

# 巢式病例对照研究的方法学进展及其在临床研究中的应用



陆春利<sup>1,2</sup>, 夏婷婷<sup>1,2</sup>, 邬娜<sup>1,2</sup>, 向颖<sup>1,2</sup>, 李亚斐<sup>1,2</sup>

1. 陆军军医大学 (第三军医大学) 军事预防医学系军队流行病学教研室 (重庆 400038)
2. 陆军军医大学 (第三军医大学) 临床流行病学和循证医学中心 (重庆 400038)

**【摘要】**巢式病例对照研究 (NCCS) 是传统病例对照研究的一种衍生类型, 作为观察性研究方法, 其具有诸多优势, 可为临床研究提供一种高效、经济且可靠的方法。本文从临床流行病学方法学角度, 阐述 NCCS 的进展及其在临床研究中的应用, 全面介绍 NCCS 的原理、分类、应用条件、研究设计、实施要点及优缺点。尤其是针对目前在设计和实施过程中研究类型和对照组选择上存在的关键问题, 本文提出了合理化的建议, 以期加深临床医生对 NCCS 的正确认识并规范其在临床研究中的应用, 为提高 NCCS 证据的科学性和可靠性提供方法学参考。

**【关键词】**巢式病例对照研究; 分类标准; 发病密度; 匹配; 研究设计

**【中图分类号】**R181.2+3 **【文献标识码】**A

## Methodological advances and application in clinical research of nested case-control studies

LU Chunli<sup>1,2</sup>, XIA Tingting<sup>1,2</sup>, WU Na<sup>1,2</sup>, XIANG Ying<sup>1,2</sup>, LI Yafei<sup>1,2</sup>

1. Department of Military Epidemiology, Faculty of Military Preventive Medicine, Army Medical University (Third Military Medical University), Chongqing 400038, China

2. Center for Clinical Epidemiology and Evidence-Based Medicine, Army Medical University (Third Military Medical University), Chongqing 400038, China

Corresponding author: LI Yafei, Email: liyafei2008@tmmu.edu.cn

**【Abstract】**Nested case-control study (NCCS) is a derived type of traditional case-control study. As an observational research method, it has many advantages and provides an efficient, economic and reliable method for clinical research. This article elaborates on the progress of NCCS and their applications in clinical research from the perspective of clinical epidemiological methodology. The principles, classifications, application conditions, research design, implementation points, as well as advantages and disadvantages of NCCS are comprehensively introduced. Especially regarding the key issues in selecting research types and control groups for current design and implementation, this article has put forward rational suggestions. This article aims to deepen clinical doctors' correct understanding and standardized application in clinical research of NCCS, and provide methodological reference for improving the scientific and reliable evidence of NCCS.

DOI: 10.12173/j.issn.1004-5511.202603034

基金项目: “癌症、心脑血管、呼吸和代谢性疾病防治研究” 国家科技重大专项 (2025ZD0551600)

通信作者: 李亚斐, 博士, 教授, 博士研究生导师, Email: liyafei2008@tmmu.edu.cn

**【Keywords】** Nested case-control study; Classification standard; Incidence density; Matching; Research design

巢式病例对照研究 (nested case-control study, NCCS) 是将队列研究与病例对照研究相结合的一种研究方法, 该方法基于事先确定好的队列, 在随访观察的基础上, 采用病例对照研究的设计思路进行研究和分析, 研究对象均选自于同一队列。当前我国临床流行病学研究逐渐重视高质量队列的建立, 基于社区的自然人队列和基于医院的专病队列日益增多<sup>[1-3]</sup>。NCCS可以针对这些队列资源深度复用, 尤其在当前精准医学由“关联发现”向“机制驱动”转型的背景下, NCCS通过高效利用队列资源和生物样本、强化因果推断、支撑多组学整合, 正在成为解析复杂疾病分子机制、推进生物标志物临床转化的重要方法之一<sup>[4-5]</sup>。

然而, 目前临床医生对NCCS的方法学仍存在诸多认知误区。例如, 其对研究类型的名称与定义认识模糊, 导致对各类型的设计要点阐述不明; 高估或低估该研究类型的科研价值等。为帮助科研工作者正确理解和规范应用NCCS, 本文从临床流行病学方法学角度出发, 明确其定义, 更新并提出了新的分类标准及其优缺点, 以减少该研究类型分类中存在的混乱; 同时, 明确不同分类的适用场景, 以及研究设计与实施要点, 旨在推动科研工作者对NCCS形成统一理解, 为提高该方法应用的科学性和规范性提供重要的方法学参考。

