的纵向发展轨迹。”

见方法 2.3.3: “本研究的调节变量涉及: (1)连续变量: 样本平均年龄、样本男性占比、文献质量、CHR-P 各个亚组占比(即 APS、BLIPS、GRD、COGDIS、COPER)、受教育年限; (2)分类变量: 随访时间。为确保调节效应结果的代表性, 各调节变量的效应量应不少于 5 个(孟现鑫 等, 2023)。”

意见 8. 文献筛选流程图并不符合 PRISMA 2020 的规范 (<https://www.prisma-statement.org/prisma-2020-flow-diagram>)。另外, 排除文献中的“其他”是指什么? 需交待清楚。

回应:感谢您宝贵的意见。我们已经按照 PRISMA 2020 的规范修改了文献筛选流程图; 并且我们已经将阅读全文排除文献的标准进行进一步解释说明, 最终排除原因有: 人群不符合 ($n = 16$)、结果变量获取不完整 ($n = 154$)、非纵向研究 ($n = 12$)、重复样本 ($n = 2$)、接受特定干预 ($n = 10$)、非期刊论文 ($n = 2$)。

意见 9. 发表偏倚检验的结果, 需指明数据所对应的表格, 方便读者查阅。

回应:感谢您仔细的审阅。我们已经在原文中指明发表偏倚检验的结果所对应的表格。

见结果 3.2: “如表 2 所示, 大部分元分析的 Eggers 回归检验的 p 值均大于 0.05(除一般精神病理症状 $p = 0.027$), 表明发表偏差不明显; 各个元分析的 Rosenthal 失安全系数 N 均大于 $5k+10$ (k 为效应量数目), 也表明不存在严重的发表偏差。”

意见 10. 表 2 主效应分析结果中的 p 值, 0.000 一般替换成 <0.001 。

回应:感谢您的指正。我们已将表二中 $p = 0.000$ 改为 $p < .001$ 。

意见 11. 社会功能的改善不显著, 但效应量的值却很大 (Hedges' $g = 0.664$)。如何看待这一结果?

回应:感谢您的提问, 这个结果确实值得深思。Hedges' g 效应量在 0.5 到 0.8 的范围内属于中等效应量。我们对进行社会功能元分析纳入的研究进行了进一步分析, 认为这个结果可能反映出: 元分析结果显示社会功能改善有中等程度的效应量(Hedges' g 值为 0.664), 但这种

改善并没有达到统计学上的显著。可能是因为在样本中存在着相对较大的差异，即可能是由于个别研究的参与者在社会功能改善方面的变化比较显著，导致整体效应较大。但由于纳入的研究之间存在很大的差异，这种差异性可能使得整体效应并不显著。我们发现，在针对计算社会功能改变效应量改变纳入的 9 个研究（16 个效应量）中，社会功能改变的效应量差异很大(Hedges'g range = 0.04 ~ 3.96)，这种情况下，即使效应量看起来不算小，也可能由于差异性过大而不能达到统计学显著性的标准。该结果提示着精神病高危人群在社会功能改善上存在较大的差异，某些高危人群在社会功能方面有较大的改善，而有些高危人群的社会功能可能并没有很明显的改善。

意见 12. 症状与功能的纵向相关关系的元分析中，所纳入的研究数量非常少，最多只有 6 个研究。结果的可靠性是令人怀疑的。

回应：感谢您仔细的审阅。症状与功能的纵向相关关系分析可纳入的研究数目较少，这的确是我们元分析的一个局限。因此，我们把这部分作为一个系统综述，描述性地阐述结果，也是作为一个初步的探索性结果，提示未来需要更多相关的实证研究，去验证结果的可靠性。我们在局限中也补充了相应表述。

见讨论 4.4 意义与局限：“第四、症状与功能的纵向相关关系结果还需进一步验证。由于现有相关研究数量限制，本研究仅对于症状与功能的纵向相关关系进行了系统综述，提示未来需要更多相关的实证研究验证结果的可靠性，以及进行全面地量化整合”。

