

Mendeliaanse randomisatie

Hugoline G. de Haan, Bob Siegerink en Astrid van Hylckama Vlieg

Klassieke observationele studies naar de causale relatie tussen een risicofactor en een ziekte resulteren soms in tegenstrijdige en foutieve bevindingen. Dit komt door versturende factoren.

Uit de resultaten van klassieke observationele studies is niet goed op te maken of een bepaalde risicofactor een geschikt aangrijpingspunt kan zijn voor toekomstige behandelingen.

Een oplossing hiervoor is een analyse op basis van Mendeliaanse randomisatie, waarbij genetische variatie als surrogaatmarker voor de risicofactor wordt gebruikt.

Mendeliaanse randomisatie is gebaseerd op de gedachte dat eigenschappen en omgevingsfactoren evenredig verdeeld zijn over dragers en niet-dragers van verschillende genetische varianten.

Mendeliaanse randomisatie mag alleen toegepast worden als er een robuust verband is tussen de genetische variant en de risicofactor, als de genetische variant niet geassocieerd is met andere factoren die het verband tussen risicofactor en ziekte vertekenen, en als de genetische variant alleen via de risicofactor invloed uitoefent op de ziekte, dus niet via andere biologische mechanismen.

Als je een nieuwe behandelstrategie wil ontwikkelen, dan zou je idealiter van tevoren al willen vaststellen of de risicofactor waar je je strategie op gaat richten, daadwerkelijk een oorzaak is van het ziekteproces. Als die risicofactor geen oorzaak is, dan zullen eventuele behandelstrategieën immers weinig effect hebben op het ziekteproces.

Neem bijvoorbeeld het C-reactief proteïne (CRP). Enkele jaren geleden vond men in verschillende observationele onderzoeken een verband tussen de CRP-concentratie en het risico op cardiovasculaire ziekten: bij een hogere CRP-concentratie was het risico hoger.^{1,2} Maar een verband betekent nog geen causaal verband. Uitsluitel over het oorzakelijk verband tussen de CRP-concentratie en de verhoogde kans op cardiovasculaire ziekten was daarom wenselijk.

Om duidelijkheid te krijgen over een oorzaak-gevolgrelatie kan men een methode genaamd 'Mendeliaanse randomisatie' toepassen.^{3,4} Deze methode gebruikt genetische variatie als 'proxy' (surrogaatmarker) voor de risicofactor; in termen van ons voorbeeld: we gebruiken variatie in het CRP-gen als proxy voor CRP-concentraties in bloed. Hierdoor kan men aan de hand van observationeel onderzoek toch uitspraken doen over de causaliteit van het verband.

Dit artikel beschrijft wat Mendeliaanse randomisatie inhoudt, geeft een historisch voorbeeld en bespreekt wanneer deze methode wel of juist niet te gebruiken is.

Leids Universitair Medisch Centrum, afd. Klinische Epidemiologie, Leiden.

H.G. de Haan, MSc, onderzoeker in opleiding; dr. B. Siegerink en dr. A. van Hylckama Vlieg, onderzoekers en epidemiologen.

Contactpersoon: dr. B. Siegerink (b.siegerink@lumc.nl).

UITLEG

'Reverse causation'

Omdraaiing van oorzaak en gevolg: de ziekte beïnvloedt de risicofactor en niet andersom. Dit fenomeen kan optreden bij een klassieke observationele studie. Hierdoor kunnen onderzoeksresultaten vertekend zijn.

Confounding

Verstoring van het verband tussen risicofactor en ziekte door andere factoren die zowel gerelateerd zijn aan de risicofactor als aan de ziekte. Dit komt vaak voor bij klassieke observationele studies en kan vertekende onderzoeksresultaten opleveren.

Instrumentele-variabele-analyse

Instrumentele-variabele-analyse is een onderzoeksmethode in observationeel onderzoek waarbij een proxy ('instrument') voor een risicofactor wordt gebruikt om te bestuderen in hoeverre er een causale relatie tussen de risicofactor en ziekte is. Mendeliaanse randomisatie is hier een bijzondere vorm van.

Mendeliaanse randomisatie

Observationele onderzoeksmethode waarbij genetische variatie wordt gebruikt als proxy voor een risicofactor. Omdat reverse causation en confounding minder snel de resultaten van deze methode verstoren is het mogelijk om de causale relatie tussen risicofactor en ziekte te bestuderen.