## 1 巢式病例对照研究设计原理的发展

NCCS方法最早由美国流行病学家Mantel于1973年提出, 命名为综合性病例对照研究 (synthetic case-control study)。1982年, 该方法被正式定名为NCCS, 又称套叠式、嵌入式病例对照研究或队列内病例对照研究<sup>[6-7]</sup>。其名称反映了设计核心, 即将病例对照研究“嵌套”在队列研究的设计框架内。

该方法的设计原理为研究者首先需确定某队列人群为研究对象, 收集队列中研究对象的信息并进行随访, 以队列随访中发生某疾病 (或具有某种健康状态) 的研究对象作为病例组, 尚未发生该疾病 (或不具有该种健康状态) 的研究对象

作为对照组, 通过比较病例组与对照组中研究因素的暴露比例差异有无统计学意义, 判断暴露因素与疾病 (或健康状态) 的关联性和关联强度。综上, NCCS是一种观察性的分析性流行病学研究方法<sup>[6-7]</sup>, 其通过方法学创新, 平衡了队列研究的严谨性与病例对照研究的成本效益, 现已成为现代流行病学和临床研究的重要工具。

## 2 巢式病例对照研究分类存在的问题和建议

目前, NCCS缺乏统一的分类标准, 研究者对于分类与应用方面的观点存在分歧。部分文献未明确标注NCCS设计, 或未详细说明对照的选择方法; 少数研究甚至将传统病例对照研究或队列研究误标为NCCS。分类不规范已对学术交流和研究结果的解读造成不利影响。

现有分类主要包括两类: 按队列建立与随访时序分为前瞻性NCCS (prospective nested case-control study, PNCCS) 与回顾性NCCS (retrospective nested case-control study, RNCCS), 该分类在教材与文献中较为常用; 按对照的选择方法可分为匹配与不匹配NCCS<sup>[1-3]</sup>。

在分析性流行病学研究中, 对照的选择是影响研究真实性的关键环节, NCCS亦如此。最新版流行病学词典将NCCS中对照的选择方法分为累积抽样和密度抽样<sup>[9]</sup>。基于此, 本文建议将NCCS分为累积NCCS (cumulative nested case-control study) 和发病密度NCCS (incidence-density nested case-control study) 两类。两种类型的对照选择均可采用与病例匹配或者不匹配的方式, 资料的收集也可采用前瞻性或回顾性的方式。

累积NCCS是在随访的某一时间点或结束时, 将队列内全体或部分发病 (或具有某种健康状态) 人群作为病例组, 从研究期内未发病者中选取对照, 可采用匹配或非匹配方式 (图1-A)。该方法的对照明确、选择简便, 但未考虑不同个体风险时间与发病概率差异。发病密度NCCS则采用密度抽样选取对照, 需同时考虑观察人时, 通过随机抽样或匹配设计, 保证病例与对照在观察时间、发病风险层面具有可比性 (图1-B); 该设

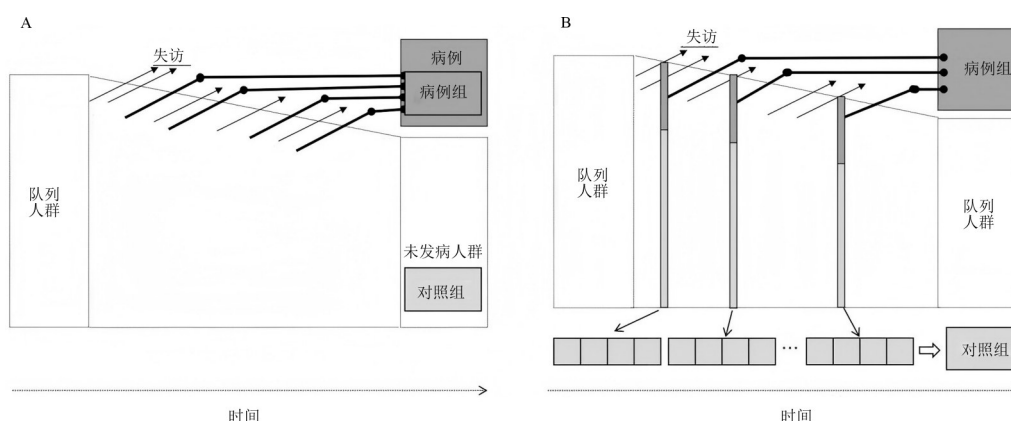


图1 累积/发病密度巢式病例对照研究的设计原理

Figure 1. Design principle of cumulative/incidence-density nested case-control study

