



Universiteit  
Leiden  
The Netherlands

## Optimal cardiovascular treatment strategies in kidney disease: casual inference from observational data

Fu, E.L.

### Citation

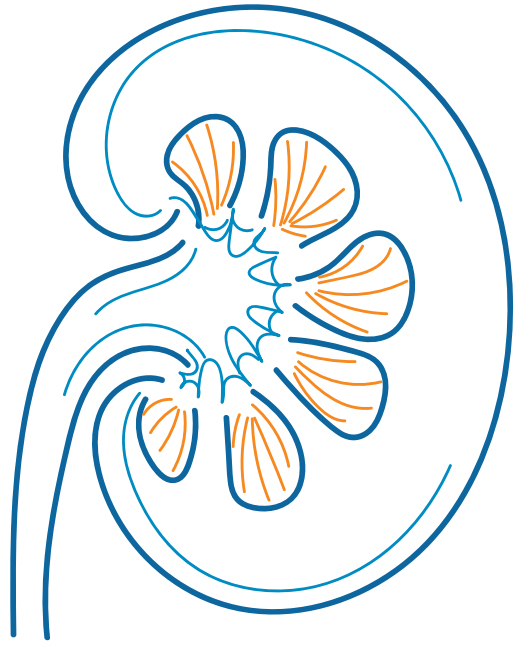
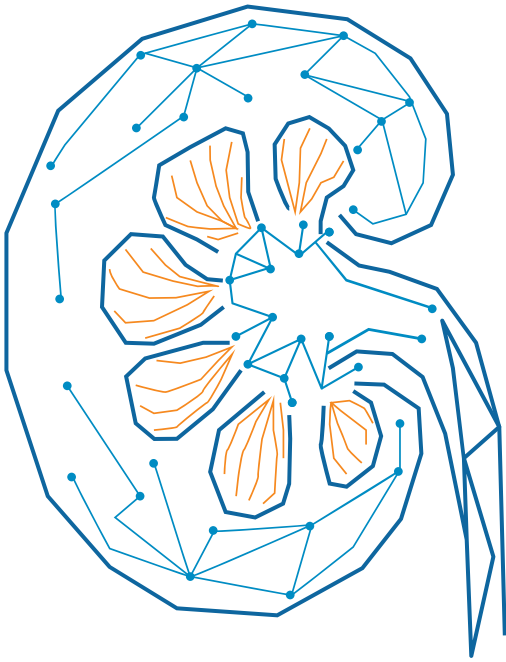
Fu, E. L. (2021, October 28). *Optimal cardiovascular treatment strategies in kidney disease: casual inference from observational data*. Retrieved from <https://hdl.handle.net/1887/3221348>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3221348>

**Note:** To cite this publication please use the final published version (if applicable).



# APPENDICES

## Nederlandse samenvatting (Dutch summary)

### Samenvatting van belangrijkste bevindingen

Het doel van dit proefschrift was om een substantiële bijdrage te leveren aan de besluitvorming bij patiënten met nierziekten. In dit proefschrift zijn een aantal klinische vragen beantwoord over de effectiviteit en veiligheid van therapieën door gebruik te maken van state-of-the-art methoden en routinematig verzamelde zorgdata. Daarnaast is getracht het bewustzijn te vergroten voor potentiële systematische vertekeningen die kunnen optreden bij het gebruik van observationele data, en hoe deze biases verminderd kunnen worden.