Pleiotropie

Situatie waarin een genetische variant via meerdere en verschillende mechanismen een effect heeft op het lichaam. Als er sprake is van pleiotropie, is het niet mogelijk om een uitspraak te doen over de precieze causale mechanismen die leiden tot de ziekte.

Kanaliserende

Biologische aanpassingsmechanismen die in werking treden bij bepaalde genetische varianten en die de effecten van genetische variatie compenserend.

Voor alle duidelijkheid: we gaan niet in op de identificatie van genetische risicofactoren als voorspellers van ziekte.

PROBLEMEN IN KLASSIEKE OBSERVATIONELE STUDIES

Het aantonen van een causaal verband is niet altijd mogelijk in klassiek observationeel onderzoek. 'Confounding' en 'reverse causation' kunnen namelijk een statistisch verband tussen de risicofactor en de uitkomst verklaren (figuur 1).^{5,6}

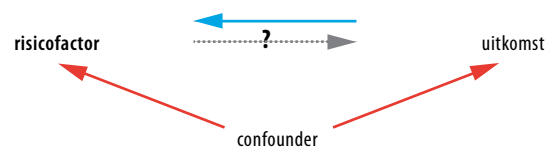
Bij reverse causation zijn oorzaak en gevolg omgedraaid. Een ziekte of een subklinische vorm daarvan veroorzaakt een verandering in de parameter die beschouwd wordt als risicofactor, in plaats van andersom. In het CRP-

voorbeeld dacht men dat verhoogde CRP-concentraties een oorzaak konden zijn van cardiovasculaire ziekten, bijvoorbeeld doordat het CRP de hechting van monocytten aan de vaatwand en hun migratie door de vaatwand zou bevorderen; op die manier zou het CRP indirect plaquevorming stimuleren.⁷ Aan de andere kant zouden CRP-concentraties ook verhoogd kunnen zijn door ontsteking van de vaatwand als gevolg van beginnende atherosclerotische processen. In dat geval is de ziekte de oorzaak en de verhoogde CRP-concentratie het gevolg. Bij confounding verstoren andere factoren de onderzochte associatie.⁵ Dat zijn dan factoren die geassocieerd zijn met zowel de risicofactor als het ziekteproces. Hierdoor lijkt er – ten onrechte – een verband tussen de risicofactor en het ziekteproces te zijn. Mogelijke bronnen van confounding in ons CRP-voorbeeld zijn roken en veroudering. Traditioneel kan men confounding tegengaan door aangepaste analyses uit te voeren.⁵ Als je echter niet alle bronnen van confounding volledig meeneemt in de analyses, blijft de associatie tussen risicofactor en uitkomst verstoord.

Waar klassieke observationele studies een vertekend resultaat kunnen opleveren door reverse causation en confounding, zijn deze problemen geminimaliseerd bij een Mendeliaanse randomisatie, zoals geïllustreerd met het historische voorbeeld in de tabel.

PRINCIPES VAN MENDELIAANSE RANDOMISATIE

Zoals gezegd ligt het principe van Mendeliaanse randomisatie in het gebruik van genetische variatie als proxy (of 'instrument') voor de variatie in de risicofactor waarvan men graag wil vaststellen of deze daadwerkelijk een oorzaak is van de ziekte. En daarmee is Mendeliaanse randomisatie een bijzondere vorm van een zogenoemde



FIGUUR 1 Relatie tussen risicofactor en uitkomst bij het gebruik van klassieke observationele analyses. In klassieke observationele analyses wordt de risicofactor (bijvoorbeeld 'CRP-concentratie') gerelateerd aan de uitkomst ('risico op cardiovasculaire ziekten'). Vaak is het niet mogelijk om vast te stellen of het gevonden verband ook een oorzakelijke relatie weergeeft. Dit komt doordat er verstorende factoren aanwezig zijn, zoals 'confounders' (bijvoorbeeld de leeftijd en de BMI, rode pijl) en 'reverse causation' (blauwe pijl), waarbij de uitkomst (bijvoorbeeld atherosclerose) de veronderstelde risicofactor beïnvloedt (hogere CRP-concentraties).