注：A. 累积巢式病例对照研究；B. 发病密度巢式病例对照研究。

计的组间可比性更强、统计效率更高，应用更普遍。总之，二者在适用场景、实施特点及统计效率上存在明显差异，为便于清晰对比，现将两类设计要点汇总于表1。

上述两种设计类型均可采用前瞻性或回顾性的资料收集方式。其中，PNCCS是指从当前时间点开始对队列进行前瞻性随访，在随访过程中收集新发病例，数据来源于前瞻性随访，因此回忆偏倚较小、可信度较高，但研究周期较长<sup>[10]</sup>。RNCCS则利用已建成队列的历史基线与随访资料，研究结局已发生，可快速开展，效率更高、更省时，但对历史资料的完整性要求更高<sup>[7]</sup>。

综上，NCCS可从2个维度进行分类：①按对照抽样策略，分为累积抽样与发病密度抽样；②按资料收集时序，分为前瞻性与回顾性。两类分类

标准并非相互排斥，而是可交叉组合形成4种规范设计类型。临床研究人员在撰写方案与报告时，应明确标注设计类型与资料收集方式，建议采用四分类规范命名，即前瞻性累积、回顾性累积、前瞻性发病密度、回顾性发病密度NCCS。该命名体系有助于规范研究设计、统一报告标准、提升研究可重复性，并促进学术交流。

### 3 巢式病例对照研究在临床研究中的适用场景

当前瞻性队列研究的随访开始后出现了新的研究假设，而研究因素未能在所有队列成员中测量，或对队列中每个成员进行暴露因素测量的成本过高时，可采用NCCS，以较低成本快速验证新的研究假设。当队列研究预期结局发生率较低

表1 累积/发病密度巢式病例对照研究的比较

Table 1. Comparison of cumulative/incidence-density nested case-control study

对比维度	累积巢式病例对照研究	发病密度巢式病例对照研究
适用场景	1. 队列随访时间较短、失访较少； 2. 研究实施追求简便快捷	1. 随访时间长、个体观察人时差异大； 2. 暴露因素随时间变化； 3. 常见病或发生率较高的结局； 4. 需准确估计发病风险与效应值
优点	1. 对照选择简单直观、实施方便； 2. 设计易于理解与操作； 3. 对照人群明确，工作量较小	1. 考虑观察人时，组间可比性更好； 2. 统计效率更高，结果更可靠； 3. 更适合暴露随时间变化的场景
缺点	1. 未考虑不同个体观察时间与发病风险差异； 2. 组间可比性相对较差； 3. 难以处理时变暴露，易产生偏倚	1. 设计与抽样更复杂，需计算人时； 2. 对资料完整性要求更高； 3. 同一对照可能被多次抽取
统计分析差异	1. 采用传统病例对照分析思路； 2. 不考虑人时，难以精确校正时间差异	1. 基于人时构建风险集，可估计发病密度； 2. 常用条件Logistic回归； 3. 效应估计更准确，检验效能更高

时,也可选择NCCS,通过增加匹配对照组的人数,提高统计检验效能。在研究生物标志物与疾病(或健康状态)的关联时,常采用NCCS,可直接调用已积累的生物样本(如血液、组织库)和标准化表型数据(如临床指标、影像数据),在筛选出病例并匹配对照后,对选定样本进行全基因组测序、蛋白质组或代谢组分析,从而大幅降低高昂的组学检测成本<sup>[11-12]</sup>。此外,在临床试验研究中,若感兴趣的暴露因素未经过随机分配,可将全部试验人群视为一个队列,在探讨该暴露因素与结局的关联时,也可通过NCCS对试验数据进行二次分析<sup>[13]</sup>。

值得注意的是,累积NCCS和发病密度NCCS分别适用于不同的研究场景。累积NCCS更适用于队列随访时间较短、失访率低,且研究实施追求简便快捷的场景;而发病密度NCCS则更适合随访时间长、个体观察人时差异大、暴露因素随时间变化、结局发生率较高或为常见病,且需要准确估计发病风险与效应值的研究。研究者应根据队列类型、暴露特征、疾病类型和研究资源,选择合适的研究设计类型。在明确适合开展的NCCS类型后,研究者下一步应重点关注研究的设计与实施要点。

## 4 巢式病例对照研究设计与实施的核心要点

良好的设计是研究成功实施的前提和保障。了解NCCS的设计与实施要点,有助于提高研究的科学性、控制混杂因素、增强结果的可信度与推广性。同时也是研究者规范研究设计、避免常见错误的必要前提。以下具体介绍NCCS的设计与实施6个方面的要点。