意见 13. 调节效应的讨论显得过于单薄。

回应：感谢您的指正。我们对于调节效应结果进行了进一步解释，并深入阐述从结果中得到的启示。

见讨论 4.3：“调节效应结果表明个体差异在 CHR-P 人群症状改善的重要性。CHR-P 个体年龄越大，情感症状随时间的改善越小，可能显示了年轻个体具有较高的神经可塑性，能够更好地自然恢复或响应治疗；受教育程度越高的 CHR-P 个体，其轻微阳性症状和阴性症状的改善更大，可能表明高教育水平的个体具有更好的应对机制，帮助个体有效管理和改善症状，也提示提升受教育水平作为辅助治疗的可行性(Bora, 2018)。年龄和教育水平等个体特征对症状变化起到显著的调节作用，提示对于 CHR-P 人群进行个性化干预的重要性，对于 CHR-P 人群的研究也更应整合个体差异的因素，以更加全面地理解其症状的发展进程(Fusar-Poli et al., 2020)。另外，在控制了年龄后，APS 和 BLIPS 亚型的比例对情感症状变化的调节作用并不显著，表明 CHR-P 亚型对情感症状变化的影响可能受到其他因素的干扰。本研究并没有发现随访时间对临床结局改变的调节作用，与 De Pablo 等人(2021)的结果不一致，这可能是因为本研究纳入的样本随访时间大部分在 2 年以内(77.3%)。由于数据量的限制，未来研究可以进一步探究异质性的来源，如工作状况(Cotter et al., 2017)、共病(Usui et al., 2022)、基线的药物使用和随访时间(De Pablo et al., 2021)。为了澄清调节效应，也需要更多针对性的随访研究。”

Bora, E. (2018). A neurodevelopment and neuroplasticity-based framework for early intervention in psychotic disorders. *Psychological medicine*, 48(3), 353-361.

Fusar-Poli, P., De Pablo, G. S., Correll, C. U., Meyer-Lindenberg, A., Millan, M. J., Borgwardt, S., ... & Arango, C. (2020). Prevention of psychosis: advances in detection, prognosis, and intervention. *JAMA psychiatry*, 77(7), 755-765.

.....

审稿人 2 意见：本研究通过元分析方法对精神病临床高危 (CHR-P) 人群临床症状和功能改变

的纵向研究进行定量整合分析，纳入最新的研究样本，更新 CHR-P 人群纵向临床症状和功能结局，探究潜在的调节变量，并阐明 CHR-P 症状对功能的纵向影响，为精神病早期识别和干预提供指导，具有重要的临床意义。然而，某些重要的术语描述欠规范或有歧义，某些数据和资料需要澄清或补充（具体详见稿件上修改模式的标记）。

意见 1. 摘要“全面整合 CHR-P 人群各个临床症状和功能的纵向改变”中的“各个”建议改为“各维度”。

回应：感谢您的建议。“各维度临床症状”确实表述更为准确，我们已将摘要和全文中相应的“各个临床症状”改为“各维度临床症状”。

意见 2. 摘要“CHR-P 人群减弱的阳性症状”中，attenuated 建议翻译为“轻微”，下同。

回应：感谢您对于专业术语准确性的建议。我们已将摘要和全文中“减弱的阳性症状”改为“轻微阳性症状”。

意见 3. 摘要“基线的阴性症状显著负向预测追踪的整体功能”中，follow-up 建议翻译为“随访”，下同。

回应：感谢您对于专业术语准确性的指正。我们已将摘要和全文中对于 follow-up 的翻译改为“随访”。

意见 4. 摘要“本研究首次关注 CHR-P 人群特定功能的变化以及症状对功能的纵向预测作用，表明阴性症状的持续和社会功能损害是 CHR-P 人群的核心特征”中，阴性症状的持续？根据上文，阴性症状是也有改善的：Hedges' $g = -0.450$ 。这句总结需要重新改写。

回应：感谢您的指正。我们将本句总结改写为“本研究首次关注 CHR-P 人群特定功能的变化以及症状对功能的纵向预测作用，表明阴性症状较小的改善和社会功能损害是 CHR-P 人群的核心特征”。

意见 5. 引言第一段“在这段时期，妄想、幻觉、形式性思维障碍 等这些精神分裂症阳性症状轻微或短暂出现”中的“形式性思维障碍”建议改为：思维形式障碍。

回应：感谢您对于专业术语准确性的指正。我们已经将原文中“形式性思维障碍”修改为“思维形式障碍”。

意见 6. 引言第一段“目前研究将同时符合基本症状(basic symptoms, BS)诊断与 UHR 标准的人群定义为精神病临床高危 (clinical high-risk for psychosis, CHR-P)人群(Fusar-Poli et al., 2016)”：这个描述可能有误，需要仔细阅读原文并改写。按照通常的理解，临床高危和超高危描述的是同一群体；甄别和诊断这一群体有两个途径：基本症状的标准，或 UHR 症状的标准。

在这里请简述两种途径的理论差异，以及各自的常用评估工具。

回应：感谢您的指正。我们重新阅读相关文献，并对 CHR-P 人群的定义进行了修改，并且描述了鉴定 CHR-P 人群的两种标准——UHR 标准和 BS 标准——的理论差异及各自常用的评估工具。

