

· 循证研究与临床转化 ·

网络Meta分析研究进展系列（二十）：网络Meta分析的样本量计算及精确性评估

杨智荣^{1,2}, 武珊珊³, 董圣杰⁴, 张天嵩⁵, 田金徽⁶, 孙凤⁷

【摘要】本文着重介绍网络Meta分析所需样本量的估算方法,包括有效合并研究数量法、有效样本量法和等效样本量法,并辅以实例演示计算步骤。在此基础上,我们进一步把样本量和效应估计值的精确性联系起来,介绍网络Meta分析的GRADE证据分级系统中对效应估计值精确性的评估。研究者在进行网络Meta分析后应评估当前样本量是否满足检验效能的要求,并据此对证据的精确性的升降级给出判断意见。

【关键词】网络Meta分析; 样本量计算; 精确性; GRADE证据分级

【中图分类号】R4

【文献标志码】A

开放科学(源服务)标识码(OSID)



Advances in methodology of network meta-analysis (20): sample size calculation and precision assessment in network meta-analysis Yang Zhirong*, Wu Shanshan, Dong Shengjie, Zhang Tiansong, Tian Jinhui, Sun Feng. *Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. Corresponding authors: Sun Feng, E-mail: sunfeng@bjmu.edu.cn

[Abstract] This review introduces the methods of calculating sample size required in network meta-analysis, including the approaches for the effective number of trials, the effective sample size and the equivalent sample size. Given the intrinsic link between sample size and estimate precision, we further introduce the precision assessment in GRADE system for network meta-analysis. It is recommended that whether sample size available in the network meta-analysis satisfies the requirements for statistical power should be usually assessed, which would be helpful for the grading assessment of evidence precision.

[Key words] Network meta-analysis; Sample size calculation; Evidence precision; GRADE

与其他类型的临床研究一样,Meta分析也存在样本量和检验效能的问题^[1]。当纳入的样本量不足时,Meta分析的检验效能较低,所估计的效应值可能会出现假阴性或假阳性的情况,结果的精确性较低(体现在置信区间的跨度较大)^[2]。

近年来,Meta分析的样本量和检验效能的问题逐渐得到关注。在传统Meta分析中,通常采用序贯分析来估算样本量^[3]。它按照纳入研究的发表时间顺序进行累计,通过调整随机误差和研究间异质性,最终估算得出确切结论时所需要的最小信息量(即纳入研究对象的总数量或结局事件发生数)。在网络Meta分析(network meta-analysis, NMA)中,也同样存在样本量与统计效能计算的问题^[4]。通过模拟研究发现,NMA合并结果往往因

统计效能不足而缺乏可信性,研究者及证据使用者应谨慎地评价NMA合并结果的统计效能,这对判断证据的真实性和临床价值尤为重要。

目前,NMA的样本量及统计效能的计算方法主要有三种:有效合并研究数量法、有效样本量法和等效样本量法^[4,5]。在此基础上,研究者可进一步评价证据的精确性^[5]。下面以我们课题组发表的一篇NMA为例^[6],介绍上述各种样本量计算方法的使用(证据网络结构如图1)。该例子是以二分类指标作为结局,下述所有方法也适用于连续性结局指标。

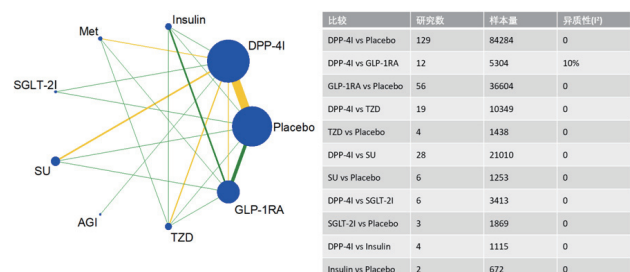


图1 证据网络结构图

注:为便于举例讲解,图中的研究数和样本量只显示与二肽基肽酶-4抑制剂(DPP-4I)及安慰剂相关的直接比较对,且把DPP-4I vs. 胰高血糖素样肽-1受体激动药(GLP-1RA)的 I^2 由0改为10%(实例原文中图1显示的直接比较的异质性 I^2 均为0)。Insulin:胰岛素制剂; Met:二甲双胍; Placebo:安慰剂; SGLT-2I:钠-葡萄糖共转运蛋白-2抑制剂; SU:磺脲类; TZD:噻唑烷二酮类; AGI:α-糖苷酶抑制剂