### 4.1 研究队列

NCCS基于明确的队列人群,且该人群在进入队列时未发生所研究的疾病。该人群应具备测量暴露指标的条件,并有一定的发病风险,且其数量能够满足NCCS的样本量需求。在临床研究中,也可基于医院人群构建队列,例如医院的体检人群、门诊或住院人群、特定疾病登记人群

等。若研究某罕见病或结局的相关因素,可能需要选择多家医院联合组成的较大规模患者队列,以确保有足够的样本量。

### 4.2 暴露因素和结局指标

研究可对多个感兴趣的暴露因素进行广泛探索,也可聚焦于某一明确的暴露因素。暴露因素的确定需具备充分的科学依据,通常来源于描述性研究中的病因线索或病例对照研究中的初步病因假设。研究应明确规定暴露因素的定义、测量方法与时间、暴露剂量及持续时间。NCCS属于观察性研究,所有的暴露因素均为基于真实世界人群的自然暴露,并非研究者施加的干预。

研究结局是研究者期望观察到的事件。在病因研究中,结局通常指疾病的发生;在疾病预后或干预措施效果评价的研究中,结局可能是死亡、疾病进展等不良事件。研究结局的确定应基于研究目的、时间、财力和人力等因素进行全面客观的考量。结局的判定需遵循明确且统一的标准,并在研究过程中严格执行。

### 4.3 样本量

NCCS的样本量估算方法与经典的病例对照研究一致。对于非匹配设计,样本量估算需要确定显著性水平( $\alpha$ )、检验效能( $1-\beta$ )、待研究暴露因素在病例组的暴露率( $P_1$ )和对照组的暴露率( $P_0$ )。而对于采用匹配设计的NCCS,样本量估算还需确定匹配比例。详细的样本量估算公式可参见相关流行病学书籍<sup>[14]</sup>,也可采用PASS软件输入相关参数后进行计算。

针对非匹配设计与1:1个体匹配设计两种常用情形,基本的样本量公式如下:①非匹配NCCS可通过公式1估算所需病例数。公式1中, $P_1$ 为病例组暴露率, $P_0$ 为对照组暴露率; $\bar{P}=(P_1+P_0)/2$ ;  $Z_{1-\alpha/2}$ 与 $Z_{1-\beta}$ 分别为标准正态分布下双侧检验与检验效能对应的分位数。对照组样本量通常按病例数的1:1至1:4比例确定。②对于1:1匹配的NCCS,可采用公式2估算配对设计所需的样本量。其中 $m$ 为不一致对子数,再通过公式3估算所需的总对子数( $M$ )。在1:1匹配设计下,每1例病例匹配1例对照,总样本量为 $2M$ 。

$$n = \frac{\left[ Z_{1-\alpha/2} \sqrt{2\bar{P}(1-\bar{P})} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_0(1-P_0)} \right]^2}{(P_1 - P_0)^2} \quad \text{公式1}$$

$$m = \frac{\left[ Z_{1-\alpha/2} \sqrt{2\bar{P}(1-\bar{P})} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_0(1-P_0)} \right]^2}{(P_1 - P_0)^2} \tag{公式2}$$

$$M = \frac{m}{P_1(1-P_0) + P_0(1-P_1)} \tag{公式3}$$

#### 4.4 病例组和对照组的选择

病例组和对照组选择的要点在于明确纳入与排除标准，确保两组的代表性与可比性，并避免选择偏倚和过度匹配。

NCCS 的病例组来自队列中发生目标结局的人群，需事先制定严格的纳入与排除标准。累积 NCCS 是在某一规定时点或随访结束时，从累积发生结局的病例中，纳入符合条件的研究对象；而发病密度 NCCS 则是研究对象一旦符合条件并发生结局时即可纳入病例组，同时选择符合条件的对照组。

对照组的主要作用是病例组提供比较基准，帮助研究者评估暴露因素与结局的关联性，故对照组的选择是研究设计的关键环节，选择不当可能会导致偏倚、混杂或错误结论。累积 NCCS 可采用非匹配或匹配的方式选择对照人群。常用的匹配方法是依据常见或重要的混杂因素，将所有病例和对照以 1:1 或 1:n 的比例进行匹配。在队列研究中，匹配的对照数目一般不超过 4 个，因为超过 4 个后，其提升的统计检验效能有限，同时会显著增加研究成本和工作量，降低研究效率<sup>[15]</sup>。然而，在 NCCS 中，由于依附队列研究已收集的资料，在暴露和结局资料可直接获取的情况下，增加匹配数目并不会额外增加研究本身的成本，且在发病人数较少的情况下，纳入更多对照人群反而有助于提高研究结论在研究人群中的内部真实性和代表性，但在暴露和结局资料尚未测量时，匹配数目的选择仍需考虑研究的成本<sup>[16]</sup>。