见引言第一段：“处在这一阶段的人群被称作精神病临床高危 (clinical high-risk for psychosis, CHR-P)人群(Fusar-Poli et al., 2016)。目前有两种识别 CHR-P 人群的标准：超高危(ultra-high risk, UHR)标准和基本症状(basic symptoms, BS)标准。UHR 标准以精神病的临床阶段模型为基础，包括出现以下一项或多项：轻微阳性症状(attenuated positive

symptoms, APS)、短暂间歇性精神病性症状(brief limited intermittent psychotic symptoms, BLIPS)、遗传风险和衰退综合征(genetic risk and deterioration syndrome, GRD)(McGorry et al., 2006; Fusar-Poli et al., 2013; 杨学智 等, 2023)。UHR 标准的测量工具主要包括高危精神状态综合评估(comprehensive assessment of at-risk mental states, CAARMS, Yung et al., 2006) 和精神病风险综合征定式访谈(structured interview for psychosis-risk syndromes, SIPS, McGlashan et al., 2010)。BS 指在动机、情感、思维、语言、躯体、感知、运动等方面主观经历的亚临床障碍。该标准以期在功能损害和精神病前驱症状出现之前尽早发现精神病风险 (Poletti et al., 2019), 包括认知-知觉基本症状(cognitive-perceptive basic symptoms, COPER)标准和认知障碍(cognitive disturbances, COGDIS)标准, 评估 BS 的工具主要是成人和儿童青少年版本的精神分裂症倾向测试(Schizophrenia Proneness Instrument, Adult (SPI-A, Schultze-Lutter et al., 2007) or Child & Youth version (SPI-CY, Schultze-Lutter et al., 2012)。因为 UHR 和 BS 标准涉及互补的临床特征: BS 标准可能识别出更早期的前驱状态, 而 UHR 标准则反映了稍后一些的阶段, 所以目前这两种标准倾向于被同时用于评估 CHR-P 个体(Fusar-Poli et al., 2013)。”

意见 7. 引言第一段“CHR-P 人群成为精神病临床研究和服务的主要关注对象”中,“精神病”建议改为“精神病早期干预领域的”。

回应:感谢您对语言表达准确性的建议,我们已经做了相应修改。CHR-P 是精神病早期阶段,改为“精神病早期干预领域的临床研究”更为准确。

见引言第一段:“CHR-P 人群成为精神病早期干预领域的临床研究和服务的主要关注对象”。

意见 8. 引言 1.1.1 中,“然而,最近也有研究发现,阴性症状并没有在追踪过程中改善(Gifford et al., 2021; Leanza et al., 2020)”。请明确,这个结论是描述未转化者,还是混合样本(未转化+转化者)?

回应:感谢您的提问。经过再一次回顾原文,我们确认了 Gifford 等人(2021)和 Leanza 等人(2020)的这两篇研究纳入的是混合样本(未转化者+转化者)。我们也对原文的描述进行了补充,消除潜在的歧义。

见引言 1.1.1:“然而,最近也有研究发现,CHR-P 人群的阴性症状并没有在追踪过程中改善 (Gifford et al., 2021; Leanza et al., 2020)”。

意见 9. 引言 1.1.2 中,“以往元分析结果表明,CHR-P 人群整体功能水平随着时间会有所改善(De Pablo et al., 2021; De Pablo et al., 2022),但上述研究未区分不同功能类型”。另外,这些研究是否区分了转化者和未转化者?转化与否对功能显然是有影响的。

回应:感谢您指出的问题。经过再一次回顾原文,我们发现 De Pablo 等人(2021)的元分析研究并没有区分 CHR-P 人群是否转化为精神病,但是 De Pablo 等人(2022)的元分析区分了转化者和未转化者,结果发现未转化者的功能随时间改善显著改善,且相比于转化者有显著的改善。将两项针对不同人群的结果笼统的描述确实欠妥,我们对原文表述进行了修改。

见引言 1.1.2:“以往元分析结果表明,CHR-P 人群整体功能水平随着时间有所改善(De Pablo et al., 2021),并且相比于转化为精神病的 CHR-P 样本,未转化为精神病的 CHR-P 人群整体功能显著改善(De Pablo et al., 2022)”。

意见 10. 引言 1.1.2 中,“而且,未有干预措施可以有效改善社会功能(Lee et al., 2015),社会功能的损害可能是精神分裂症的潜在预测指标”。Lee 2015 这篇引文并非研究社会功能,

而是社会认知(social cognitive function)，而且也并非随访研究。请确认，在本研究中，纳入 meta 分析的所有研究均为 social function 的研究，而非 social cognitive function.若有 social cognitive function 的被纳入，请剔除

回应：感谢您认真的审阅。很抱歉这是我们的错误，我们已经将这篇引文删除，并对相应文段做了修改。同时，我们也检查了所有纳入元分析的研究，确保都是社会功能的研究，而不是社会认知功能。

见引言 1.1.2：“Velthorst 等人(2018)发现相比于角色功能，CHR-P 人群社会功能随时间改善较小；Cornblatt 等人(2007)也发现社会功能不因治疗和时间的推移而改善，是 CHR-P 个体一种稳定的特征，因此社会功能的损害可能是精神分裂症的潜在预测指标。”