基金项目:国家自然科学基金(72074011);中国药品监管科学行动计划第二批重点项目

作者单位:¹ 518055 深圳,中国科学院深圳理工大学;² CB1 8RN 剑桥,英国剑桥大学临床医学院初级医疗中心;³ 100050 北京,首都医科大学附属北京友谊医院临床流行病学与循证医学研究室,国家消化系统疾病临床医学研究中心;⁴ 264000 烟台,烟台市烟台山医院骨关节科;⁵ 200040 上海,复旦大学附属静安区中心医院中医科;⁶ 730000 兰州,兰州大学循证医学中心甘肃省循证医学与临床转化重点实验室;⁷ 100191 北京,北京大学循证医学中心北京大学公共卫生学院流行病学与卫生统计学系

通讯作者:孙凤,E-mail:sunfeng@bjmu.edu.cn

doi: 10.3969/j.issn.1674-4055.2022.01.01

为了评估GLP-1RA和DPP-4I两类新型降糖药在2型糖尿病患者中的心血管事件,该NMA纳入了281项治疗时间 ≥ 12 周的随机对照试验(约180 000例患者),以心血管死亡、心肌梗塞、中风和心力衰竭等事件作为心血管复合结局(简称MACE)^[6]。研究发现,GLP-1RA与安慰剂和SU相比,结局风险降低,比值比(OR)分别是0.89(95%CI: 0.80~0.99)和0.76(95%CI: 0.59~0.99),而DPP-4I与安慰剂相比未显示出统计学差异(OR=0.92, 95%CI: 0.83~1.01)。研究的结论是,与SU或安慰剂相比,GLP-1RA可降低MACE发生风险,而DPP-4I则对MACE发生风险没有影响。但该NMA的结论是否有足够的样本量和检验效能来支撑?这就需要通过以下方法进行回答。

1 有效合并研究数量法

应用有效合并研究数量法进行计算时,前提是纳入合并的各个研究方差相等且具有同质性,即认为纳入研究的对象均来自同一个总体。其计算主要包括以下步骤^[4]:

1.1 根据间接比较的研究数目来确定精确度比率
先假设干预措施A和B比较的效应值均来自间接比较A和C的 n_{AC} 个研究与B和C的 n_{BC} 个研究,且每个研究的效应值方差均为 v ,那么根据间接比较的方差计算公式,A和B比较的效应值方差为

$$v_{\text{间接}} = \frac{v}{n_{AC}} + \frac{v}{n_{BC}} \quad (\text{公式 1})$$

又假设有 k 个($k=n_{AC}+n_{BC}$)A和B直接比较的研究,每个研究的效应值的方差均为 v ,在不存在异质性的情况下,那么由固定效应模型估算的A和B合并值的方差就为

$$v_{\text{直接}} = \frac{v}{n_{AC}+n_{BC}} \quad (\text{公式 2})$$

而所谓的精确度比率 R ,就是公式1和2之比,即

$$R = (n_{AC} + n_{BC})^2 / n_{AC}n_{BC} \quad (\text{公式 3})$$

该比值表示在间接比较中需要多少个AC和BC比较的研究才能达到一个AB直接比较的效应值的精确度。根据公式3,当 n_{AC} 和 n_{BC} 的比值为1:1时,可算得 R 为4;当比值为1:2时, R 为4.5;当比值为1:3时, R 为5.33。如此类推,可计算任意比值下的 R (表1)。