发病密度 NCCS 常用的对照选择方式是发病密度抽样，即根据队列中每个病例出现结局的时间点，在相同时长的随访期内，将未发病（或未发生某种健康状况）、未死亡、未失访的队列人群作为对照的候选，再按照规定的匹配方式从候选人群中确定对照。这一候选人群在其他研究中也称为风险集或危险集<sup>[8]</sup>。在具有相同风险集的人群中选择对照，能有效减少因 2 组研究对象纳入时间不一致所导致的偏倚。例如，疾病并发

症的发生风险会随时间延长而逐渐增加，采用发病密度抽样可以控制时间这一混杂因素对因果推断的影响。该设计的优势在于病例与对照拥有相同的观察人时基础，风险集由事件发生时刻仍处于随访且存在风险的个体构成，对照组的暴露比例能准确反映目标风险人群的瞬时暴露分布。因此，由此得到的比值比（odds ratio, OR）可直接解释为队列研究的发病密度比（incidence rate ratio, IRR）。同时，基于风险集抽样的条件 Logistic 回归在数学上等价于 Cox 比例风险模型的偏似然估计<sup>[17]</sup>。在比例风险假设成立的前提下，理论上满足 OR=IRR= 风险比（hazard ratio, HR），因此 NCCS 获得的 OR 可作为 HR 的无偏估计。不过该设计也相应增加了研究实施与统计分析的复杂性。

#### 4.5 资料收集与随访

研究资料来源于队列研究中基线与随访中的暴露、混杂因素和结局资料。暴露因素的资料应尽可能全面收集，包括暴露的时间、频率、强度等。如需检测样本，应抽取队列中相应研究对象的生物学样本（如血液、尿液或组织等）。混杂因素的相关资料也需详细收集，通常包括且不限于研究对象的基础人口学资料（如姓名、年龄、性别、职业、民族、文化程度、婚姻状况、住址）、生活方式、家族史和疾病史等。PNCCS 需对队列人群进行一定时间的随访，以获取研究结局指标。随访时间的长短取决于研究结局发生的时间、所需样本量的大小和研究成本等<sup>[18]</sup>。

#### 4.6 统计分析

NCCS 的统计分析主要包括 3 个部分，即描述性分析、差异性分析及暴露与结局的关联分析。描述性分析主要是对纳入研究的变量进行统计学描述，分类变量采用频数和百分比描述；连续性变量若符合正态分布，采用均值和标准差描述，若不符合正态分布，则采用中位数和四分位间距描述。差异性分析用于比较暴露信息和基线资料在组间的差异，分类变量使用  $\chi^2$  检验；连续性变量在满足正态分布且方差齐性时，采用 *t* 检验或

单因素方差分析,不满足时采用秩和检验。暴露与结局的关联分析通常采用回归分析方法,探讨1个或多个暴露因素与结局的关联,并评估关联强度。常用方法包括 Logistic 回归或 Cox 比例风险回归模型等<sup>[19-20]</sup>。具体计算方法可参见相关流行病学书籍<sup>[14]</sup>。在统计分析阶段,控制混杂因素的方法包括分层分析、多因素分析或倾向性评分等<sup>[21]</sup>。此外,为检验研究结果的稳健性,还可以开展敏感性分析<sup>[22-23]</sup>。

## 5 巢式病例对照研究的优点与局限性

NCCS 依托前瞻性队列设计,在兼顾研究效率与因果推断可靠性方面具有突出优势,尤其适用于资源有限、检测成本较高或结局较为罕见的研究场景,能有效平衡成本与科学性,因此在流行病学与临床研究中应用广泛。与此同时,该设计也存在一定局限性,例如对前期队列建设要求较高、研究设计与实施流程相对复杂,且由于仅对部分样本进行分析,可能存在一定的统计检验效能损失与信息利用不充分等问题(表2)。