Observationele studies met een causaal doel kunnen geplaagd worden door verschillende biases, waarvan enkele worden beschreven in **hoofdstuk 2**. Alhoewel confounding door indicatie een gevaar vormt bij alle observationele studies die als doel hebben de causale effecten van behandelingen te schatten, kunnen vragen over onbedoelde behandel-effecten en vragen met een actieve controlegroep minder gevoelig zijn voor confounding. De resterende confounding moet geadresseerd worden in de statistische analyse, waarbij confounders geïdentificeerd moeten worden op basis van vakinhoudelijke kennis, bijvoorbeeld door gebruik te maken van causale diagrammen. Er kan alleen gecorrigeerd worden voor gemeten confounding en dit kan gedaan worden met multivariabele regressie, standaardisatie of propensity score methoden. Voor eenmalige behandelingen kunnen alle methoden even goed corrigeren voor gemeten confounders. In de setting van langdurige behandelingen en behandeling-confounder feedback zijn speciale methoden gebaseerd op standaardisatie of wegen nodig. De impact van residuele confounding kan worden onderzocht door een E-waarde te berekenen, door middel van een kwantitatieve bias analyse, of door het uitvoeren van een negatieve controle uitkomst of positieve cohort analyse. Naast confounding zijn prevalent user bias en immortal time bias belangrijke limitaties in veel observationele studies. Deze biases ontstaan wanneer de start van follow-up en de start van de blootstelling niet overeenkomen. Het expliciet emuleren van een gerandomiseerde trial kan deze biases elimineren aangezien dit forceert dat de start van follow-up en blootstelling overeenkomen. Ten slotte komen missende waarden en meetfouten veelvuldig voor in routinematig verzamelde zorgdata. Het wordt daarom aanbevolen om methoden zoals multi-pele imputatie of kwantitatieve bias analyse toe te passen.

In **hoofdstuk 3** bediscussiëren we het concept en de voor- en nadelen van propensity score (PS) methoden, een populaire statistische methode om te corrigeren voor gemeten confounding. Er bestaan verschillende soorten PS methoden. Dit zijn PS matching, PS stratificatie, multivariabele regressie waarbij de

PS als variabele wordt toegevoegd, en PS wegen. PS methoden hebben een aantal voordelen ten opzichte van conventionele multivariabele regressie: het is mogelijk om de balans van variabelen te checken, om een specifieke target populatie te kiezen, en om individuen te excluderen die zich in de niet-overlappende delen van de PS distributie bevinden. Daarnaast hebben PS methoden de voorkeur in de setting van hoog-dimensionele data waarbij er veel confounders zijn en relatief weinig uitkomsten.

In **hoofdstuk 4** besteden we aandacht aan immortal time bias in een gepubliceerde observationele studie die als doel heeft om het causale effect van metformine te schatten in patiënten met diabetische nefropathie. We stellen een aantal oplossingen voor die toegepast hadden kunnen worden om deze bias te elimineren. Dit zijn onder andere het modelleren van metformine als tijdsafhankelijke blootstelling, landmarken of het gebruik van grace periodes in combinatie met de cloning-censoring-weighting methode. Een andere optie is het emuleren van een trial met de sequentiële trial methode.

In **hoofdstuk 5** bestuderen we de associatie tussen acute stijgingen in creatinine na het starten van RAS-blokkers en cardiovasculaire en nieruitkomsten. Patiënten met een hogere creatininesijging na de start van RAS-blokkers hadden een hoger risico op sterfte, hartfalen, myocardinfarcten en nierfalen. We vonden ook dat slechts 18% van de individuen die RAS blokkade startten de aanbevolen creatinine monitoring kregen, en dat creatininesijgingen tussen de 10-29% regelmatig voorkwamen bij gemonitorde individuen. Deze studie adresseert niet de kwestie van het stoppen van RAS-blokkers bij creatininesijgingen, maar suggereert wel dat patiënten met een creatininesijging een hoger risico hebben op cardiovasculaire en nieruitkomsten.

**Hoofdstuk 6** is een vergelijkend effectiviteitsonderzoek tussen RAS-blokkers en calciumantagonisten op het starten van nierfunctie-vervangende therapie (dialyse of niertransplantatie), sterfte en cardiovasculaire uitkomsten (een combinatie van cardiovasculaire sterfte, myocardinfarct of beroerte) in patiënten met gevorderde chronische nierziekten. We vonden dat het starten van RAS blokkade was geassocieerd met een lager risico op nierfunctie-vervangende therapie en vergelijkbare risico's op sterfte en cardiovasculaire uitkomsten vergeleken met calciumantagonisten. Daarnaast is een analyse uitgevoerd in een positief controle cohort van patiënten met milde tot matige chronische nierziekten, waarvan de resultaten overeenkwamen met bewijs uit eerdere gerandomiseerde onderzoeken. Een negatieve controle analyse met kanker als uitkomst suggereerde geen residuele confounding door body mass index of roken. Deze bevindingen zouden de klinische besluitvorming rond de keuze van antihypertensiva kunnen informeren in patiënten met gevorderde chronische nierziekten.