意见 11. 引言 1.2 中，“Carrión 等人(2016)认为基线的阳性症状水平会与更差的病理发展有关”中，“更差的病理发展”建议改为“精神病转化”。

回应：感谢您的建议。这里改为“精神病转化”确实更为恰当。已修改为“Carrión 等人(2016)认为基线的阳性症状水平会与精神病转化有关”。其他相关部分也做了相应的修改。

意见 12. 引言 1.3 中，“CHR-P 人群可以分为减弱的阳性症状组 APS 、短暂间歇性精神病性症状组 BLIPS、遗传风险和衰退综合征组 GRD 和认知障碍组(cognitive disturbances , COGDIS)四个亚组 (Fusar-Poli et al., 2013)。”首次出现缩写时要注明英文原文，下同。

回应：感谢您的指正。这里对于 CHR-P 亚组的划分与引言第一段中提到的 UHR 的纳入标准一致，在之前提到过了英文的全称，故在这里没有重复英文全称

见引言第一段：“UHR 的纳入标准包括出现以下一项或多项：轻微阳性症状(attenuated positive symptoms, APS)、短暂间歇性精神病性症状(brief limited intermittent psychotic symptoms, BLIPS)、遗传风险和衰退综合征(genetic risk and deterioration syndrome, GRD)”。

意见 13. 引言 1.2 中，“CHR-P 人群可以分为减弱的阳性症状组 APS 、短暂间歇性精神病性症状组 BLIPS、遗传风险和衰退综合征组 GRD 和认知障碍组(cognitive disturbances , COGDIS)四个亚组 (Fusar-Poli et al., 2013)”。前三组是使用 SIPS 或 CAARMS 访谈进行诊断确认的，和 COGDIS 组的分组方法和评估工具均不同（UHR 症状标准 VS 基本症状法），不能笼统的放在一起说成四组。

回应：感谢您的指正。将 UHR 和 BS 标准笼统的放在一起确有不妥，根据意见 6，我们也对 CHR-P 亚组进行了重新划分，修改如下。

见引言 1.3：“CHR-P 人群不同亚组也可能是调节变量。根据 UHR 标准，CHR-P 人群可以分为轻微阳性症状(APS)亚组、短暂间歇性精神病性症状(BLIPS)亚组、遗传风险和衰退综合征(GRD)亚组；根据 BS 标准，CHR-P 人群可以分为认知-知觉基本症状(COPER)亚组和认知障碍(CIGDIS)亚组(Fusar-Poli et al., 2013)。APS 亚组指过去一年内出现轻微的精神病性症状，但未达到精神病诊断标准的人群；BLIPS 亚组指过去一年内经历了一两次短暂的、间歇性的精神病性症状的人群；GRD 亚组指有精神病家族史或分裂型人格障碍，并且在过去一年内出现明显的功能损害的人群；COGDIS 亚组指根据 BS 诊断的纳入标准，在过去一段时间内经历一系列认知功能异常的人群。以往元分析的结果显示，CHR-P 样本中更低的 BLIPS 亚型占比与更好的恢复率相关(De Pablo et al., 2022)；与 APS 亚组相比，BLIPS 亚组有更高的精神病转变风险；GRD 亚组不常见，转变为精神病的风险较低；而 COGDIS 亚组比 COPER 亚组更能预测精神病的转化 (Schultze-Lutter et al., 2007)，因此，不同亚型 CHR-P 人群可能会有着不同的临床症状和功能的纵向发展轨迹。”

意见 14. 引言 1.4 中，“全面量化整合 CHR-P 人群病理发展变化的大小和一致性”。病理发展变化的大小？需要更规范的描述。

回应：感谢您的指正。结合审稿人 1 意见 3，已将引言 1.4 进行修改如下。

见引言 1.4：“综上，本研究的目的是：(1)纳入最新有关 CHR-P 人群临床结局纵向发展的研究，量化整合 CHR-P 人群各维度的临床症状和整体功能水平变化，并首次关注社会功能和角色功能的纵向发展，并探讨潜在的调节变量对 CHR-P 人群临床纵向结局的影响；(2)系统综述 CHR-P 人群临床症状对功能的纵向影响。”

意见 15. 方法 2.1 中，“文献排除标准如下：(1)横向研究、综述、元分析、临床病历研究、会议记录、研究方案等”。“横向研究”的规范的描述应该是：横断面研究。

回应：感谢您指出的问题。已将原文中“横向研究”改为“横断面研究”。

意见 16. 方法 2.2 中，“(14)基线和追踪的临床结果”。Clinical outcome 规范的翻译应为“临床结局”，下同。

回应：感谢您对表述规范性的指正。已将原文中“临床结果”的表述均改为“临床结局”。

意见 17. 方法 2.2 中，“追踪流失率 (losses to follow-up)”规范的翻译应为：随访脱落。

回应：感谢您对表述规范性的建议。已将原文中的“追踪流失率”改为“随访脱落”。

意见 18. 方法 2.3.1 中，请陈述一下使用 Hedges' g 而不用 Co'en's d 的理由。

回应：感谢您的提问。本研究采用 Hedges' g 理由如下：(1) 本元分析中纳入的有些研究样本量较小(<30)，采用 Co'en's d 在小样本情况下容易高估效应量，这会导致在合并效应量时