上述实例当中(图1),DPP-4I和安慰剂分别与GLP-1RA比较的研究分别有12和56个,按照上述公式3,可知 R 为6.88。即在满足纳入的各个研究的效应值方差相等且具有同质性的前提下,对于DPP-4I与安慰剂的比较来说,需约7个间接比较的研究才能达到1个直接比较的效应值的精确度。

1.2 计算由间接比较所贡献的有效合并研究数量
该指标是指在已有间接比较的研究数目下,可等同于多少个类似的直接比较研究。可根据间接比较的实际纳入研究和精确度比率计算,即

$$N_{\text{eff}} = (n_{AC} + n_{BC}) / R \quad (\text{公式 4})$$

例如上述实例当中(图1),对于DPP-4I与安慰剂的比较来说,由以GLP-1RA作为共同比较组的间接比较所贡献的有效合并研究数约为10个(按公式4计算,结果为9.88)。

为快捷地进行近似计算,Thorlund等制定了有效合并研究数量的表格(表1)^[4]。研究者通过查表,可迅速知道NMA中的间接比较贡献了多少有效合并研究数量。例如上述实例当中(图1),以GLP-1RA作为共同比较组的间接比较的研究数约为1:5(12和56的比值),合计68个研究,与表中的72(12:60)的数值最为接近,所对应的有效合并研究数为10个。与上述公式4的计算结果类似。大部分的间接比较都能在表1中找到近似的有效研究合并数,如果出现极端的间接比较研究数比值(如1:20),超出了表1的范围,此时需根据公式3和公式4来确定有效合并研究数。

有效合并研究数量法只适用于粗略估计间接比较中的有效研究数。如果需要更精确的估算,则应进一步计算有效样本量或等效样本量(即研究对象数目),具体见下述。

表1 有效合并研究数与所需的间接比较研究数

间接比较研究数比值	精确度比率	有效合并研究数					
		1	2	3	4	5	10
1:1	4	4 (2:2)	8 (4:4)	12 (6:6)	16 (8:8)	20 (10:10)	40 (20:20)
1:2	4.5	6 (2:4)	9 (3:6)	15 (5:10)	18 (6:12)	24 (8:16)	45 (15:30)
1:3	5.33	8 (2:6)	12 (3:9)	16 (4:12)	24 (6:18)	28 (7:21)	56 (14:42)
1:4	6.25	10 (2:8)	15 (3:12)	20 (4:16)	25 (5:20)	35 (7:28)	65 (13:52)
1:5	7.2	12 (2:10)	18 (3:15)	24 (4:20)	30 (5:25)	36 (6:30)	72 (12:60)
1:6	8.17	14 (2:12)	21 (3:18)	28 (4:24)	35 (5:30)	42 (6:36)	84 (12:72)
1:7	9.14	16 (2:14)	24 (3:21)	32 (4:28)	40 (5:35)	48 (6:42)	96 (12:84)
1:8	10.1	18 (2:16)	27 (3:24)	36 (4:32)	45 (5:40)	54 (6:48)	108 (12:96)
1:9	11.1	20 (2:18)	30 (3:27)	40 (4:36)	50 (5:45)	60 (6:54)	120 (12:108)
1:10	12.1	22 (2:20)	33 (3:30)	44 (4:40)	55 (5:50)	66 (6:60)	121 (11:110)

注:括号内表示 n_{AC} 和 n_{BC} 的数量

2 有效样本量法

该方法将NMA证据网络中的每一个比较对视为一个临床研究，通过估算每一个比较对的有效样本量来计算间接比较的统计效能和精确性^[4]。该法包括非校正和异质性校正两种方式，其计算主要包括：①根据样本量比值计算精确度比率；②分析各比较组是否存在异质性；③对具有同质性的比较组，用总体间接样本量乘以精确度比率即可获得有效间接样本量，对于存在异质性的比较组，则通过异质性校正因子对实际样本进行处理后，再计算有效间接样本量。以下介绍常用的通过有效样本量法来计算检验效能的步骤。