In **hoofdstuk 7** vonden we dat het gebruik van  $\beta$ -blokkers bij ziekenhuisontslag geassocieerd was met een lager risico op sterfte, en een gecombineerd eindpunt van cardiovasculaire sterfte en ziekenhuisopnames voor hartfalen in individuen met hartfalen met een verminderde ejectiefractie en gevorderde chronische nierziekten. Een positieve controle analyse in individuen met hartfalen met een verminderde ejectiefractie en milde tot matige chronische nierziekten lieten vergelijkbare reducties in uitkomsten zien voor gebruikers van  $\beta$ -blokkers. Deze voordelen werden niet gezien in individuen met gevorderde chronische nierziekten en matige of behouden ejectiefractie, alhoewel deze analyses niet eenduidig waren door lage power. Deze bevindingen suggereren dat  $\beta$ -blokkers effectief zijn in patiënten met hartfalen met een verminderde ejectiefractie over het gehele spectrum van nierfunctie.

In **hoofdstuk 8** is het effect van stoppen versus doorgaan van RAS-blokkers in individuen met gevorderde chronische nierziekte onderzocht op de uitkomsten sterfte, cardiovasculaire uitkomsten, en nierfunctie-vervangende therapie. We observeerden dat individuen die stopten met RAS blokkade een hoger risico hadden op sterfte en cardiovasculaire uitkomsten, maar een lager risico hadden op nierfunctie-vervangende therapie. Vergelijkbare bevindingen werden verkregen wanneer gebruik van RAS-blokkers werd gemodelleerd als tijdsafhankelijke variabele in een marginal structural model, wanneer daarnaast gecorrigeerd werd voor kalium en albumine-creatinine ratio, en in subgroepen, inclusief het stoppen van RAS-blokkers bij een hogere (nierfunctie tussen 20-30%) of lagere eGFR (nierfunctie <20%). Deze bevindingen suggereren dat, terwijl de resultaten van de gerandomiseerde STOP-ACEi trial worden afgewacht, niet routinematig moet worden gestopt met RAS-blokkers in patiënten met gevorderde chronische nierziekten.

**Hoofdstuk 9** adresseert de vraag of er een optimale nierfunctie is om te starten met dialyse. In dit onderzoek konden we met observationele data de bevindingen van de gerandomiseerde IDEAL trial repliceren. Daarnaast toonden we aan dat vroeg starten met dialyse geassocieerd was met een bescheiden vermindering in sterfte en cardiovasculaire uitkomsten (met een nierfunctie tussen de 15-16% versus een nierfunctie tussen 6-7%). Dit vertaalde zich naar een hogere levensverwachting van 1,6 maanden, terwijl dialyse gemiddeld 48 maanden eerder gestart zou moeten worden. We tonen ook aan dat eerdere observationele studies leden aan lead time bias, selectiebias en immortal time bias, dat deze systematische vertekeningen voorkomen hadden kunnen worden als het target trial emulatie framework was toegepast, en dat incorrecte analyse van onze eigen data tot dezelfde verkeerde resultaten leidt. Als geheel wijzen deze resultaten erop dat er weinig voordeel is van het vroeg starten van dialyse op basis van nierfunctie alleen. In toekomstige studies kan worden onderzocht of dialyse gestart moet worden op basis van symptomen om klinische uitkomsten verder te verbeteren.

## Toekomstperspectieven

Het aantal observationele studies dat gebruik maakt van routinematig verzamelde zorgdata neemt almaar toe. In dit proefschrift is toegelicht dat het gebruik van dit soort data om de klinische praktijk te informeren een tweesnijdend zwaard betreft: aan de ene kant biedt het enorme kansen om te bestuderen hoe behandelingen in de routine klinische praktijk werken, om vragen te bestuderen die moeilijk te beantwoorden zijn in gerandomiseerde trials, en om populaties te bestuderen die zijn ondervertegenwoordigd in trials. Aan de andere kant kunnen verscheidene biases de bevindingen uit observationele studies vertekenen: confounding bias, immortal time bias en prevalent user bias om er slechts een paar te noemen.