产生偏差。而 Hedges' g 包含一个校正因子 $J = 1 - \frac{3}{4N-9}$ ，通过这个校正因子可以减小由于小样本带来的偏差，从而提供更为准确和保守的效应量估计；(2) 本元分析中纳入的研究的样本量相差较大，Hedges' g 通过校正小样本效应量的高估问题，使得不同研究之间的效应量更具可比性，减少了异质性，能确保效应量估计更加一致和可靠；(3) 本领域以往的元分析研究均使用 Hedges' g (e.g., De Pablo et al., 2021; De Pablo et al., 2022)，便于不同研究之间的比较，以及以往研究的建议(Rosnow et al., 2003)，故本研究选择 Hedges' g 作为效应量指标。

参考文献：Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, 57(3), 221.

我们已将此理由简单陈述在 2.3.1 中：“因 Hedges' g 包含一个样本校正因子 $J = 1 - \frac{3}{4N-9}$ ，可矫正部分研究小样本带来的偏差，更有利于不同研究之间的汇总比较 (Rosnow et al., 2003)，本研究采用 Hedges' g 值作为效应量。”

意见 19. 方法 2.3.1 中，请明确临床症状评估工具，如果是 PANSS 的话，如何界定“瓦解症状”、“情感症状”？用的是 PANSS 5 因子法吗？

回应：感谢您的指正和提问。我们已经在表 1 中补充了临床症状的评估工具，包括 PANSS、SIPS、BPRS、SANS、HAMA、HAM-D、SANS、CAARMS。采用 PANSS 进行症状评估采用的是 PANSS 的五因子模型，包含瓦解症状(disorganization symptoms)和情感症状(emotional distress)的测量(van der Gaag et al., 2006)。

参考文献：van der Gaag, M., Hoffman, T., Remijns, M., Hijman, R., de Haan, L., van Meijel, B., ... & Wiersma, D. (2006). The five-factor model of the Positive and Negative

Syndrome Scale II: a ten-fold cross-validation of a revised model. Schizophrenia research, 85(1-3), 280-287.

我们已在方法 2.3.2 中提及 PANSS5 因子法。

意见 20. 方法 2.3.3 中, 可否增加两个重要的调节变量: (1) 未转化样本 vs. 转化 (或混合样本)? (2) 干预与否

回应: 感谢您认真的审阅。(1) 精神病转化确实是 CHR-P 人群的症状和功能变化的潜在调节变量, 但由于纳入大部分研究并没有在研究中区分转化者和未转化者, 无法分别提取转化者和未转化者的症状和功能的数据。我们在之前计算了每个研究 CHR-P 人群的精神病转化率, 并将其作为调节变量进行分析, 但是我们认为将临床结局的改变作为因变量, 精神病转化率作为调节变量在逻辑上难以自治, 因此没有将其最终作为本研究的调节变量; (2) 本研究的研究目的旨在探究 CHR-P 人群临床结局的自然纵向改变, 因此在纳入研究的时候, 排除了对 CHR-P 人群的干预研究, 虽然排除了一些特定干预的影响, 但许多研究的被试仍来源于大型临床服务, 不能排除临床服务中的一般治疗对 CHR-P 人群的症状和功能的改善作用, 因此也无法明确纳入研究中的样本接受干预的比例。但您提到的这两个变量确实是 CHR-P 人群临床结局重要的潜在变量, 我们也将将其作为我们研究的局限进行说明。

见讨论 4.4: “第三、对于结果异质性的探究存在局限。由于数据量限制, 本研究并没有充分解释临床结局纵向变化的异质性, 未来研究可以纳入更多潜在的调节变量澄清结果的异质性, 例如是否转化为精神病、是否接受干预”。

意见 21. De Pablo 的 Meta 分析最终纳入了 75 篇文章, 为什么当前这篇只有 53 篇? 与 De Palo 的筛选标准有哪些不同? 如有不同, 就不能说“本研究在 De Pablo 等人元分析的基础上, 更新…”。

回应: 感谢您认真的审阅。我们的研究筛选标准和 De Pablo 等人(2021)的筛选标准大体上是相同的 (具体文献纳入标准见原文 2.1), 只是纳入的临床结局有所不同。De Pablo 等人(2021)除了纳入了症状 (轻微阳性症状 (26 篇)、阴性症状 (29 篇)、抑郁症状 (17 篇)) 和功能 (32 篇) 之外, 还纳入了恢复率(Remission, 19 篇); 而本研究并没有纳入恢复率作为临床结局, 仅纳入症状 (轻微阳性症状 (40 篇)、阴性症状 (42 篇)、瓦解症状 (12 篇)、情感症状 (7 篇)、一般精神病理症状 (16 篇)) 和功能 (整体功能 (35 篇)、社会功能 (9 篇)、角色功能 (10 篇))。另外, 为探究 CHR-P 人群症状和功能的随时间的自然纵向变化, 我们排除了明确说明对 CHR-P 人群进行特定干预的研究。故本研究故可能文献总数少于 De Pablo 等人(2021)的研究。我们的筛选标准与 De Pablo 等人(2021)略有不同, 采用“更新”确有不妥, 我们已经删除了“在 De Pablo 等人(2021)研究的基础上更新……”的表述

