

Statistiska metoder för utvärdering av heterogenitet i kausala effekter. Två exempel med diabetesläkemedel baserar på registerdata från Nationella Diabetesregistret (NDR)

2002-08-24

SAMMANFATTNING

Projektet handlar om att undersöka den externa validiteten hos kliniska fas3-prövningar av två orala diabetesläkemedel, samt att utvärdera graden av heterogenitet i behandlingseffekt av två orala diabetesläkemedel i ett observationellt datamaterial.

Två metoder används, där den första utvärderar enbart generaliserbarhet och bygger på klassiska epidemiologiska propensity-score-metoder för att studera av effekten både bland personer som uppfyller urvalskriterier för en klinisk prövning och därmed är inkluderbara i studien, samt personer som inte uppfyller dessa kriterier. Den andra metoden utvärderar graden av heterogenitet i effekt och bygger på en kombination av maskininlärning och kausal inferens för att skatta en betingad behandlingseffekt givet en uppsättning variabler.

Båda metoderna appliceras på samkörda data från Nationella Diabetesregistret (NDR), Patientregistret, läkemedelsregistret och LISA innehållande personer som initierar behandling med liraglutide, dapagliflozin eller SU i tillägg till metformin. Urvalskriterier till två kliniska prövningar NN2211-1575 som utvärderar liraglutide jämfört med SU i tillägg till metformin och D1690C00004 som utvärderar dapagliflozin jämfört med SU används som exempel på urvalskriterier.

Resultaten antyder att utfallet från både NN2211-1575 och D1690C00004 kan generaliseras utanför den population som inkluderas i respektive prövning och att effekten av både liraglutide och dapagliflozin med avseende på förändring av HbA1c är heterogen så att effekten beror av pre index HbA1c i båda fallen och även av kön för liraglutide.

Projektet är en fortsättning på tidigare projekt kring generaliserbarhet av kliniska prövningar av orala diabetesläkemedel. De tidigare projekten handlade om att justera resultaten från en klinisk prövning för det urvalsfel som kan uppstå som en konsekvens av att individer inom studiens urvalskriterier kan ha olika benägenhet att komma med i prövningen vilket kan göra att studien sammansättning av patienter skiljer sig från vad man finner i kliniks praxis även inom studien urvalskriterier.

Introduktion

Projektet handlar om att undersöka den externa validiteten hos kliniska fas3-prövningar av två orala diabetesläkemedel, samt att utvärdera graden av heterogenitet i behandlingseffekt av två orala diabetesläkemedel i ett observationellt datamaterial.

Projektet utgör en fortsättning på två tidigare projekt kring generaliserbarhet av resultat från kliniska prövningar. Det första projektet var en litteraturgenomgång och det andra ett praktiskt exempel där de statistiska metoder som identifierades i det första projektet testades in praktiken med hjälp av data från två autentiska fas3-studier av två olika orala diabetesläkemedel i kombination med data från Nationella Diabetesregistret (NDR), läkemedelsregistret och patientregistret. Rapporter från de två tidigare projekten kan laddas ner från TLV.

[https://www.tlv.se/download/18.3d61c3c3167c695651ef749/1545305039255/Pilot2-Metod for utvardering av behandlingseffekter i klinisk vardag med en jamforelsegrupp fran en klinisk provning.pdf](https://www.tlv.se/download/18.3d61c3c3167c695651ef749/1545305039255/Pilot2-Metod%20for%20utvardering%20av%20behandlingseffekter%20i%20klinisk%20vardag%20med%20en%20jamforelsegrupp%20fran%20en%20klinisk%20provning.pdf)

[https://www.tlv.se/download/18.2c1b1aa6167c5e758eb27031/1545390635087/Uppf%C3%B6ljning lakemedelsanvandning och behandlingseffekter i klinisk%20vardag.pdf](https://www.tlv.se/download/18.2c1b1aa6167c5e758eb27031/1545390635087/Uppf%C3%B6ljning%20av%20lakemedelsanvandning%20och%20behandlingseffekter%20i%20klinisk%20vardag.pdf)

De två tidigare projekten var inriktade på att metoder för att utvärdera generaliserbarhet av en klinisk prövning inom en given population, definierad av inklusionskriterier och exklusionskriterier för den aktuella prövningen genom att korrigera det fel som uppstår när en prövning inte är ett obundet slumpmässigt urval ur den population som skall studeras vilket gör att fördelningen av individers karaktäristiska skiljer sig mellan de som inkluderas i prövningen och den population man vill generalisera till även inom urvalskriterierna till prövningen.

En naturlig följdfråga är hur effekten av behandlingen ser ut bland personer som inte kunde varit med i prövningen. Detta kan inte beläggas med hjälp av data från prövningen utan blir en extrapolation utanför den undersökta populationen. Frågan om effekten av en intervention är lika stor i den del av den del av populationen som inte kunde varit med i prövningen i relation till de som kunde varit med handlar om heterogenitet i effekten av interventionen.

Användning av förskrivningsläkemedel kan spåras via läkemedelsregistret och för flera specifika sjukdomar finns relevanta kliniska patientdata i kvalitetsregister. Här har vi valt att studera två orala diabetesläkemedel, liraglutid och dapagliflozin, i jämförelse med en referensbehandling.

Frågeställningar

Den primära frågeställningen är att beskriva hur effekten av en intervention varierar i en population, och speciellt hur stor effekten av interventionen är bland personer som kunde, respektive inte kunde varit med i en specifik klinisk prövning av den intervention som undersöks och därmed avgöra om resultatet från prövningen kan vara generaliserbart utanför den populations som inkluderades i prövningen. Primära effektvariabler är i samtliga delar förändring i HbA1c, vikt, systoliskt blodtryck och eGFR efter 6 månaders uppföljning.