instrumentele-variabele-analyse (zie uitlegkader).¹⁰

In theorie kan men genetische variatie gebruiken als proxy voor allerlei risicofactoren, van eiwitten in het bloed tot leefstijl en psychologische factoren.¹¹ Terug naar ons voorbeeld: als er een causale relatie bestaat tussen CRP-concentraties en cardiovasculaire ziekten, dan is het aannemelijk dat variaties in het *CRP*-gen – die samenhangen met de CRP-concentraties in het bloed – ook een oorzaak zijn van cardiovasculaire ziekten. Anders geformuleerd: doordat dragers van een bepaalde variant in het *CRP*-gen levenslang licht verhoogde CRP-concentraties hebben, zouden deze dragers een verhoogd risico op cardiovasculaire ziekten moeten hebben als het CRP daadwerkelijk een oorzaak is.

Het minimaliseren van confounding en reverse causation bij analyses met Mendeliaanse randomisatie voert terug op de tweede wet van overerving van Gregor Mendel. Deze wet stelt dat genen, maar ook genetische variatie, onafhankelijk van elkaar van ouders op kind overerven. Analooq hieraan zijn niet alleen alle genetische eigenschappen, maar ook omgevingsfactoren evenredig verdeeld over de dragers van de verschillende genetische varianten van het *CRP*-gen. Zo zal het percentage rokers even groot zijn onder dragers als onder niet-dragers van een variant in het *CRP*-gen. Hierdoor wordt confounding geminimaliseerd. Maar ook reverse causation wordt geëlimineerd, aangezien de genetische variatie van een individu wordt vastgelegd bij de conceptie en dus niet wordt beïnvloed door ziekte.

VOORWAARDEN VOOR MENDELIAANSE RANDOMISATIE

Mendeliaanse-randomisatie-analyses zijn gebaseerd op een aantal aannames (figuur 2). Men kan nagaan of deze geldig zijn, maar de geldigheid is vaak niet te garanderen.^{10,12} De juiste interpretatie van een analyse met Mendeliaanse randomisatie hangt hier echter wel van af. Hier bespreken we kort deze voorwaarden en manieren om de validiteit van een Mendeliaanse-randomisatie-analyse te controleren.

ROBUUST VERBAND MET RISICOFACOR

Allereerst dient er een robuust verband te zijn tussen de genetische variant en de risicofactor. In ons voorbeeld houdt dit in dat variaties in het *CRP*-gen een voldoende groot effect moeten hebben op de CRP-concentraties. Dat is niet altijd het geval, aangezien vaak vele genetische varianten, in één of zelfs meerdere genen, bijdragen aan de variatie in de risicofactor. Zo worden de CRP-concentraties ook beïnvloed door variaties in andere genen, bijvoorbeeld in *HNF1a* en *LEPR*.¹³

Het gebruik van een genetische variant die een zeer zwakke associatie heeft met de risicofactor kan leiden tot bias.¹⁴ Dit kun je voorkomen door de verschillende genetische varianten in één risicoscore te combineren. Verder is het mogelijk dat de veronderstelde associatie tussen genotype en risicofactor ‘fout-positief’ is. Bij het gebruik van zo’n genetische variant in de Mendeliaanse-randomisatie-analyse zal, wellicht ten onterechte, geconcludeerd worden dat de risicofactor géén oorzaak is van de

TABEL Historisch voorbeeld van de toepassing van Mendeliaanse randomisatie

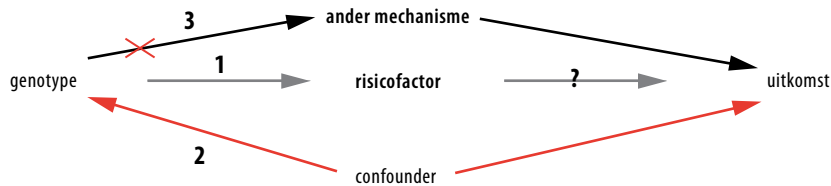
De relatie tussen serumcholesterolconcentratie en kanker

Mendeliaanse randomisatie werd voor het eerst beschreven in de jaren 80, al werd die methode toen nog niet zo genoemd. In een brief aan *The Lancet* beschreef Katan hoe met behulp van variatie in het gen voor apolipoproteïne E (apoE) de causale relatie tussen cholesterolconcentraties en kanker onderzocht kon worden.⁸ Resultaten van diverse klassieke observatieve studies hadden geleid tot de gedachte dat lagere serumcholesterolconcentraties mogelijk kanker konden veroorzaken. Hiermee was cholesterolverlagende medicatie in een kwaad daglicht komen te staan.