见讨论 4.4: “本研究的理论和实践意义如下: 第一、本研究纳入最新的研究, 综合分析 CHR-P 人群各维度的临床症状, 并关注了特定的社会功能和角色功能, 更全面地综述 CHR-P 人群的临床结局的纵向发展, 有利于理解 CHR-P 人群精准的预后轨迹和提供更针对性的临床服务。”

意见 22. 表 1 应添加“诊断工具”, 以及“症状评估工具及评分”。另外, 干预与否对症状改善、功能变化, 及精神病转化与否有重要影响, 请列出干预手段 (药物? 心理? 个案管理? 自然观察? 或“未描述”)

回应: 感谢您的指正。我们已经在表 1 中添加了每个研究对于 CHR-P 人群的诊断工具, 以及症状和功能的评估工具。为了探究 CHR-P 人群症状和功能的随时间的自然纵向变化, 我们的元分析纳入的文章排除了仅对 CHR-P 个体进行特定干预的研究, 因此大部分纳入的文

章并没有对 CHR-P 人群是否接受干预进行描述, 仅有少数 RCT 研究, 我们是纳入了空白对照组的数据, 排除了接受特定干预的干预组人群; 但许多研究的被试来源于大型临床服务, 不能排除临床服务中的一般治疗对 CHR-P 人群的症状和功能的改善作用。因此我们认为列出具体的干预手段可能意义不大。

意见 23. 表 1 中前后要一致, 前文用 CHR-P, 表格中不要改成 CHR。表下面要有缩写的注释, 下同。

回应: 感谢您宝贵的意见。我们已经将表 1 中 CHR 修改为 CHR-P, 并且检查文中的所有表格, 补充表格中出现的英文缩写的全称注释。

意见 24. 讨论 4.4 应补充局限性: 未专门探讨“未转化者”, 而转化对症状和功能具有显著影响; 另外, 未探讨“干预类型”对症状改善和功能变化的影响。

回应: 感谢您的建议。我们通过先前的文献阅读发现, 已经有一项元分析研究对未转化者的症状和功能的纵向变化进行了分析(De Pablo et al., 2022), 但是本研究并没有区分社会功能和角色功能, 因此对于未转化者的社会功能和角色功能的纵向发展尚不清楚, 这也是本研究的一个局限, 我们在原文中也加以补充说明

见讨论 4.4: “其次, 虽然 De Pablo 等人(2022)对未转化为精神病的 CHR-P 个体的临床结局纵向发展进行了元分析, 但未来还需要专门探讨未转化者的社会功能和角色功能的发展。”

对于干预类型来说, 以往有元分析和综述表明, 目前还没有特定干预能够有效缓解 CHR-P 人群的症状和功能(De Pablo et al., 2021), 这提示未来还需要更多实证研究去开发对于 CHR-P 人群症状和功能的有效干预, 因此本研究并没有将干预类型作为调节变量进行分析。

参考文献:

De Pablo, G. S., Soardo, L., Cabras, A., Pereira, J., Kaur, S., Besana, F., ... & Fusar-Poli, P. (2022). Clinical outcomes in individuals at clinical high risk of psychosis who do not transition to psychosis: a meta-analysis. *Epidemiology and psychiatric sciences*, 31, e9.

De Pablo, G. S., Besana, F., Arienti, V., Catalan, A., Vaquerizo-Serrano, J., Cabras, A., Pereira, J., Soardo, L., Coronelli, F., & Kaur, S. (2021). Longitudinal outcome of attenuated positive symptoms, negative symptoms, functioning and remission in people at clinical high risk for psychosis: A meta-analysis. *EClinicalMedicine*, 36.