在图1的由DPP-4I、GLP-1RA和安慰剂三种干预构成的闭合环，DPP-4I与安慰剂、DPP-4I与GLP-1RA、GLP-1RA与安慰剂直接比较的样本量分别为84284、5304和36604，I²分别为0、10%和0。现在需要计算DPP-4I与安慰剂在此闭合环的NMA中的检验效能。

第一步：计算直接比较所需的样本量。在不考虑异质性的情况下，计算方法等同于一项RCT的样本量计算

$$n = C \times (Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})^2 \times \frac{V^2}{d^2} \quad (\text{公式 5})$$

d为预期效应值（如两组结局发生率的差值），V是预期效应值的方差（若结局为二分类时， $V^2 = \bar{P}(1-\bar{P})$ ， \bar{P} 为两组结局发生率的平均值）， $Z_{1-\frac{\alpha}{2}}$ 和 $Z_{1-\beta}$ 分别是正态分布在 $1-\frac{\alpha}{2}$ 和 $1-\beta$ 百分位上的Z值。一般假设两组样本量相等，此时C=4。若存在异质性，此时需要在n的基础上乘以异质性校正系数 $\frac{1}{1-H}$ ， $H=I^2$ 。

根据上述公式，若预计安慰剂组的MACE事件发生率为2%，DPP-4I组的发生率能降低20%（即发生率为1.6%）时认为有实际意义， $\alpha=0.05$ ， $1-\beta=90\%$ ，根据公式5，在不存在异质性的情况下，图1中DPP-4I与安慰剂直接比较的所需样本量为

$$n = 4 \times (1.96 + 1.28)^2 \times [(0.02 + 0.016)/2] \times [(1 - (0.02 + 0.016)/2)/(0.02 - 0.016)]^2 = 46389。$$

第二步：计算间接比较的有效样本量 n_{ind} ，也分为存在和不存在异质性两种情况。若不存在异质性时，可基于AC和BC的间接比较（如图1中的DPP-4I与GLP-1RA、安慰剂与GLP-1RA）计算AB（如图1中的DPP-4I与安慰剂）间接比较的有效样本量：

$$n_{ind} = \frac{n_{AC} \times n_{BC}}{(n_{AC} + n_{BC})} \quad (\text{公式 6})$$

若有异质性存在，则需要先对 n_{AC} 和 n_{BC} 进行调整，即分别乘以相应的 $1-I^2$ 。然后把调整后的 n_{AB}

和 n_{AC} 代入上述公式6。

例如，图1中 $n_{AC} = 5304 \times (1 - 0.10) = 4774$ （ $I^2=10\%$ ）， $n_{BC}=36604$ （ $I^2=0$ ），代入公式6，即得：

$$n_{ind} = (4774 \times 36604) / (4774 + 36604) = 4223。$$

同理，我们可以基于其他闭合环计算 n_{ind} 。根据图1中的数据 and 公式6，可得出DPP-4I-TZD-安慰剂、DPP-4I-SU-安慰剂、DPP-4I-SGLT-2I-安慰剂、DPP-4I-Insulin-安慰剂这四个闭合环为DPP-4I与安慰剂比较所贡献的 n_{ind} 分别为1263、1182、1208和419。

第三步：计算合计的有效样本量N。此时只需把DPP-4I与安慰剂直接比较的样本量（84284例，图1）和间接比较的有效样本量相加，即 $N=84284+4223+1263+1182+1208+419=92579$ ，多于经公式5计算得到的所需样本量46389，说明NMA实际纳入的样本量已满足DPP-4I与安慰剂比较的统计学要求。

第四步：计算检验效能 $1-\beta$ 。

$$1 - \beta = \Phi^{-1} \left(-Z_{1-\frac{\alpha}{2}} + \sqrt{\frac{Nd^2}{CV^2}} \right) \quad (\text{公式 7})$$