Metod

I projektet används två olika typer av metoder för att utvärdera heterogenitet i kausala effekter. Den enklaste studerar bara skillnaden i effekt mellan två grupper av personer, de som kunde varit med i en prövning givet urvalskriterier och de som inte kunde varit med hjälp av propensityscoreviktning som appliceras på två disjunkta delar av populationen. Den andra metoden bygger på maskininlärning där flexibla prediktionsmodeller används tillsammans med metoder för kausal inferens för att direkt skatta en betingad kausal effekt för varje individ givet individdata.

Den hittills vanligaste metoden för att undersöka om effekten av en intervention är heterogen, dvs är olika i olika subgrupper, är att utföra ett antal subgruppsanalyser där den undersökta populationen delas in i ett antal olika grupper varpå man gör en analys inom varje grupp. Detta leder ofta till att man vill undersöka ett stort antal subgrupper och det kan vara svårt att veta hur man på ett bra sätt skall dela in populationen i grupper, speciellt med avseende på kontinuerliga variabler som ålder eller kliniska variabler. I ett observationellt datamaterial måste dessutom en utvärdering av en intervention justeras för störfaktorer (eng. confounding) vilket då skall göras separat i varje subgrupp.

Båda metoderna som används i projektet bygger på liknande antagande om att det inte finns andra variabler som påverkar exponering, utfall ock specifikt även benägenheten att vara med i en klinisk prövning än de vi har data på och inkluderar i modellerna. Vidare bygger analyserna på att benägenheten att sakna värden på de variabler pre index som används endast beror på andra variabler vi har data på, dvs missing at random. Benägenheten att sakna uppgift om effektvariabler antas vara helt oberoende av andra variabler, dvs missing completely at random. Slutligen antas de som har data på effektvariabler i de tidsfönster som används vara representativa för samtliga personer.

För att beskriva metoderna behöver vi lite matematiska verktyg och vi utgår från Rubins ramverk för kausal inferens, namngett av Holland (1986) och utvecklat i en serie artiklar av Rubin (, 1974; 1975; 1976; 1977; 1978; 1979; 1980) som bygger på potentiella utfall. Vi koncentrerar oss på fallet med en binär exponering, eller intervention, som alltså har två nivåer, vilka vi kallar 0 och 1. För dessa båda nivåer av exponering definierar vi för varje individ $i = 1, \dots, N$ de potentiella utfallen $Y_i(0)$ och $Y_i(1)$. Dessa båda utfall är stokastiska variabler, eller i dagligt tal slumpvariabler, och $Y_i(0)$ representerar det utfall som individ i skulle uppleva om hen får interventionen 0, medan $Y_i(1)$ representerar det utfall som hen skulle upplevt under interventionen 1. Med hjälp av dessa två potentiella utfall kan vi definiera den individuella kausala effekten av interventionen som skillnaden mellan de två potentiella utfallen, dvs som $\Delta_i = Y_i(1) - Y_i(0)$. Vi utgår från att varje person bara kan erhålla en av de två interventionerna vilket innebär att vi egentligen inte kan observera den individuella kausala effekten direkt, men konstruktionen som sådan hjälper oss att förstå vad den kausala effekten är, dvs skillnaden i utfall för en individ om hen skulle välja den ena eller den andra av två möjliga interventioner.

På grupp nivå kan vi med hjälp av de individuella effekterna definiera den genomsnittliga effekten i en population, population average treatment effect, eller PATE som

$$\Delta = \frac{1}{N} \sum_{i=1}^N (Y_i(1) - Y_i(0)).$$

Inte heller PATE kan observeras direkt, men vi kan göra en naiv uppskattning av hur stor den kan vara genom att jämföra det genomsnittliga utfallet bland de som får interventionen 0 med det genomsnittliga utfallet bland de som får interventionen 1. Med hjälp av den observerade

interventionen T_i (som är 0 eller 1 för varje individ) och det observerade utfallet Y_i för individ i får vi den enklaste skattningen av PATE som

$$\hat{\Delta}_{sim} = \frac{1}{n_1} \sum_{i=1}^N Y_i T_i - \frac{1}{n_0} \sum_{i=1}^N Y_i (1 - T_i).$$

En sådan enkel skattning fungerar bra om det inte finns några systematiska skillnader mellan personer som får interventionerna, vilket kan vara fallet i en väl designad randomiserad prövning, men denna enkla skattning fungerar sällan eller aldrig i ett observationellt datamaterial. Baserat på Rubins ramverk har det utvecklats en rad metoder för att under mer eller mindre realistiska antaganden skatta PATE även baserat på registerdata. Det centrala antagandet handlar om att endast en begränsad uppsättning variabler som vi känner till och har data på får påverka både valet av intervention och utfallet. Denna typ av variabler kallas störfaktorer på svenska eller confounders på engelska och antagandet kan formuleras som att det inte får finnas några observerade, eller dolda störfaktorer. Matematiskt kan detta beskrivas som att exponering T och utfall är Y oberoende betingat på en uppsättning variabler $\mathbf{X} = (X_1, \dots, X_p)$, dvs $Y \perp T \mid \mathbf{X}$.

Eftersom vi vill använda just registerdata för att utvärdera effekten av en intervention och speciellt hur denna varierar i en population så blir antagandet om att det inte får finnas några dolda störfaktorer gemensamt för alla metoder vi använder i projektet.