Katan motiveerde hoe de resultaten uit observatieve studies mogelijk verklaard kunnen worden door ‘reverse causation’: de lage serumcholesterolconcentraties bij kankerpatiënten zijn mogelijk een gevolg van de aanwezigheid van de tumor, zelfs als deze nog in een subklinisch stadium verkeert. Maar ook confounding speelt een rol: factoren als leeftijd en leefstijl zijn gerelateerd aan zowel de hoogte van de cholesterolconcentratie als aan het risico op kanker.

Om tot een definitieve uitspraak te komen suggereerde Katan om gebruik te maken van natuurlijke variaties in het *apoE*-gen, waarvan bekend was dat dragers hun leven lang gemiddeld lagere serumcholesterolconcentraties hebben dan niet-dragers, onafhankelijk van hun leefstijl of andere factoren die cholesterolconcentraties beïnvloeden. Katan redeneerde als volgt: als lage cholesterolconcentraties oorzakelijk zijn voor tumorgroei, dan zullen mensen met deze *apoE*-genvarianten ook een verhoogd risico op kanker hebben.⁸

De voorgestelde Mendeliaanse-randomisatie-analyse werd jaren later uitgevoerd in een studie waarin DNA-materiaal, cholesterolbepalingen en gegevens over kankerincidentie van bijna 3000 individuen beschikbaar waren. Uit deze analyse bleek dat lagere cholesterol- en LDL-concentraties weliswaar geassocieerd waren met een hoger risico op kanker, maar dat variatie in het *apoE2*-gen niet geassocieerd was met kankerincidentie en kankermortaliteit.⁹ Hieruit kan geconcludeerd worden dat lage LDL- en cholesterolconcentraties geen oorzaak zijn van kanker. De eerder gevonden associaties in de klassieke epidemiologische analyses waren dus het gevolg van confounding of reverse causation.



FIGUUR 2 Voorwaarden waaronder Mendeliaanse randomisatie toegepast mag worden. Allereerst dient er een sterke associatie te zijn tussen het genotype (bijvoorbeeld een bepaalde variant van het *CRP*-gen) en de risicofactor (in dit voorbeeld: de *CRP*-concentratie)(pijl 1). De tweede voorwaarde houdt in dat het genotype niet geassocieerd mag zijn met een andere risicofactor (bijvoorbeeld etniciteit) die gerelateerd is aan de uitkomst (in dit geval: cardiovasculaire ziekte); anders zou confounding ontstaan (pijl 2). Tot slot mag het genotype niet geassocieerd zijn met de uitkomst via een ander mechanisme (pleiotropie, pijl 3).

ziekte. Dit kun je voorkomen door alleen genetische varianten te gebruiken waarvan in meerdere, onafhankelijke studies is vastgesteld dat zij geassocieerd zijn met de risicofactor.

GEEN ASSOCIATIE MET ANDERE CONFOUNDERS

De tweede voorwaarde is dat de genetische variant niet geassocieerd mag zijn met andere factoren die het verband tussen de risicofactor en de ziekte vertekenen. Dit zou namelijk weer leiden tot confounding. Variaties in het *CRP*-gen mogen dus niet vaker gepaard gaan met andere factoren die het risico op cardiovasculaire ziekten beïnvloeden dan men op basis van toeval zou verwachten. De frequentste voorbeelden hiervan zijn genetische varianten die tegelijk overerven met de genetische variant die in de Mendeliaanse randomisatie onderzocht wordt; er is dan sprake van 'linkage disequilibrium'. Dit gebeurt wanneer de varianten in dezelfde regio op hetzelfde chromosoom liggen. Dat is met behulp van publieke databases en biologische kennis redelijk goed te bestuderen.