把上述所有参数代入公式7，得到 $1-\beta=98\%$ 。这说明基于当前样本量的NMA，有足够的检验效能发现DPP-4I与安慰剂比较的MACE事件发生率的差异。

3 等效样本量法

近年来，随着GRADE证据评级系统的发展，GRADE工作组对NMA结果的精确度评价作了相应的规定，提出等效样本量的计算方法^[5]。

该方法首先忽略整个证据网络中的直接证据和间接证据，将NMA的效应估计值和CI当作单个研究的结果，重点关注要达到该CI的精确度所需要的样本量，然后在特定基本假设下，反推计算产生该NMA效应估计值的等效样本量，跟所需样本量（通过公式5计算）作对比，从而确定等效样本量是否满足需求。以下分别介绍以RR、OR和均数差作为效应量时的等效样本量计算。

（1）假设纳入NMA的均为两组样本量相等的试验， $\log(RR)$ 的标准误可计算为：

$$SE_{\text{trial}} = SE(\log[RR]) = \sqrt{\frac{1}{n} \left(\frac{1}{p_c} + \frac{1}{RR \times p_c} - 2 \right)} \quad (\text{公式 8})$$

其中，n是每组样本量， p_c 是对照组中观察到的结局发生比例，RR是相对风险。NMA估计值 $\log(RR)$ 的标准误可以根据CI上、下界来计算：

$$SE_{NMA} = \frac{\log(CI_{upper}) - \log(CI_{lower})}{3.92} \quad (\text{公式 9})$$

假设两个标准误相等，即可算出样本量n：

$$n = \frac{\left(\frac{1}{p_c} + \frac{1}{RR \times p_c} - 2\right)}{(SE_{NMA})^2} \quad (\text{公式 10})$$

p_c 可通过网状结构中的相关比较组的比例（经Meta分析合并）来估计，也可通过简单计算结局发生数和患者数来近似估计 p_c 。

(2) 假设纳入NMA的均为两组样本量相等的试验， $\log(OR)$ 的标准误可以估计为：

$$SE_{\text{trial}} = SE(\log[OR]) = \sqrt{\frac{1}{n} \left(\frac{1}{p_c} + \frac{1}{1-p_c} + \frac{1}{p_t} + \frac{1}{1-p_t} \right)} \quad (\text{公式 11})$$

其中 n 为每组样本量， p_c 为对照组发生结局的比例， p_t 为观察组发生结局的比例。 p_t 也可以由 $p_t = p_c * OR / (1 - p_c + p_c * OR)$ 来计算，NMA估计值 $\log(OR)$ 的标准误可以根据CI上、下界来计算：

$$SE_{NMA} = \frac{\log(CI_{upper}) - \log(CI_{lower})}{3.92} \quad (\text{公式 12})$$

假设两个标准误相等，可以求解 n ：

$$n = \frac{\frac{1}{p_c} + \frac{1}{1-p_c} + \frac{1}{p_t} + \frac{1}{1-p_t}}{(SE_{NMA})^2} \quad (\text{公式 13})$$

(3) 假设纳入NMA的均为两组样本量相等的试验，两均数差的标准误可以估计为：

$$SE_{\text{trial}} = SD \times \sqrt{\frac{2}{n}} \quad (\text{公式 14})$$

其中 n 为每组的样本量，SD为组内合并标准差（可采用所有 k 个比较臂的SD的合并值），NMA均数估计值的标准误可以根据CI上、下界来计算：

$$SE_{NMA} = \frac{CI_{upper} - CI_{lower}}{3.92} \quad (\text{公式 15})$$

假设两个标准误相等，可以求解 n ：

$$n = \frac{2 \times SD^2}{(SE_{NMA})^2} \quad (\text{公式 16})$$

例如，图1的实例当中，DPP-4I与安慰剂相比MACE事件发生率的差异没有统计学差异（ $OR=0.92$ ，95%CI：0.83~1.01）。安慰剂组MACE事件的发生风险（各研究安慰剂组的中位数）为1.2%。根据公式11至13，可计算出满足OR、基线风险和CI的单个研究样本为每组70 506例，因此该NMA案例其估计值对应的等效样本量为141 012例，多于经公式5计算得到的所需样本量46 389，说明当前等效样本量满足统计要求。