Under antagandet om att det inte finns några dolda störfaktorer kan PATE skattas med flera olika metoder. Vanliga val är att anpassa en regressionsmodell som modellerar utfallet som en (parametrisk) funktion av exponering och störfaktorer. Som exempel kan vi använda en linjär modell för ett kontinuerligt utfall och en enkel sådan modell skulle kunna vara

$$Y_i = \mu + \tau T + \beta_1 X_1 + \dots + \beta_p X_p,$$

där τ är effekten av intervention 1 jämfört med intervention 0. Vi får en skattning av den kausala effekten genom att skatta τ , dvs

$$\hat{\Delta}_{reg} = \hat{\tau}.$$

En annan vanlig metod, som återigen bygger på antagande om att det inte finns några dolda störfaktorer, för att skatta den kausala effekten är viktning med hjälp av propensity score, inverse probability of treatment weighting (IPTW) på engelska. Propensity score definieras som sannolikheten att få intervention 1, dvs som $e_i(\mathbf{X}_i) = P(T = 1 \mid \mathbf{X}_i)$. Propensity score måste skattas baserat på data och det är vanligt att detta sker med en logistisk regression som innehåller alla potentiella störfaktorer som oberoende variabler, men man kan också använda någon form av flexibel prediktionsmodell av typen maskininlärning vilket har flera fördelar. När man väl skattat sina propensity score får man en skattning av den genomsnittliga kausala effekten i populationen som

$$\hat{\Delta}_{iptw} = \frac{\sum_{i=1}^N \frac{Y_i T_i}{\hat{e}_i}}{\sum_{i=1}^N \frac{T_i}{\hat{e}_i}} - \frac{\sum_{i=1}^N \frac{Y_i (1 - T_i)}{(1 - \hat{e}_i)}}{\sum_{i=1}^N \frac{1 - T_i}{(1 - \hat{e}_i)}}.$$

Detta kan se lite avskräckande ut men är egentligen ingen annat än differensen mellan två viktade medelvärden där vikterna definieras som sannolikheten att få den intervention man faktiskt fick.

Propensity score kan också användas som grund för stratifierade analyser eller för att skapa matchade analysdataset där man direkt kan skatta den kausala effekten, men det tar vi inte upp här.

Alla tre sätten $\hat{\Delta}_{sim}$, $\hat{\Delta}_{reg}$ och $\hat{\Delta}_{iptw}$ ger en enda skattning av den genomsnittliga kausala effekten, PATE och kan inte användas för att utvärdera om det finns heterogenitet i den kausala effekten. Vi

började med att definiera en individuell kausal effekt som vi snabbt konstaterade att vi inte kan observera och därmed inte uppskatta, och har nu sett tre sätt att skatta den genomsnittliga effekten vilket är relevant om de individuella effekterna inte skiljer sig åt (så mycket). Syftet med detta projekt var att undersöka heterogenitet i effekten så vi behöver se oss om efter andra metoder är de klassiska.

Prezler & Kaizar

Prezler och Kaizar (2011) har publicerat en metod för att utvärdera generaliserbarhet av resultat från en del av populationen till hela populationen genom att dela in populationen i två delar. En del I innehåller N_I individer som kunde varit med i en given klinisk prövning om de fått chansen. Dvs dessa individer uppfyller studiens inklusionskriterier men inte något av dess exklusionskriterier. Vi kan kalla dessa individer för inkluderbara, eller eligible på engelska. Den resterade delen av populationen E består av N_E individer som inte kunde varit med i prövningen. För var och en av dessa två delpopulationer kan man definiera en genomsnittlig kausal effekt, en subpopulation average treatment effect (SPATE) som

$$\Delta_I = \frac{1}{N_I} \sum_{i=1}^N (Y_i(1) - Y_i(0)) I_i$$

och

$$\Delta_E = \frac{1}{N_I} \sum_{i=1}^N (Y_i(1) - Y_i(0)) (1 - I_i)$$

där I_i är lika med 1 för personer som tillhör I och alltså inte kunde varit med i prövningen om de fått chansen och 0 för övriga individer. Med hjälp av dessa kan man definiera skillnaden i effekt bland de som kunde varit med i prövningen och hela populationen som

$$\gamma = \Delta_I - \Delta = \pi_E (\Delta_I - \Delta_E),$$

där $\pi_E = N_E / (N_E + N_I)$ är andelen personer som är inkluderbara. För att få en skattning av generaliseringsfelet γ kan man ersätta parametrarna Δ_I och Δ_E med skattningar $\hat{\Delta}_I$ och $\hat{\Delta}_E$ baserade på någon av metoderna ovan, dvs en skattning med en regressionsmodell eller en skattning med hjälp av propensity score-viktning. Till detta projekt har vi valt viktning med hjälp av propensity score där propensity score skattas med maskininlärning, i detta fall gradient boosting med shrinkage satt till 0.01 och det antal träd som ger optimal balans.

Metoden bygger på både ett antagande om att det inte finns några okända störfaktorer som påverkar exponering och utfall och på antagandet att det inte finns några andra faktorer som påverkar benägenheten att vara med i en prövning och utfallet av interventionen än de vi kan observera och inkludera i vår analys

Causal Forest

Vi har tidigare konstaterat att vi inte varken observera eller på något sätt skatta den individuella kausala effekten eftersom vi för varje person bara kan observera det ena av de två potentiella utfallen $Y_i(0)$ och $Y_i(1)$. Låt oss nu definiera den betingade kausala effekten som

$$\Delta(\mathbf{X}) = E[Y_i(1) - Y_i(0) | \mathbf{X}],$$

vilket vi kan tolka som den genomsnittliga kausala effekten för personer har en specifik uppsättning egenskaper $\mathbf{X} = (X_1, \dots, X_p)$. Uppgiften blir nu att hitta en bra skattning av $\Delta(\mathbf{X})$. Om det inte finns några andra störfaktorer än variablerna \mathbf{X} så vore det rimligt att dela in alla personer i små grupper där alla har ungefär samma värden på \mathbf{X} och i var och en av dessa grupper skapa en skattning av $\Delta(\mathbf{X})$. Detta är i praktiken så man idag utvärderar heterogenitet i effekten av en intervention, nämligen att dela in populationen i subgrupper och göra en analys i varje subgrupp. Detta har fler nackdelar. Subgrupperna kan lätt bli ganska många och därmed små och det är svårt att veta vilka subgrupper som är relevanta.