Een andere bron van confounding kan optreden wanneer men een Mendeliaanse-randomisatie-analyse uitvoert bij een onderzoekpopulatie met mensen van verschillende etnische achtergronden, elk met hun eigen genetische variaties en eigen basisrisico op ziekte. Deze vorm van confounding – ook wel 'populatiestratificatie' genoemd – kan men voorkomen door aanpassingen in de onderzoeksopzet of door correcties in de data-analyse. Dit is natuurlijk alleen mogelijk als de verschillende etnische achtergronden binnen de onderzoekpopulatie voldoende nauwkeurig vastgesteld kunnen worden.

GEEN ANDERE BIOLOGISCHE MECHANISMEN

Als derde en laatste voorwaarde dient de genetische variant alleen via de tussenliggende risicofactor geassocieerd te zijn met de ziekte, en dus niet via andere biologische mechanismen. Dit betekent dat variaties in het *CRP*-gen niet mogen leiden tot veranderingen in serumcholeste-

rolconcentraties of andere factoren die geassocieerd zijn met cardiovasculaire ziekten. Als dat wél het geval is, is er sprake van 'pleiotropie' en is het niet langer duidelijk voor welke risicofactor de genetische variatie nu als proxy functioneert. De aanwezigheid van pleiotropie is soms bekend uit de literatuur en kan, soms worden vastgesteld aan de hand van eigen data, maar men kan dit fenomeen nooit uitsluiten.

Een ander fenomeen dat het verband tussen het genotype en de uitkomstmaat kan vertekenen is 'kanalisatie'. Dit houdt in dat er biologische, epigenetische aanpassingsmechanismen in werking komen, bijvoorbeeld verhoging van de concentratie *IL-10* – een cytokine met anti-atherosclerotisch effect –, die de effecten van de genetische variatie in *CRP* compenseren. Het is moeilijk in te schatten in hoeverre kanalisatie plaatsvindt en in hoeverre dit ook daadwerkelijk het verband tussen genotype en uitkomstmaat vertekent.

VOOR- EN NADELEN VAN MENDELIAANSE RANDOMISATIE

Het grootste voordeel van Mendeliaanse randomisatie is het minimaliseren van bekende en onbekende confounding en reverse causation. Daarnaast kan men het effect van een levenslange blootstelling bepalen, iets wat in klassiek observationeel onderzoek vaak niet mogelijk is. Om meer duidelijkheid te krijgen over de causale relatie tussen risicofactor en ziekteproces kan men ook kiezen voor een gerandomiseerde, gecontroleerde trial (RCT), waarbij de onderzoeker de blootstelling aan de risicofactor direct beïnvloedt. Door de randomisatie worden alle mogelijke versturende factoren in principe gelijk verdeeld over de behandelgroepen en daardoor is confounding geminimaliseerd. Deze onderzoeksopzet is echter niet altijd mogelijk vanwege ethische en praktische overwegingen. Bovendien is de generaliseerbaarheid van een RCT doorgaans beperkt door strenge in- en exclusiecriteria en relatief gezonde deelnemers. In observationele studies daarentegen is het mogelijk om een representa-

tievare steekproef van de algemene bevolking te includeren om daarmee de klinische toepasbaarheid van de resultaten te vergroten.

De methode van Mendeliaanse randomisatie kent ook beperkingen.^{3,4,12} Zo moet de onderzoekspopulatie vaak erg groot zijn om met enige zekerheid de associaties – zowel tussen proxy en risicofactor als tussen proxy en ziekte – te kunnen bepalen. En zoals we al hebben aangegeven, steunt Mendeliaanse randomisatie op een aantal voorwaarden. Als aan deze voorwaarden niet voldaan wordt, kan de methode vertekende resultaten opleveren. Het is aan de onderzoeker om de lezer ervan te overtuigen dat aan alle voorwaarden zo goed als mogelijk is voldaan.

CONCLUSIE

Klassieke observationele studies naar de causale relatie tussen een risicofactor en een ziekte resulteren soms in tegenstrijdige en foutieve bevindingen door de aanwezigheid van versturende factoren (bias en confounding) of reverse causation. Hierdoor blijft het onduidelijk of de risicofactor een geschikt aangrijpingspunt kan zijn voor toekomstige behandelingen. In die situatie worden steeds vaker analyses op basis van Mendeliaanse randomisatie toegepast.