## 6 对临床研究者的实践建议

NCCS 在兼具成本效益与因果推断价值的同时,也对研究设计与实施质量提出了更高要求。基于上述特点,可为临床研究者提供以下针对性实践建议。

在开展临床研究,尤其是生物标志物、慢性病及罕见结局相关研究时,研究者应优先评估现

有队列的数据质量,包括随访完整性、样本保存状况、暴露测量规范性及结局诊断一致性等,确保研究基础可靠。在设计阶段,需根据研究目的选择适宜的抽样与匹配方案:①针对生存结局与时间依赖性暴露,建议采用发病密度抽样与风险集设计,使 OR 可近似为 HR,提升结果的解释效力;②对于探索性关联分析,可采用非匹配或简单匹配设计以简化实施流程。在资源有限的前提下,可合理设置对照比例,在控制检测成本与样本量的同时保障统计检验效能。实施过程中应严格规范病例与对照的选取流程,控制选择偏倚与信息偏倚;统计分析阶段需根据设计类型选用条件 Logistic 回归等恰当模型,并重视混杂因素控制与敏感性分析。研究者还应客观认识该设计在统计效能与信息利用上的局限性,避免结果过度解读。

临床研究者在选择研究设计时,需明确 NCCS 与经典队列研究、传统病例对照研究之间的核心差异,结合研究目的与资源条件科学决策。经典队列研究虽结果全面、统计检验效能高,能直接计算发病率与 HR,但研究周期长、人力物力成本高,更适用于暴露明确、样本量易获取、随访难度小的研究。传统病例对照研究实施快捷、成本低廉,无需长期随访,适合缺乏队列资源时的快速探索性研究,但暴露与结局的时间顺序难以确定,易受回忆偏倚、入院率偏倚的影响。NCCS 依托前瞻性队列,兼具因果推断可靠、成本效益突出的优势,仅需检测病例与部分

表2 巢式病例对照研究的优点与局限性

Table 2. Advantages and limitations of nested case-control study

优点	局限性
1. 节约研究资源与时间:研究基于已建立的前瞻性队列开展,可显著缩短研究周期,降低人力与物力成本;尤其在生物标志物检测中,仅需对病例与部分对照样本进行检测,能大幅节约实验检测费用;	1. 研究设计与实施要求较高:该设计需依托已建立的前瞻性队列,在设计、随访及样本管理等环节需制定更为细致的方案,对研究实施与质量控制的要求相对更高;
2. 检验假设能力强:暴露资料在研究结局发生前已收集,符合因果推断的时间顺序,相较于病例对照研究具有较强的假设检验能力;	2. 存在观察性研究固有的偏倚风险:作为观察性研究,仍可能受到未知混杂因素或未测量混杂因素的影响,从而干扰暴露与结局关联的准确推断;
3. 样本代表性和组间可比性较好:研究对象比病例对照研究更能代表源人群;病例和对照来自同一队列,尤其是采用发病密度抽样匹配对照时,2组间的可比性较好;	3. 研究质量高度依赖队列基础:研究结果的可靠性受原有队列研究质量制约,包括生物样本保存状况、长期随访完整性、暴露测量准确性等;若基线暴露测量存在误差,或疾病诊断标准在随访期间发生变化,仍可能引入明显的信息偏倚;
4. 适用于少见或罕见结局的研究:通过长期随访队列,可逐步累积足够数量的罕见事件病例;同时可通过增加每名病例匹配的对照数量,进一步提升统计检验效能	4. 存在一定的统计检验效能与信息损失:由于仅对队列中部分病例与对照样本进行分析,其统计检验效能略低于直接对全队列进行分析,可能在一定程度上遗漏较弱或次要的关联信号

对照样本,尤其适合生物标志物、组学等检测成本较高的研究;采用发病密度抽样时,OR可直接作为HR的无偏估计,无需罕见病假设,是资源有限情况下的优选设计。通过科学评估数据基础、合理选择设计方案、规范实施分析流程,并充分理解不同研究设计的适用边界,临床研究者可在保证研究质量的前提下,实现资源优化配置与因果推断可靠性的平衡。

## 7 展望

NCCS凭借因果时序清晰、成本效益突出、实施效率较高等特点,在当前医疗大数据与精准医学快速发展的时代背景下,为大规模队列资源的高效利用提供了重要设计思路,也为复杂健康问题的研究开辟了更灵活可行的路径。

在精准医学与多组学研究中,该设计具有独特的应用潜力。未来可依托高质量前瞻性队列,结合基因组学、蛋白质组学、代谢组学等多维组学技术,在控制检测成本的同时高效筛选与疾病早期预警、预后判断及治疗反应相关的生物标志物。通过整合个体遗传特征、免疫状态、环境暴露等多层面信息,进一步解析治疗应答的异质性,为个体化干预和精准诊疗提供更具针对性的科学依据。但同时也应认识到,高维组学数据伴随的多重检验问题易导致假阳性结果膨胀,如何通过多重校正、惩罚回归、机器学习等策略提升标志物筛选的稳定性与可靠性,将是该领域亟待解决的重要问题<sup>[24]</sup>。