Causal Forest (Athey 2019) är en sammansmältning av en metod inom maskininlärning, Random Forest, och kausal inferens. Först skapas ett stort antal urval av populationen utan återläggning. I varje urval används en del av datamaterialet till att stegvis skapa en gruppuppdelning och en annan del av materialet till att inom varje grupp bestämma en skattning av den kausala effekten av interventionen inom gruppen. Gruppuppdelningen sker på en så att heterogeniteten i de skattade kausala effekterna mellan grupperna maximeras. Detta resulterar en stor mängd olika uppdelningar och effektskattningar som sedan vägs samman till en skattning av den betingade kausala effekten för varje individ utifrån dennes egenskaper. Denna skattning förses också med en uppskattning av den statistiska osäkerheten i form av ett standardfel. De individuella betingade kausala effekterna kan sedan analyseras för att hitta variabler som är associerade med variabilitet i den kausala effekten och därpå kan de betingade kausala effekterna aggregeras i lämpliga grupper som män och kvinnor, unga och gamla eller personer med vissa värden på kliniska variabler som exempelvis HbA1c eller vikt.

I matematiska termer handlar det om att först skapa $b = 1, \dots, B$ stycken slumpmässiga urval $S_b \subseteq \{1, \dots, n\}$ från populationen, för varje urval skapa ett träd via rekursiv partitionering och sedan skattas den betingade kausala effekten som.

$$\hat{\Delta}(x) = \frac{\sum_{i=1}^N \alpha_i(x) (Y_i - \hat{m}_i^{(-i)}(X_i)) (T_i - \hat{e}_i^{(-i)}(x_i))}{\sum_{i=1}^N \alpha_i(x) (T_i - \hat{e}_i^{(-i)}(x_i))^2},$$

där $m_i(x) = E[Y_i | X_i = x]$ är en modell för utfallet Y som en funktion av variablerna X oavsett intervention och $e_i(x) = P[T_i = 1 | X_i = x]$ är propensity score. Notationen $(-i)$ indikerar att observation i inte används i beräkning och refererar till att skattningen är baserad på out-of-bag-observationer enligt terminologi för random forest. Återstående storhet $\alpha_i(x)$ styr hur skattning viktas samman och definieras som

$$\alpha_i(x) = \frac{1}{B} \sum_{b=1}^B \frac{1(X_i \in L_b(X), i \in S_b)}{|\{i: X_i \in L_b(X), i \in S_b\}|}$$

där $L_b(X)$ är det löv i träd b som innehåller X .

Metoden är som synes relativt komplicerad att beskriva matematiskt varför den intresserade läsaren hänvisas till Athey & Wager (2019).

Athey & Wagers metod resulterar i individuella betingade kausala effekter, dvs en skattning av den betingade kausala effekten för varje individ i den analyserade populationen, tillsammans med en standardfelet för varje skattning. Dessa kan användas för att identifiera variabler som är associerade med heterogenitet i de kausala effekterna, exempelvis genom maskininlärning. Det skattade kausala

effekterna kan också plottas mot olika bakgrundsvariabler och aggregeras i olika grupper samt presenteras med konfidensintervall.

Programvara

Alla analyser utförs med R 3.6.1. Imputation sker med mice, propensity score skattas med twang och causal random forest skattas med hjälp av grf.

Data

I studien används data från NDR, läkemedelsregistret, patientregistret och LISA. Samkörningen av data har skett på SCB och Socialstyrelsen och efter att personnumret krypterats har data sänts till Registercentrum Västra Götaland där det samkörda materialet bearbetats i en säker analysmiljö. Utifrån samkörda data har två studiepopulationer skapats. Den första innehåller 18587 personer som hämtat ut minst två recept på liraglutide (n=2971) eller SU (n=15616) i tillägg till metformin under tidsperioden 2010-01-31 till 2017-12-31 och den andra innehåller 10226 personer som hämtat ut minst två recept på dapagliflozing (n=1051) eller SU (n=9175) i tillägg till metformin under tidsperioden 2013-07-09 to 2017-12-31.

I båda studierna definieras indexdatumet som datum för det första uttagen av liraglutide, dapagliflozin eller SU. Behandling med metformin definieras som minst två uttag inom 210 dagar innan indexdatum.

Pre indexobservationer för kliniska variabler definieras som det sista observerade värdet under året innan indexdatumet. Post indexobservationer efter 6 månaders uppföljning definieras som värdet närmast 6 månader efter indexdatum inom intervallet 3 till 9 månader efter index.

I analysdata ingår observationer på RCT eligibility, ålder, kön, BMI, vikt, rökning, HbA1c, systoliskt blodtryck, diastoliskt blodtryck, LDL, HDL, triglycerider, eGFR, retinopati, mikroalbuminuri, makroalbuminuri, fysisk aktivitet, blodtryckssänkande behandling, lipidsänkande behandling, utbildning, inkomst, civilstånd, födelseplats (Sverige, EU, resten av världen), tidigare CHD, tidigare AMI, tidigare stroke, tidigare CVD, tidigare HF, tidigare AF, tidigare njursvikt, tidigare hyperglykemi, tidigare psykiatrisk sjukdom, tidigare alkohol eller drogmissbruk, tidigare cancer samt tidigare demens.

RCT eligibility har definierats utifrån inklusionskriterier och exklusionskriterier från två kliniska studier, NN2211-1575 som utvärderar liraglutide jämfört med SU i tillägg till metformin och D1690C00004 som utvärderar dapagliflozin jämför med SU i tillägg till metformin.