4 精确性的评级

样本量的大小与效应估计值的精确性紧密相连，样本量越大，精确性越高。在GRADE证据分级中，效应估计值的精确性是其中重要的考虑因素之一^[7]。

按照NMA的GRADE分级规定，证据精确性的降级有三种情况^[5]。①当效应估计值的95%CI的上限和下限之比小于3（如实例中的效应估计值的95%CI为0.83~1.01，上下限比值为1.22），此时需要计算等效样本量和所需样本量（通过公式5计算），若等效样本量大于所需样本量，则证据的精确性不降级（例如图1实例中的DPP-4I与安慰剂MACE发生风险的比较），否则应降一级。②当效应估计值的95%CI的上限和下限之比大于3时，不需要计算等效样本量，因为这种情况下 SE_{NMA}^2 过大（远大于公式5中的 V ），使得由公式10、13、16计算得到的等效样本量过小，一般难以满足由公式5计算的所需样本量，因此证据的精确性应降两级。③当效应估计值的95%CI极其宽，不需要计算等效样本量，直接把证据的精确性降三级^[8]。

需要注意的是，GRADE证据分级对精确性的评价不一定要依赖等效样本量，也可以设定跟临床决策相关的效应阈值，然后根据95%CI是否包含该阈值来进行分级（具体请查看GRADE相应的文献^[5]）。

5 讨论

本文综述了NMA样本量计算的常用方法。目前这些方法都是采取简便的近似计算，帮助研究者评估NMA所估计的效应值是否满足检验效能的要求^[4,5]。

除了只是简便的近似计算以外，上述各种计算方法都有各自的前提假设和优缺点^[4,5]。有效合并研究数量法假设每个研究的样本量相等，效应值的方差相等，且不存在异质性。该方法只能非常粗略地估计在间接比较中需要纳入多少个研究才能达到一个直接比较研究的效应估计值的精确度。在实际操作中，更多使用有效样本量法和等效样本量法。这两种方法均考虑以研究个体作为单位计算样本量，而且均把NMA看作是单个研究，然后基于特定的假设来计算有效样本量或等效样本量；主要区别在于，前者是基于间接比较的样本量换算成与直接比较相当的样本量，后者则是基于NMA效应估计值的置信区间反推直接比较的样本量。此外，前者方法还可进一步对可能存在的异质性进行调整，而后者已把异质性的大小蕴含在随机效应模型所估算的置信区间之中。由于计算原理不同，这两种方法的结果可能会存在一定差异。研究者可根据实际情况选取一种或同时使用两种方法进行样本量计算。若这两种方法所得到的样本量均大于所需样本量，更有理由说明当前NMA中某两种特定干预的比较具有足够的检验效能。如本文的实例，对于DPP-4I和安慰

剂的比较,不管是有效样本量法还是等效样本量法,计算结果均大于所需样本量,因此在对证据的精确性进行评价时可不降级。

目前已经发表的NMA,包含样本量计算的研究并不多。但如果要回答NMA是否有足够的检验效能,则应对样本量进行计算,包括所需样本量和NMA的有效或等效样本量。在使用GRADE分级系统对证据的精确性进行分级时,效应估计值的精确性跟样本量密切相关,分级时可能也会涉及样本量的计算。建议研究者在进行NMA后应评估当前样本量是否满足检验效能的要求,并对效应估计值的精确度性进行评价。