在数据科学与多学科交叉趋势下,NCCS也将迎来更广阔的应用空间。随着医疗信息化水平的提升、多中心数据共享机制的逐步完善,研究者可基于更大规模、更具代表性的队列开展设计,联合生物信息学、机器学习等方法深度挖掘临床表型、暴露因素与疾病结局之间的复杂关联。同时,该设计可与倾向性评分、因果森林等方法相结合,并进一步引入逆概率加权、靶向最大似然估计等现代因果推断手段,在巢式抽样框架下更灵活地调整混杂因素,借助双稳健估计提升因果效应的准确性与可靠性,从而在观察性研究数据中更严谨地区分关联与因果关系,为阐明疾病机制、优化公共卫生策略提供更高等级的证据支撑<sup>[25-26]</sup>。

随着人群队列建设的不断完善、组学技术与

统计方法的持续革新,NCCS将在慢性病防控、罕见结局研究、生物标志物开发等领域发挥愈加重要的作用,为临床研究与公共卫生实践提供高效、可靠的研究范式。

综上,NCCS有效融合了队列研究与传统病例对照研究的优势,在保证较强因果推断能力的同时显著提升研究效率、节约研究成本,在临床与流行病学研究中具有重要应用价值。本文系统梳理了该设计的类型、实施要点、适用场景及其优缺点,并针对抽样策略、样本量估算、统计分析等关键环节进行阐述,以为临床研究者提供更规范、可操作的方法学参考。在实际研究中,建议研究者根据研究目的、暴露与结局特点、资源条件合理选择设计类型,规范病例与对照的选取流程,清晰报告设计方案与统计模型,以提高研究结果的真实性与可信度。深入理解并规范运用NCCS设计,可为疾病危险因素识别、生物标志物筛选及精准医学研究提供高质量的循证证据,对推动临床研究更规范、高效地开展具有持续而重要的意义。

伦理声明:不适用

作者贡献:论文撰写:陆春利、夏婷婷;论文修订与质量控制:邬娜、向颖;研究指导与论文审定:李亚斐

数据获取:不适用

利益冲突声明:无

致谢:不适用

## 参考文献

- 1 Wang R, Yang Y, Lu J, et al. Cohort profile: ChinaHEART (health evaluation and risk reduction through nationwide teamwork) cohort[J]. *Int J Epidemiol*, 2023, 52(5): e273-e282.
- 2 Gao M, Lv J, Yu C, et al. Metabolically healthy obesity, transition to unhealthy metabolic status, and vascular disease in Chinese adults: a cohort study[J]. *PLoS Med*, 2020, 17(10): e1003351.
- 3 Jiang Y, Lu L, Liu Z, et al. The Taizhou longitudinal study: a population-based biobank resource of genetic and biochemical biomarkers for precision medicine in China[J]. *J Genet Genomics*, 2026, 7: S1673-8527(26) 00048-2.
- 4 Fang C, Guan P, Yang H, et al. Metabolome-wide association study to decipher the risk of exogenous chemical exposure to liver cancer in a Chinese nested case-control cohort[J]. *Environ Int*, 2026, 211: 110274.
- 5 Zhang L, Bai Y, Cheng Y, et al. The potential carcinogenicity of long-term polymetallic exposure on the development of primary liver cancer: nested case-control evidence from the Jinchang cohort study[J]. *Environ Sci Technol*, 2025, 59(27): 13624-13637.
- 6 叶冬青.巢式病例对照研究的设计及分析[J]. *疾病控制杂志*, 2001, 5(1): 65-68. [Ye DQ. Design and analysis of nested case-control study[J]. *Chinese Journal of Disease Control & Prevention*, 2001, 5(1): 65-68.]
- 7 黄平,俞守义.巢式病例对照研究[J]. *继续医学教育*, 2006, 20(31):