Saknade pre indexdata har imputerats med hjälp av en enkel stokastisk imputation med hjälp av multiple chained equations medan endast personer med valida observationer av varje effektvariabel ingår i respektive analys vilket får till följd att det verkliga antalet observationer som ingår i varje analys varierar något

Datamaterialet är samma som använts i ett av de två tidigare projekten vilket gjort att resurs och tidsåtgång för att skapa analysdatat till detta projekt varit relativt begränsad. För en beskrivning av resurs och tidsåtgång för att skapa datamaterialet till det tidigare projektet hänvisas till tidigare rapporter.

Resultat

Studien av liraglutide jämfört med SU omfattar totalt 18587 varav 2971 initierat behandling med liraglutide och 15616 initierat behandling med SU. Bland de totalt 18587 personerna i studien var 5119 (27.5%) möjliga att inkludera i NN2211-1575 om de fått chansen medan resterande 13468 (72.5%) inte kunde varit med.

Studien av dapagliflozin jämfört med SU omfattar totalt 10226 varav 1051 initierat behandling med dapagliflozin och 9175 initierat behandling med SU. Bland de totalt 10226 personerna i studien var 2672 (26.1%) möjliga att inkludera i D1690C00004 om det fått chansen medan resterande 7554 (73.9%) inte kunde varit med.

Sammantaget är personer som initierat behandling med liraglutid eller dapagliflozin något yngre och har något kortare duration än personer som initierar behandling med SU (tabell1). De är också tyngre och har högre HbA1c.

Skillnaderna mellan personer som kunde varit med i respektive klinisk prövning och de som inte kunde varit med är med få undantag, såsom makroalbuminuri och tidigare cancer som utgör ett exklusionkriterium samt cancer tidigare cancer ändå relativt liten (tabell 2)

Antalet observationer som ingår i analyserna av respektive effektvariabel varierar något eftersom endast personer med observerade värden på respektive effekt variabel kan ingå i analyserna (tabell 3).

Beskrivande statistik

Beskrivande statistik för de observationer som ingår i respektive analys kan hittas i appendix 1.

Tabell 1 Beskrivande statistik redovisat som medelvärden med standardavvikelse för kontinuerliga variabler eller antal med procent för kategoriska variabler uppdelat på exponering inom de två analyspopulationerna

Variabel	Liraglutide vs SU		Dapagliflozin vs SU	
	SU	liraglutide	SU	dapagliflozin
n	15616	2971	9175	1051
Age	62.97 (10.56)	56.18 (10.43)	63.66 (10.70)	59.45 (10.01)
Female gender	5867 (37.6)	1226 (41.3)	3398 (37.0)	409 (38.9)
Diabetes duration	7.50 (5.68)	6.42 (5.33)	8.22 (5.97)	7.44 (5.69)
BMI	30.57 (5.13)	35.84 (6.03)	30.21 (5.11)	32.50 (5.47)
Weight(kg)	90.36 (17.66)	107.77 (20.57)	89.34 (17.67)	97.34 (19.59)
Smoking	1673 (16.4)	297 (14.7)	975 (15.5)	96 (13.9)
HbA1c (mmol/mol)	62.85 (12.60)	65.12 (14.58)	62.51 (12.80)	66.37 (14.84)
SBP(mmHg)	135.80 (15.03)	135.29 (14.91)	135.60 (14.90)	135.73 (15.09)
DBP(mmHg)	78.72 (9.50)	81.06 (9.48)	78.66 (9.54)	81.14 (9.64)
LDL (mmol/L)	2.57 (0.90)	2.63 (0.92)	2.51 (0.91)	2.59 (0.94)
HDL (mean (SD))	1.18 (0.33)	1.13 (0.31)	1.19 (0.34)	1.16 (0.31)
Triglycerides	2.09 (1.39)	2.39 (1.68)	2.09 (1.53)	2.19 (1.37)
Retinopathy	1749 (21.6)	336 (20.7)	1278 (22.9)	141 (22.9)
Micro albuminurea	1573 (18.6)	319 (20.2)	1020 (19.6)	118 (21.0)
Macro	359 (4.3)	81 (5.2)	180 (3.7)	18 (3.4)
Anti hypertensive	8484 (74.5)	1650 (74.2)	5197 (74.1)	603 (78.4)
Creatinin (mmol/L)	72.28 (17.15)	70.32 (16.82)	72.12 (17.23)	69.41 (15.72)
Physical Activity				
< time per week	1159 (12.6)	331 (18.4)	725 (12.4)	94 (14.8)
1-2 times per week	2054 (22.4)	406 (22.6)	1281 (21.9)	131 (20.6)
3-5 times per week	2200 (24.0)	405 (22.5)	1356 (23.1)	143 (22.5)
Daily	2505 (27.3)	347 (19.3)	1651 (28.2)	156 (24.5)
Never	1262 (13.7)	309 (17.2)	848 (14.5)	112 (17.6)
Lipid lowering	7493 (65.9)	1376 (62.4)	4661 (66.5)	504 (65.8)
Education				