第二轮

审稿人 1 意见:

作者进行了较为全面的修改, 文章的质量有了一定程度的提升, 但仍有两个问题:

意见 1. 文献质量评估的编码一致性为 $Kappa = 0.32$, 如此低的一致性, 与之相关的分析结果的可信度和稳定性会受到质疑。

回应: 感谢您的意见。 $Kappa = 0.32$ 是对文献质量评估的一致性, 这个系数说明几位研究者对文献质量的主观评估可能存在不一致性, 质量评估均值见表 1。但文献质量评估并不涉及元分析结果效应量计算所需的数据编码, 因此并不影响元分析主要结果的可信度和稳定性; 文献数据编码一致性在 0.63 到 0.99 之间, 而且几位研究者对于编码不一致的内容经过协商之后均达成一致, 对于最终的元分析结果的可信度和稳定性并无影响。具体见 2.2 文献编码与质量评估。

意见 2. 作者遗漏了“认知-知觉基本症状(COPER)亚组”的界定。

回应:感谢您宝贵的指正。已在引言 1.3 中加入对于 COPER 亚组的界定。

见引言 1.3:“COGDIS 亚组指根据 BS 的诊断标准,在过去三个月内至少出现九种基本症状(主要涉及认知功能的显著变化)中的两种症状的人群;COPER 亚组与 COGDIS 亚组诊断涉及的症状有重叠,指在过去三个月内至少出现十种基本症状(主要涉及认知和感知方面的异常)中的一种症状的人群。”

审稿人 2 意见:

该作者已经对我的问题给予很好回应,同意发表。

第三轮

审稿人 1 意见: 作者做出了恰当的回应。同意发表。

编委 1 意见:

经多轮修改,该稿件质量有较大的提升。几点注意:

核心问题: 本研究是 De Pablo 等人(2021)元分析的更新,加入了 14 篇新文献,但研究的结果与其较为一致,那么重要的新发展是什么,需要有引言和讨论中有重点论述。

回应:感谢您宝贵的意见。

这个元分析,除了在 De Pablo 等人(2021)元分析的基础上增加了 14 篇文献,探讨了 CHR-P 人群症状和功能结局的纵向变化以外,还在两个方面有新进展:

首先,是对功能做了区分,见引言 1.1.2 的重点论述:“……但上述研究未区分不同功能类型。社会功能(social functioning)和角色功能(role functioning)的损害是 CHR-P 人群的核心特征(Meyer et al., 2014)。”，“两种不同的功能损害随时间改善程度不同”，“因此，本研究首次关注 CHR-P 人群各种功能的纵向改变并进行定量整合，旨在澄清 CHR-P 人群特定功能的发展轨迹。”

其次,系统阐述 CHR-P 人群症状对特定功能的影响。见引言 1.2“CHR-P 人群的症状对功能的影响”的重点论述。

引言 1.4,“研究目的与研究问题”对这两点进行了小结:“本研究在 De Pablo 等人(2021)元分析的基础上进行文献更新并区分功能,同时系统阐述 CHR-P 人群症状对特定功能的影响。”

对于这两点的新进展,见结果 3.3.2 和表 3。

针对第一点,讨论 4.1 的第四段做了不同功能纵向改变的重点讨论。

针对第二点,讨论 4.2 对症状与不同功能纵向关系做了重点讨论。

讨论 4.4 也已对研究意义做了详尽的表述“本研究使用三水平元分析的方法对 CHR-P 人群的临床症状和功能的纵向发展进行全面的量化分析,并系统综述了基线症状对功能的预测作用。本研究的理论和实践意义如下:第一、本研究纳入最新的研究,综合分析 CHR-P 人群各维度的临床症状,并关注了特定的社会功能和角色功能,更全面地综述 CHR-P 人群的临床结局的纵向发展,有利于理解 CHR-P 人群精准的预后轨迹和提供更针对性的临床服务。第二、本研究首次综述了 CHR-P 人群各维度基线症状对随访特定功能的预测作用,拓展了症状和研究类型(Devoe et al., 2020),揭示了阴性症状与 CHR-P 人群社会功能的重要关

系,为 CHR-P 人群的社会功能干预提出新思路。第三、本研究揭示年龄、受教育程度、CHR-P 亚型对于功能改变的显著调节作用,这为 CHR-P 人群症状与功能改变异质性背后可能的原因提供了实证依据。”

其他问题:

意见 1. 论文摘要需修改,以符合“进展”的要求。

回应:感谢您的建议。我们已经参考“投稿指南”对摘要进行修改,尽量精简字数,以达到《心理科学进展》的要求。

意见 2. Comprehensive Meta-Analysis 需标注版本,并引用;metafor 包同理。

回应:感谢您的指正。我们已在方法 2.3.2 中加入 Comprehensive Meta-Analysis 的版本和 metafor 包的引用表述:

见方法 2.3.2:“本研究首先采用软件 Comprehensive Meta-Analysis Version 3.0 (CMA)软件将提取的临床结局计算出 Hedge' s g 值.....之后采用 R X64 4.3.2-win 版本的 metafor 包 (Viechtbauer, 2010),参照 Assink 和 Wibbelink (2016)的教程改编 R 语法,进行临床症状和功能的三水平元分析。最后采用 CMA 软件进行症状与功能相关系数的两水平元分析。”

补充参考文献: Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, 36(3), 1-48.