Hoe moeten we verder om het beste bewijs te leveren voor de behandeling van patiënten met nierziekten? Natuurlijk zullen er meer gerandomiseerde trials moeten komen. In nierziekten zijn relatief weinig gerandomiseerde trials gedaan, en patiënten met nierziekten zijn overwegend geëxcludeerd uit trials in andere vakgebieden zoals de cardiologie en oncologie. Om dit probleem op te lossen hebben andere onderzoekers opgeroepen tot vermindering van de kosten en complexiteit van trials, inclusief de bureaucratische last. Dat dit mogelijk is, is recent aangetoond door de RECOVERY trial, die binnen een jaar meer dan 39.000 gehospitaliseerde COVID-19 patiënten heeft gerandomiseerd. Andere voorbeelden zijn onder andere de publicatie van zogeheten "grote, simpele trials" in de afgelopen decennia en recente innovaties in trial design zoals de registratie-gebaseerde trial. Gelukkig lijkt het vakgebied van nierziekten een inhaalslag te maken, met de publicaties van een aantal belangrijke klinische trials, terwijl andere trials nu actief bezig zijn met het includeren van patiënten. Naast het uitvoeren van meer trials kunnen nieuwe benaderingen om trialresultaten te generaliseren naar andere populaties de kloof tussen trials en de dagelijkse klinische praktijk overbruggen.

Een fundamentele vraag is of het überhaupt mogelijk is causale conclusies te trekken uit observationele data. Sommige onderzoekers zijn inderdaad van mening dat causale conclusies alleen uit gerandomiseerde trials kunnen worden getrokken, maar niet uit observationele studies, en dat alleen gerandomiseerde onderzoeken nuttig zijn. Dit is echter een valse tweedeling. Causaliteit is niet een ja/nee statement en wordt zelden geconcludeerd op basis van één studie, aangezien het altijd om de totaliteit van bewijs gaat, wat kan komen uit laboratorium onderzoek, observationele studies en gerandomiseerde trials. Daarnaast zijn er substantiële verschillen in kwaliteit tussen observationele studies, waarbij sommigen beter tot causale conclusies kunnen komen dan anderen. Aan de ene kant is het mogelijk met goed uitgevoerde observationele studies de resultaten van gerandomiseerde trials te repliceren of zelfs te voorspellen. Aan de andere kant bestaan er legio voorbeelden van observationele studies waarbij dit niet is gelukt. Dit laatste kan

vaak verklaard worden door het feit dat deze observationele studies gebrekkige methoden hebben gebruikt die onnodige systematische vertekeningen introduceerden, zoals immortal time bias en prevalent user bias, en werd in deze gevallen niet veroorzaakt door de aanwezigheid van ongemeten confounding. Deze biases ontstaan als de timing van de volgende drie elementen niet overeenkomt op baseline: de start van follow-up, de start van de behandelstrategieën en het voldoen aan alle inclusiecriteria. Omdat door randomisatie de timing van deze elementen automatisch overeenkomt, hoeven onderzoekers van gerandomiseerde trials zich hier nooit zorgen over te maken. Dit is echter niet het geval voor observationele data, waar de onderzoeker zorgvuldig moet nadenken over de baseline en hier adequaat mee om moet gaan in de analyses. Voorbeelden uit de literatuur waarbij goed uitgevoerde observationele studies in staat waren dezelfde antwoorden te krijgen als trials, terwijl observationele studies die vermijdbare biases introduceerden dit niet konden bereiken, zijn onder andere onderzoeken naar de timing van dialyse en het risico op sterfte (dit proefschrift), postmenopauzale hormoontherapie en het risico op coronaire hartziekten, statines en het risico op kanker, de timing van antiretrovirale behandeling en het risico op sterfte, dabigatran en het risico op beroertes, sodium-glucose cotransporter 2 remmers en het risico op sterfte, en colonoscopie screening en het risico op darmkanker. Het gebruik van het target trial emulatie framework kan helpen om deze onnodige studiedesign gebreken te elimineren, dwingt de onderzoeker om betekenisvolle causale vragen te stellen, faciliteert de communicatie en is leidend voor de statistische analyse. Daarnaast moeten onderzoekers de analytische methoden gebruiken die het best passen bij de klinische onderzoeksvraag. De cloning-censoring-weighting methode is bijvoorbeeld geschikt om vragen te beantwoorden die 1) verschillende initiatiemomenten van een interventie vergelijken ("Wanneer moeten we starten met de behandeling?"); 2) verschillende behandelduur vergelijken ("Hoe lang moeten we behandelen?") en 3) een grace periode bevatten ("Moet de behandeling binnen x maanden gestart worden na event y of niet?"). Wanneer het doel is om het effect van het starten van een behandeling te vergelijken met het niet starten van deze behandeling kan een random inclusiedatum gekozen worden voor de niet-starters, of zal een sequentiële trial methode gebruikt moeten worden om correct om te gaan met baseline. Onderzoekers die observationele studies uitvoeren zullen daarom moeten beschikken over geschikte methodologische expertise en een gedegen opleiding om deze methoden correct toe te passen, of iemand moeten inschakelen die deze expertise heeft. Wanneer gebrekkige methoden worden toegepast, zullen gebrekkige antwoorden worden verkregen.