Basic	5197 (35.2)	588 (22.2)	2907 (34.6)	215 (24.3)
College	6696 (45.3)	1463 (55.1)	3896 (46.3)	457 (51.6)
University	2874 (19.5)	603 (22.7)	1609 (19.1)	214 (24.2)
Marital status				
Married	8411 (55.9)	1364 (50.9)	4735 (55.2)	466 (52.1)
Separated	2661 (17.7)	516 (19.3)	1552 (18.1)	166 (18.6)
Single	2706 (18.0)	688 (25.7)	1517 (17.7)	221 (24.7)
Widowed	1260 (8.4)	111 (4.1)	776 (9.0)	41 (4.6)
Origin (%)				
Europe except	809 (5.2)	157 (5.3)	458 (5.0)	41 (3.9)
RoW	2830 (18.1)	333 (11.2)	1944 (21.2)	161 (15.3)
Sweden	11977 (76.7)	2481 (83.5)	6773 (73.8)	849 (80.8)
CHD	2272 (14.5)	347 (11.7)	1399 (15.2)	141 (13.4)
AMI	1025 (6.6)	174 (5.9)	628 (6.8)	71 (6.8)
Stroke	613 (3.9)	74 (2.5)	385 (4.2)	38 (3.6)
CVD	1575 (10.1)	242 (8.1)	975 (10.6)	105 (10.0)
HF	674 (4.3)	139 (4.7)	408 (4.4)	32 (3.0)
AF	1050 (6.7)	189 (6.4)	677 (7.4)	59 (5.6)
Renal dysfunction	134 (0.9)	35 (1.2)	90 (1.0)	11 (1.0)
Hyper glycemia	126 (0.8)	38 (1.3)	69 (0.8)	8 (0.8)
Psyciatric disorder	786 (5.0)	234 (7.9)	524 (5.7)	85 (8.1)
Alcohol or drug	290 (1.9)	82 (2.8)	193 (2.1)	27 (2.6)
Cancer	1583 (10.1)	233 (7.8)	1061 (11.6)	81 (7.7)
Dementia	32 (0.2)	2 (0.1)	31 (0.3)	1 (0.1)

Tabell 2 Beskrivande statistik redovisat som medelvärden med standardavvikelse för kontinuerliga variabler eller antal med procent för kategoriska variabler uppdelat på inkluderbarhet (eligible / non-eligible)

Variabel	Liraglutide vs SU		Dapagliflozin vs SU	
	Non eligible	Eligible	Non eligible	Eligible
N	13468	5119	7554	2672
Age	62.21 (10.98)	61.04 (10.37)	63.73 (10.77)	61.82 (10.42)
Female gender	5391 (40.0)	1702 (33.2)	2935 (38.9)	872 (32.6)
Diabetes duration	7.44 (5.82)	7.04 (5.20)	8.35 (6.11)	7.61 (5.48)
BMI	31.87 (5.93)	30.46 (4.77)	30.71 (5.41)	29.72 (4.53)
Weight(kg)	94.04 (19.98)	91.47 (17.51)	90.56 (18.49)	88.98 (16.70)
Smoking	1356 (15.9)	614 (16.5)	752 (14.9)	319 (16.4)
HbA1c (mmol/mol)	63.71 (14.37)	62.12 (9.05)	63.23 (14.30)	62.03 (9.22)
SBP(mmHg)	135.73 (15.49)	135.67 (13.84)	135.69 (15.44)	135.39 (13.48)
DBP(mmHg)	78.91 (9.73)	79.54 (9.05)	78.67 (9.76)	79.50 (9.06)
LDL (mmol/L)	2.56 (0.91)	2.64 (0.88)	2.49 (0.91)	2.59 (0.91)
HDL (mean (SD))	1.17 (0.33)	1.17 (0.31)	1.18 (0.34)	1.18 (0.32)
Triglycerides (mmol/L)	2.17 (1.46)	2.06 (1.39)	2.11 (1.53)	2.08 (1.48)
Retinopathy	1453 (21.5)	632 (21.4)	1009 (22.7)	410 (23.5)
Micro albuminurea	1361 (19.6)	531 (17.2)	891 (21.4)	247 (15.4)
Macro albuminurea (%)	440 (6.4)	0 (0.0)	198 (5.1)	0 (0.0)
Anti hypertensive	7241 (76.4)	2893 (70.0)	4316 (76.8)	1484 (68.5)
Creatinin (mmol/L)	72.58 (18.19)	70.46 (14.06)	72.49 (18.00)	70.11 (14.26)
Physical Activity				
< time per week	1079 (14.1)	411 (12.3)	630 (13.4)	189 (10.5)
1-2 times per week	1712 (22.4)	748 (22.4)	1027 (21.9)	385 (21.4)
3-5 times per week	1744 (22.8)	861 (25.8)	1036 (22.0)	463 (25.8)
Daily	1926 (25.2)	926 (27.7)	1287 (27.4)	520 (28.9)
Never	1174 (15.4)	397 (11.9)	719 (15.3)	241 (13.4)
Lipid lowering	6269 (66.4)	2600 (63.0)	3781 (67.4)	1384 (63.9)
Education				
Basic	4157 (33.1)	1628 (33.5)	2321 (33.9)	801 (32.7)
College	5884 (46.8)	2275 (46.9)	3173 (46.3)	1180 (48.2)
University	2527 (20.1)	950 (19.6)	1356 (19.8)	467 (19.1)
Marital status				

Married	7023 (54.9)	2752 (55.8)	3833 (55.0)	1368 (54.7)
Separated	2391 (18.7)	786 (15.9)	1298 (18.6)	420 (16.8)
Single	2356 (18.4)	1038 (21.1)	1197 (17.2)	541 (21.6)
Widowed	1016 (7.9)	355 (7.2)	645 (9.2)	172 (6.9)
Origin (%)				
Europe except Sweden	710 (5.3)	256 (5.0)	379 (5.0)	120 (4.5)
RoW	2338 (17.4)	825 (16.1)	1544 (20.4)	561 (21.0)
Sweden	10420 (77.4)	4038 (78.9)	5631 (74.5)	1991 (74.5)
CHD	2492 (18.5)	127 (2.5)	1488 (19.7)	52 (1.9)
AMI	1164 (8.6)	35 (0.7)	685 (9.1)	14 (0.5)
Stroke	647 (4.8)	40 (0.8)	403 (5.3)	20 (0.7)
CVD	1745 (13.0)	72 (1.4)	1047 (13.9)	33 (1.2)
HF	780 (5.8)	33 (0.6)	426 (5.6)	14 (0.5)
AF	1073 (8.0)	166 (3.2)	657 (8.7)	79 (3.0)
Renal dysfunction	158 (1.2)	11 (0.2)	99 (1.3)	2 (0.1)
Hyper glycemia	158 (1.2)	6 (0.1)	72 (1.0)	5 (0.2)
Psychiatric disorder	855 (6.3)	165 (3.2)	517 (6.8)	92 (3.4)
Alcohol or drug abuse	355 (2.6)	17 (0.3)	212 (2.8)	8 (0.3)
Cancer	1751 (13.0)	65 (1.3)	1115 (14.8)	27 (1.0)
Dementia	27 (0.2)	7 (0.1)	25 (0.3)	7 (0.3)