意见 3. 论文中英文字体需调整。

回应:感谢您的指正。我们已经按照投稿指南中的字体要求对论文字体进行了修改。

意见 4. 图 1 PRISMA 文献筛选流程图,格式乱了,建议另附。

回应:感谢您的建议。已进行流程图格式检查,并将文献筛选流程图以 300dpi 图片形式附在后面。

意见 5. 统计符号前后各有一个半角状态的空格,比如基线的阴性症状($r=-0.229$; $p=0.003$, $N=6$, $n=526$),全文校对。

回应:感谢您的指正。我们已经进行了全文校对并修改。

意见 6. 有些涉及统计方法的重要文献,不要间接引用。比如“本研究还纳入发表年份和文章质量作为调节变量,进一步检验发表偏倚的严重性(陈必忠 等, 2023)。”

回应:感谢您的意见,抱歉引起您的误解。引言 1.3 这句话想说明的是:本研究参考陈必忠(2023)等人的做法,纳入了一些研究特征变量作为调节变量,并不涉及统计方法的引用。为了表述得更加明白,我们将这句话修改为:“本研究还参照以往研究(陈必忠 等, 2023)的做法,纳入发表年份和文章质量作为调节变量,进一步检验发表偏倚的严重性。”

陈必忠(2023)等人的原文表述为:“基于结果稳健性的考虑,除了以上调节变量外,我们还纳入了研究特征变量(发表年份、文章语言、发表状态、文章质量),进一步检验选择性报告偏差的严重性。”

陈必忠,黄璇,牛更枫,孙晓军, & 蔡志慧. (2023). 学步期至青年期社交焦虑的发展轨迹和稳定性:一项基于纵向研究的三水平元分析. *心理学报*, 55(10), 1637-1652.

意见 7. 研究多处重复标注量表的英文及缩写。

回应:感谢您的指正。我们已经将正文中重复的量表的英文及缩写删除。

见方法 2.1: “(3)研究对象为 CHR-P 人群, 且对于 CHR-P 人群的判定标准出自己有成熟的心理测量工具(如高危精神状态综合评估、前驱综合征定式检查 (杨学智 et al., 2023))。”

见讨论 4.4: “本研究仍存在以下局限: 第一、对于变量的测量存在局限。本研究纳入的文献对于阴性症状的测量主要采用阳性与阴性症状量表.....可以更精确地测量阴性症状的二因素本质: 动机与愉悦、表达减少 (Forbes et al., 2010; Kirkpatrick et al., 2011; Li et al., 2022).....本研究对于功能的测量大多采用功能大体评估量表和社会和职业功能评估量表 (Social and Occupational Functioning Assessment Scale, SOFAS)进行测量, 反映的是较为客观的功能水平。”

意见 8. 参考文献需仔细校对, 比如, 中文最后一位前不需要&。英文多处缺卷或期号。

回应: 感谢您的指正。我们已经对参考文献进行了校正。部分文献没有期号的信息, 故没有标注期号。

意见 9. 英文摘要需润色, 有些语法错误, 比如动词使用不当: Clarifying the changes in symptoms and functioning in this population and their underlying relationships is beneficial for early prevention of psychosis. 再比如, 有些连词使用不当: ...to comprehensively integrate the longitudinal changes in various clinical symptoms and functioning of CHR-P individuals, and to explore the predictive role of baseline symptoms on follow-up functioning。

回应: 感谢您的指正。我们已对英文摘要进行修改、润色。

编委 2 意见:

文章结构、逻辑性和文笔通顺性良好, 建议发表。其中, 有一点建议: 由于本文所采用的研究随访 2-3 年, 高危症状逐渐改善。尽管不能说明这种改善在不同的干预方法上的效果差异, 但在方法中和文章结尾的局限中, 需要介绍一下整体文献采用干预方法的使用情况, 并指出未来的研究方向。

回应: 感谢您的指正。为了探究 CHR-P 人群临床结局的自然变化轨迹, 我们的研究剔除了明确说明对 CHR-P 人群采用特定干预的研究, 因此本研究纳入的所有研究并没有包含明确说明使用特定干预方法的研究。但由于很多研究的被试来源是一些大型的临床服务项目, 在其中可能会对 CHR-P 人群做一些干预和治疗, 我们在讨论 4.1 第三段说明了这种临床结局的改善不能排除临床服务中接受到的干预的影响。我们增加了一些对临床服务中 CHR-P 人群可能接收到的干预手段的描述。

见讨论 4.1: “本研究发现 CHR-P 人群轻微阳性症状和情感症状的较大改善, 阴性症状的改善较小, 与先前 De Pablo 等人(2021)元分析结果模式相一致。我们的元分析纳入的文章排除了仅对 CHR-P 个体进行特定干预的研究, 以期能够探究 CHR-P 人群症状和功能的随时间的自然纵向变化, 但许多研究的被试仍来源于大型临床服务, 不能排除临床服务给予的特定干预(如认知行为治疗、家庭治疗等)和一般治疗(如药物治疗、心理健康教育等)对 CHR-P 人群的症状和功能的改善作用(De Pablo et al., 2021)。”

主编意见: 同意发表。