Lopende systematische replicaties van gerandomiseerde trials met observationele data, zoals het RCT-DUPLICATE initiatief en andere inspanningen zijn daarom essentieel om aan te tonen dat observationele onderzoeken tot dezelfde conclusies leiden als gerandomiseerde onderzoeken, mits adequaat uitgevoerd.



Deze inspanningen zullen ook waardevolle inzichten geven onder welke omstandigheden (dat wil zeggen voor welke onderzoeksvragen, analytische methoden en databronnen) men zonder randomisatie tot causale conclusies kan komen in observationele data, en de effecten van behandelingen kan bestuderen zonder randomisatie. Deze initiatieven maken gebruik van principiële causal inference methoden en proberen ook de in- en exclusiecriteria van trials zo nauwkeurig mogelijk na te bootsen, om ervoor te zorgen dat de bevindingen niet verschillen door toepassing van gebrekkige methoden of verschillende patiëntpopulaties. Dergelijke calibratiestudies zullen ook op het gebied van nierziekten moeten worden uitgevoerd.

Bovendien zijn niet alle observationele studies even gevoelig voor confounding. Of een observationele studie tot causale conclusies kan komen hangt sterk af van de onderzoeksvraag, die wel eens de "axis of haphazardness" is genoemd. Farmaco-epidemiologische studies die schadelijke, onbedoelde effecten van medicijnen onderzoeken hebben minder last van (residuele) confounding dan studies die (on)bedoelde gunstige effecten onderzoeken. Regelgevende instanties maken inderdaad al langdurig gebruik van bewijs verkregen uit observationele studies om de veiligheid van medicijnen te beoordelen. Daarnaast zijn onderzoeken waarin twee geneesmiddelen met elkaar worden vergeleken die voor vergelijkbare indicaties worden voorgeschreven minder vatbaar voor confounding dan onderzoeken waarin een geneesmiddel wordt vergeleken met geen geneesmiddel. Het is moeilijk om in observationele data vragen te bestuderen waarin het starten van een medicijn wordt vergeleken met het niet starten, aangezien het starten van de behandeling een marker is voor de ernst van de ziekte of - in het geval van preventieve behandelingen - een marker van gezondheidszoekend gedrag; beide bronnen van bias zijn mogelijk niet volledig vastgelegd in de observationele data. In deze gevallen kan de selectie voor behandeling zo sterk zijn dat randomisatie noodzakelijk is. Desalniettemin bestaan er succesvolle toepassingen in de literatuur.

Het vermogen om causale conclusies te trekken hangt ook sterk af van de data die worden gebruikt en de variabelen die in de database beschikbaar zijn. Alvorens aan een onderzoek te beginnen, dienen onderzoekers zich af te vragen of de data van voldoende kwaliteit zijn voor het betreffende onderzoek. Wanneer voldoende gedetailleerdheid in blootstelling, uitkomsten of covariabelen ontbreekt, kan men ervoor kiezen om niet verder te gaan met de analyse. Specifiek voor nierziekten is de beschikbaarheid van routinematige laboratoriummetingen zoals creatinine en albuminurie vaak essentieel om adequaat te corrigeren voor confounding.