- 75-77. [Huang P, Yu SY. Nested case-control study[J]. *Continuing Medical Education*, 2006, 20(31): 75-77.]
- 8 Hulley SB, Cummings SR, Browner WS, 等主编. 彭晓霞, 唐迅, 主译. 临床研究设计, 第4版[M]. 北京: 北京大学医学出版社, 2017: 1-358.
- 9 Porta M. A dictionary of epidemiology. 6th Edition[M]. Oxford: Oxford University Press, 2014: 1-180.
- 10 Zhang L, Li L, Andell P, et al. Attention-deficit/hyperactivity disorder medications and long-term risk of cardiovascular diseases[J]. *JAMA Psychiatry*, 2024, 81(2): 178-187.
- 11 Jia J, Ning Y, Chen M, et al. Biomarker changes during 20 years preceding Alzheimer's disease[J]. *N Engl J Med*, 2024, 390(8):712-722.
- 12 Qu Y, Liu X, Lin S, et al. Maternal serum folate during pregnancy and congenital heart disease in offspring[J]. *JAMA Netw Open*, 2024, 7(10): e2438747.
- 13 Partlett C, Hall NJ, Leaf A, et al. Application of the matched nested case-control design to the secondary analysis of trial data[J]. *BMC Med Res Methodol*, 2020, 20(1): 117.
- 14 詹思延, 叶冬青, 谭红专, 主编. 流行病学, 第8版[M]. 北京: 人民卫生出版社, 2017: 1-460.
- 15 Rassen JA, Shelat AA, Myers J, et al. One-to-many propensity score matching in cohort studies[J]. *Pharmacoepidemiol Drug Saf*, 2012, 21 Suppl 2: 69-80.
- 16 潘雄飞, 王意, 叶依, 等. 流行病学研究中的样本代表性问题(一)[J]. *中华疾病控制杂志*, 2019, 23(1): 1-4. [Pan XF, Wang Y, Ye Y, et al. Sample representativeness in epidemiological studies (Part I) [J]. *Chinese Journal of Disease Control & Prevention*, 2019, 23(1): 1-4.]
- 17 Richardson DB. An incidence density sampling program for nested case-control analyses[J]. *Occup Environ Med*, 2004, 61(12): e59.
- 18 祁子凡, 温馥源, 曹寒, 等. 大型人群队列研究随访监测设计研究进展[J]. *中华流行病学杂志*, 2022, 43(1): 134-138. [Qi ZF, Wen FY, Cao H, et al. Research progress on follow-up monitoring design of large population cohort studies[J]. *Chinese Journal of Epidemiology*, 2022, 43(1): 134-138.]
- 19 Ben Abdallah H, Bech R, Fogh K, et al. Comorbidities, mortality and survival in patients with pyoderma gangrenosum: a Danish nationwide registry-nested case-control study[J]. *Br J Dermatol*, 2021, 185(6): 1169-1675.
- 20 Gutierrez-Tordera L, Papandreou C, Novau-Ferré N, et al. Exploring small non-coding RNAs as blood-based biomarkers to predict Alzheimer's disease[J]. *Cell Biosci*, 2024, 14(1): 8.
- 21 Wang S, Hu C, Lin H, et al. Association of circulating long-chain free fatty acids and incident diabetes risk among normoglycemic Chinese adults: a prospective nested case-control study[J]. *Am J Clin Nutr*, 2024, 120(2): 336-346.
- 22 Hong M, Xiong T, Huang J, et al. Vitamin D supplementation and lower respiratory tract infection in infants: a nested case-control study[J]. *Infection*, 2023, 51(1): 109-118.
- 23 Ngo LH, Inouye SK, Jones RN, et al. Methodologic considerations in the design and analysis of nested case-control studies: association between cytokines and postoperative delirium[J]. *BMC Med Res Methodol*, 2017, 17(1): 88.
- 24 Kanduri C, Mamica M, Olstad E W, et al. Beware of counter-intuitive levels of false discoveries in datasets with strong intra-correlations[J]. *Genome Biol*, 2025, 26(1): 249.
- 25 Li M, Wang X. Semiparametric model averaging prediction in nested case-control studies[J]. *J Appl Stat*, 2024, 52(10): 1904-1930.
- 26 Shi K, Xu HJ, Lin ZL, et al. The correlation between postoperative hypouricemia and postoperative complications in patients with gastrointestinal cancer: a retrospective nested case-control study[J]. *BMC Cancer*, 2025, 25(1): 1466.

收稿日期: 2026年03月05日 修回日期: 2026年04月10日  
本文编辑: 杨燕 曹越

引用本文: 陆春利, 夏婷婷, 郇娜, 等. 巢式病例对照研究的方法学进展及其在临床研究中的应用[J]. *医学新知*, 2026, 36(5): 495-502. DOI: 10.12173/j.issn.1004-5511.202603034.

Lu CL, Xia TT, Wu N, et al. Methodological advances and application in clinical research of nested case-control studies[J]. *Yixue Xinzhi Zazhi*, 2026, 36(5): 495-502. DOI: 10.12173/j.issn.1004-5511.202603034.