参考文献

- [1] Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials[J]. *Lancet* 1998,351(9095):47-52.
- [2] Imberger G, Thorlund K, Gluud C, et al. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review[J]. *Bmj Open*, 2016,6(8).
- [3] Brok J, Thorlund K, Gluud C, et al. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses[J]. *Journal of Clinical Epidemiology*, 2008,61(8): 763-9.
- [4] Thorlund K, Mills E. Sample Size and Power Considerations in Network Meta-Analysis[J]. *Value Health*, 2012,15(7):A487.
- [5] Brignardello-Petersen R, Guyatt GH, Mustafa RA, et al. GRADE guidelines 33: Addressing imprecision in a network meta-analysis[J]. *Journal of Clinical Epidemiology*, 2021,139:49-56.
- [6] Wu SS, Cipriani A, Yang ZR, et al. The cardiovascular effect of incretin-based therapies among type 2 diabetes: a systematic review and network meta-analysis[J]. *Expert Opin Drug Saf*, 2018,17(3):243-9.
- [7] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision[J]. *J Clin Epidemiol*, 2011,64(12): 1283-93.
- [8] Piggott T, Morgan RL, Cuello-Garcia CA, et al. Grading of Recommendations Assessment, Development, and Evaluations (GRADE) notes: extremely serious, GRADE's terminology for rating down by three levels[J]. *Journal of Clinical Epidemiology*, 2020,120: 116-20.

本文编辑: 孙竹

· 循证视窗 ·

老年人接受两年的生活方式干预,可有效改善认知下降和预防心脑血管疾病

为了确定对生活方式的调整是否能够预防和延缓痴呆的发展,世界各地都在开展相关的临床试验,芬兰预防认知障碍和残疾的老年干预研究(FINGER)在十年前开展的研究首次证实,2年的生活方式综合干预能够显著改善老年人的认知下降。包括MIND-CHINA在内,很多国家的认知干预试验都是在FINGER研究的基础上设计的。

心血管疾病(CVD)和痴呆症一样在老年人中常见,甚至还拥有很多共同的风险因素,健康的生活方式更是与较低的心血管疾病风险有关。

《欧洲心脏杂志》发表了FINGER研究的新数据,分析结果显示,在老年人中进行2年的生活方式综合干预可显著降低CVD事件发生率13%~20%,尤其是对于有CVD病史的参与者来说,总CVD事件发生风险降低了50%!

FINGER研究总共有1259例参与者,年龄60~77岁,随机分配到干预组和对照组,对照组仅接受定期的健康建议,干预组的干预措施主要分为四个方面:饮食干预、运动训练、认知训练、CVD/代谢风险因素的管理。

饮食方面:基于膳食指南改善饮食结构,为每个参与者量身定制合适的食谱。

运动训练:循序渐进的肌肉力量训练+改善平衡力的锻

炼+规律的有氧运动。

认知训练:定期的小组学习和计算机练习程序,根据之前的其他分析报告,可综合训练参与者的执行能力、处理速度、记忆力等。

CVD/代谢风险因素管理是指对血压、血脂、血糖和体重的管理,必要时也可应用相关药物。

根据国家卫生登记处给出的数据,有319例参与者至少有过一次CVD事件,11%发生在试验开始前,说明他们有CVD病史。这部分参与者年龄较大,受教育程度较低,BMI较高,更有可能是男性,多会使用药物,所以在入组时血压和胆固醇水平更低,在干预组和对照组均匀分布。

自试验开始平均随访7.4年间,229例参与者发生了CVD事件,干预组107例,对照组122例,干预组CVD事件发生率低于对照组;进一步分析各种不同的CVD亚型,干预组卒中/短暂脑缺血发作(TIA)风险比为0.71,冠状动脉事件风险比0.84,总CVD事件风险比0.80。

值得关注的是,研究者们还根据参与者是否有疾病史进行了进一步分析,发现总CVD事件风险降低了50%,卒中/TIA的风险降低了60%。

【网络转载,版权归原作者所有】