Hoewel de afwezigheid van ongemeten confounding een assumptie blijft die niet kan worden geverifieerd, moeten onderzoekers deze assumptie zorgvuldig en zo goed mogelijk rechtvaardigen, wat natuurlijk het moeilijke gedeelte van

epidemiologisch onderzoek is. Verschillende analyses kunnen worden gebruikt om het vertrouwen in de validiteit van bevindingen te versterken, zoals positieve en negatieve controle analyses, die zowel uitkomsten als cohorten zouden kunnen zijn waarvoor we bepaalde associaties verwachten. Deze analyses kunnen worden uitgevoerd om te onderzoeken of het haalbaar is om een bepaalde vraag te beantwoorden, zelfs voordat de primaire analyse is uitgevoerd. Indien resultaten uit gerandomiseerde trials beschikbaar zijn, kunnen de resultaten van de observationele analyse worden vergeleken met de resultaten van de trial, in acht nemend of dezelfde behandelstrategieën en studiepopulaties zijn onderzocht. Daarnaast moet er adequaat worden gecorrigeerd voor confounding. Wanneer bijvoorbeeld tijdsafhankelijke confounding aanwezig is, zijn G-methods zoals inverse probability weighting van marginal structural models vereist om juiste schattingen te krijgen. Propensity scores zijn een populaire methode om te corrigeren voor gemeten confounders. De vele ontwikkelingen in de propensity score-methodologie bieden een grote flexibiliteit voor het specificeren van de doelpopulatie (op welke populatie zijn onze resultaten van toepassing), balans in confounders, en precisie. Belangrijk is dat de balans op gemeten en ongemeten confounders voorafgaand aan de analyse moet worden gecontroleerd. In propensity score analyses worden ongemeten confounders alleen gebalanceerd voor zover ze gecorreleerd zijn met gemeten variabelen die zijn geïncludeerd in de propensity score. Daarom is een belangrijke aanpak om te corrigeren voor residuele confounding door niet-geobserveerde factoren om te corrigeren voor zoveel mogelijk proxy's van de onderliggende confounder (bijvoorbeeld diagnoses, procedures, medicijnen, aantal ziekenhuisopnames), die vóór de start van de behandeling moeten worden gemeten om correctie voor causale mediators te voorkomen. Wanneer bepaalde confounders alleen beschikbaar zijn voor een subset van de populatie en niet kunnen worden gebruikt om voor te corrigeren, kan de balans in deze variabele na het matchen/wegen van de propensity score het vertrouwen verhogen dat andere niet-gemeten variabelen ook in evenwicht zijn. Naast confounding moeten ook andere bronnen van vertekening worden onderzocht door middel van sensitiviteitsanalyses, zoals differentiële misclassificatie. Bij het onderzoeken van een 30% GFR-afname als uitkomst, moet men bijvoorbeeld controleren of beide behandelarmen dezelfde intensiteit van creatininetesten hebben tijdens follow-up. Kwantitatieve bias-analyses kunnen worden gebruikt om de invloed van resterende systematische vertekeningen te onderzoeken.

Bovendien kunnen verschillende observationele studies die verschillende causale methoden toepassen, worden gebruikt om bewijs te trianguleren, aangezien elke methode zijn eigen specifieke assumpties heeft. Op dit punt zijn er grote kansen voor uitwisseling van (quasi-experimentele) methoden uit andere wetenschapsgebieden, zoals regressiediscontinuïteit. Ten slotte kunnen verschillende andere maatregelen de reproduceerbaarheid van observationele

studies vergroten. Dit omvat preregistratie van observationele studies, de publicatie van codes en het volgen van rapportagerichtlijnen. In wezen zou het proces van het uitvoeren van observationele onderzoeken het regelgevende indieningsproces van gerandomiseerde onderzoeken moeten nabootsen. Er mogen bijvoorbeeld geen behandeling-uitkomstanalyses worden uitgevoerd totdat het protocol volledig is gespecificeerd en geregistreerd.

Concluderend zal het verkrijgen van bewijs uit niet-experimentele en experimentele studies belangrijk blijven, aangezien beide bronnen van bewijs elkaar aanvullen. Goed uitgevoerde observationele studies kunnen waardevol bewijs leveren voor de besluitvorming op het gebied van nierziekten, maar ook voor de geneeskunde in het algemeen. Het enige dat we moeten doen is het beantwoorden van de juiste vragen met de juiste methoden.