Bij Mendeliaanse-randomisatie-analyses worden de associaties tussen genotype en risicofactor en tussen genotype en ziekte gebruikt om het causale verband tussen de risicofactor en de ziekte te herleiden. Doordat deze methode uitgaat van onafhankelijke overerving van genen, zijn problemen als confounding en reverse causation geminimaliseerd. Mendeliaanse randomisatie vereist echter grote studiepopulaties, steunt op enkele cruciale aannames en kan vertekende resultaten opleveren als niet aan de voorwaarden is voldaan. Het is niet mogelijk te garanderen dat aan alle voorwaarden is voldaan, maar als er voldoende biologische kennis is over de risicofactor en het ziektemechanisme kan men op basis van

- **Het aantonen van een causaal verband tussen een risicofactor en ziekte is niet altijd mogelijk in een klassieke observationele studie.**
- **Dit komt door de aanwezigheid van versturende factoren (confounders) of doordat de ziekte de risicofactor beïnvloedt ('reverse causation').**
- **Een oplossing hiervoor is het uitvoeren van een analyse op basis van Mendeliaanse randomisatie.**
- **Bij een Mendeliaanse-randomisatie-analyse wordt genetische variatie gebruikt als surrogaatmarker voor de risicofactor.**
- **Mendeliaanse randomisatie is gebonden aan bepaalde voorwaarden, waaronder een robuust verband tussen de genetische variant en de risicofactor.**
- **Mendeliaanse-randomisatie-analyse vereist grote studiepopulaties en voldoende biologische en klinische kennis.**

analyses van eigen onderzoeksgegevens goed beoordelen in hoeverre aan de voorwaarden voor Mendeliaanse randomisatie is voldaan.

In de rubriek Stand van zaken verschijnen regelmatig bijdragen over methoden die gebruikt worden bij het opzetten van wetenschappelijk onderzoek. De artikelen in deze serie illustreren op begrijpelijke wijze wat een bepaalde methode behelst, zonder dat hier uitvoerige methodologische kennis voor nodig is. Zowel oude als nieuwe methodologische principes worden zo inzichtelijk gemaakt voor artsen die klinische onderzoeken goed willen interpreteren.

Belangenconflict en financiële ondersteuning: geen gemeld.

Aanvaard op 24 juli 2014

Citeer als: Ned Tijdschr Geneesk. 2014;158:A7547

> KIJK OOK OP WWW.NTVG.NL/A7547

LITERATUUR

- 1 Ridker PM, Buring JE, Shih J, et al. Prospective Study of C-Reactive Protein and the Risk of Future Cardiovascular Events Among Apparently Healthy Women. *Circulation*. 1998;98:731-3.
- 2 Koenig W, Sund M, Frohlich M, et al. C-reactive protein, a sensitive marker of inflammation, predicts future risk of coronary heart disease in initially healthy middle-aged men. *Circulation*. 1999;99:237-42.
- 3 Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1-22.
- 4 Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials and limitations. *Int J Epidemiol*. 2004;33:30-42.
- 5 Groenwold RH. Verstoring in observationeel onderzoek: 'confounding'. *Ned Tijdschr Geneesk*. 2012;156:A4221.
- 6 Groenwold RHH. 3 vormen van bias. *Ned Tijdschr Geneesk*. 2013;157:A6497.
- 7 Jialal I, Devaraj S, Venugopal SK. C-Reactive Protein: Risk marker or mediator in atherothrombosis? Hypertension. 2004;44:6-11.
- 8 Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet*. 1986;1(8497):507-8.
- 9 Trompet S, Jukema JW, Katan MB, et al. Apolipoprotein e genotype, plasma cholesterol, and cancer: a Mendelian randomization study. *Am J Epidemiol*. 2009;170:1415-21.

- 10 Boef AGC, le Cessie S, Dekkers OM. Instrumentele-variabele-analyse. *Ned Tijdschr Geneeskd.* 2013;157:A5481.
- 11 Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ.* 2005;330:1076-9.
- 12 Lawlor DA, Harbord RM, Sterne JA, et al. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med.* 2008;27:1133-63.
- 13 Wu Y, McDade TW, Kuzawa CW, et al. Genome-wide association with C-reactive protein levels in CLHNS: evidence for the CRP and HNF1A loci and their interaction with exposure to a pathogenic environment. *Inflammation.* 2012;35:574-83.
- 14 Burgess S, Thompson SG. CRP HD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization. *Int J Epidemiol.* 2011;40:755-64.