Tabell 3 Antal observationer som ingår i analyser av respektive utfall

Variabel	Liraglutide vs SU		Dapagliflozin vs SU	
	SU	liraglutide	SU	Dapagliflozin
Förändring i HbA1c	9541	1826	5884	605
Förändring i Vikt	8385	1659	5077	524
Förändring i SBP	9077	1726	5632	572
Förändring i eGFR	7762	1472	4678	503

Generaliserbarhetsbias enligt Prezzler & Kaizar

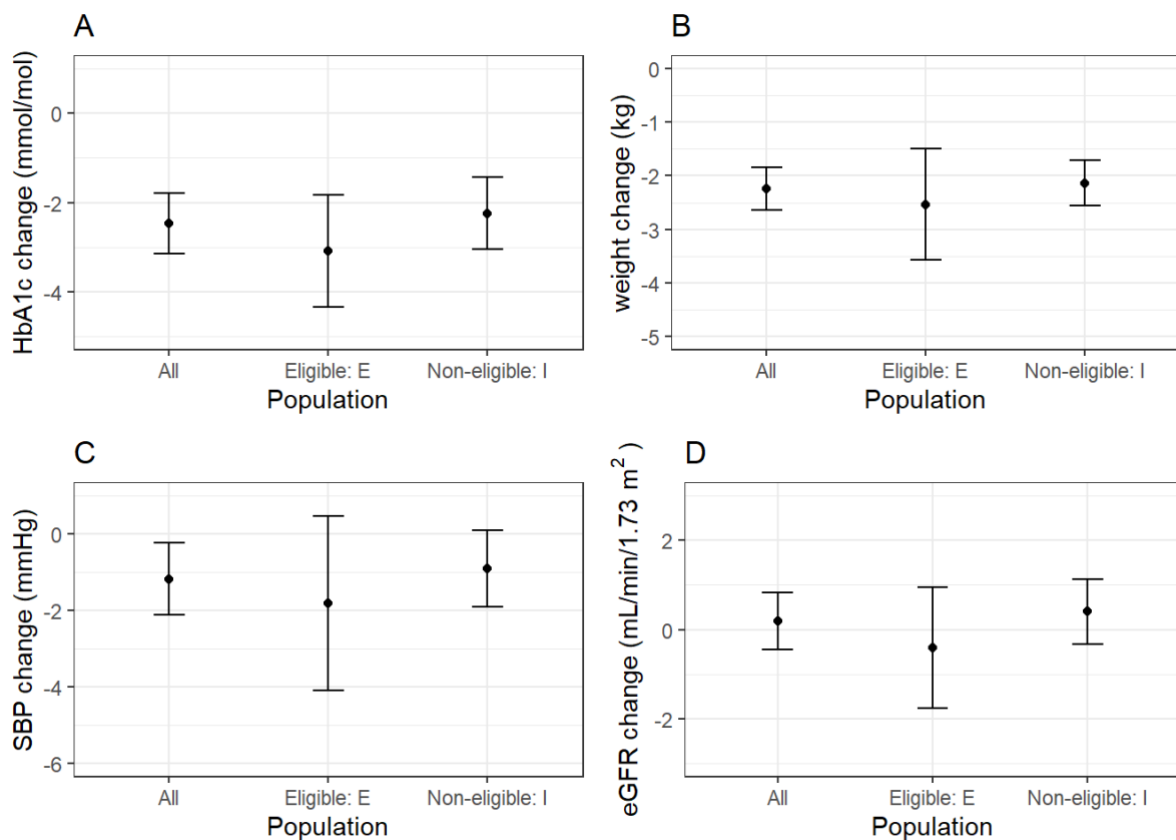
Den propensity score justerade skattade effekten vid 6 månaders uppföljning av liraglutide jämfört med SU och dapagliflozin jämfört med SU i hela populationen, bland de som kunde deltagit i respektive prövning om de fått chansen och de som inte kunde deltagit finns i tabell 4 och 5 samt i figur 1 och 2. Generellt är det små skillnader mellan personer som kunde och som inte kunde varit med i respektive prövning.

Tabell 4 Propensity-score-justerade minstakvadrat skattningar av förändring av HbA1c, vikt, systoliskt blodtryck och eGFR samt skillnad i förändring mellan liraglutide och SU med 95% konfidensintervall efter 6 månaders uppföljning

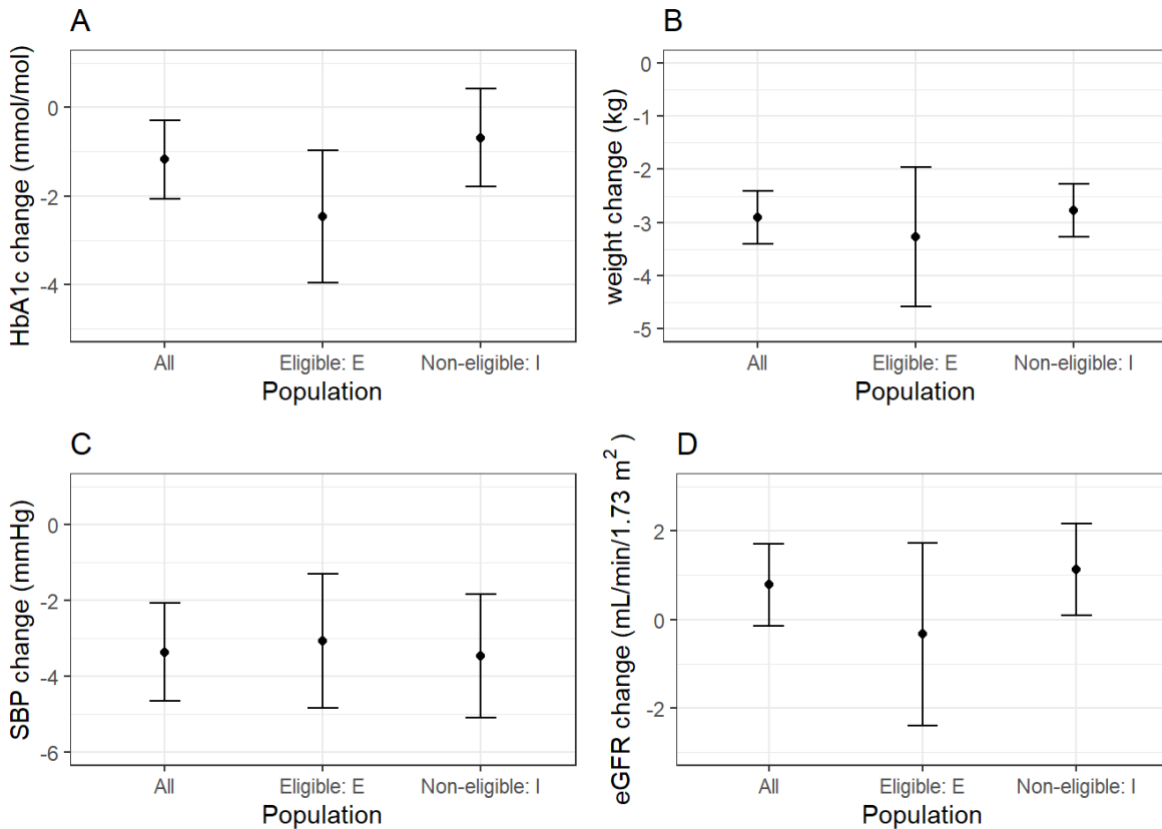
Variabel	Population	SU	liraglutide	Skattad skillnad	p-värde
HbA1c	All	-6.48[-6.71,-6.25]	-8.94[-9.58,-8.3]	-2.46[-3.15, -1.78]	<0.0001
HbA1c	Eligible: E	-6.17[-6.57,-5.77]	-9.26[-10.46,-8.07]	-3.09[-4.35, -1.84]	<0.0001
HbA1c	Non-eligible: I	-6.58[-6.87,-6.3]	-8.82[-9.57,-8.08]	-2.24[-3.04, -1.44]	<0.0001
Weight	All	-0.16[-0.31,-0.02]	-2.41[-2.78,-2.04]	-2.24[-2.65, -1.84]	<0.0001
Weight	Eligible: E	-0.06[-0.32,0.2]	-2.6[-3.57,-1.63]	-2.54[-3.58, -1.5]	<0.0001
Weight	Non-eligible: I	-0.21[-0.39,-0.04]	-2.35[-2.73,-1.97]	-2.14[-2.56, -1.71]	<0.0001
SBP	All	-0.88[-1.16,-0.61]	-2.06[-2.95,-1.16]	-1.17[-2.11, -0.24]	0.0140563
SBP	Eligible: E	-0.4[-0.87,0.07]	-2.21[-4.44,0.02]	-1.81[-4.09, 0.47]	0.1199139
SBP	Non-eligible: I	-1.12[-1.46,-0.78]	-2.01[-2.96,-1.07]	-0.89[-1.9, 0.11]	0.0809821
eGFR	All	-1.99[-2.19,-1.79]	-1.81[-2.41,-1.2]	0.19[-0.45, 0.82]	0.5673285
eGFR	Eligible: E	-1.7[-2.04,-1.35]	-2.1[-3.48,-0.72]	-0.4[-1.76, 0.96]	0.5621401
eGFR	Non-eligible: I	-2.15[-2.4,-1.91]	-1.75[-2.43,-1.07]	0.4[-0.32, 1.13]	0.2769842

Tabell 5 Propensity-score-justerade minstakvadrat skattningar av förändring av HbA1c, vikt, systoliskt blodtryck och eGFR samt skillnad i förändring mellan dapagliflozin och SU med 95% konfidensintervall efter 6 månaders uppföljning

Variabel	Population	SU	dapagliflozin	Skattad skillnad	p-värde
HbA1c	All	-6.06[-6.33,-5.79]	-7.47[-8.32,-6.62]	-1.18[-2.07, -0.28]	0.0097804
HbA1c	Eligible: E	-5.70[-6.23,-5.17]	-8.17[-9.58,-6.76]	-2.47[-3.97, -0.97]	0.0012633
HbA1c	Non-eligible: I	-6.15[-6.48,-5.82]	-6.83[-7.88,-5.79]	-0.68[-1.78, 0.42]	0.2245546
Weight	All	-0.1[-0.28,0.08]	-3.01[-3.47,-2.54]	-2.91[-3.41, -2.41]	<0.0001
Weight	Eligible: E	0.07[-0.31,0.46]	-3.2[-4.47,-1.93]	-3.27[-4.59, -1.95]	<0.0001
Weight	Non-eligible: I	-0.16[-0.35,0.04]	-2.94[-3.39,-2.49]	-2.78[-3.28, -2.28]	<0.0001
SBP	All	-0.57[-0.92,-0.22]	-3.93[-5.17,-2.69]	-3.36[-4.65, -2.07]	<0.0001
SBP	Eligible: E	0.05[-0.6,0.69]	-3.02[-4.68,-1.37]	-3.07[-4.84, -1.29]	0.0007232
SBP	Non-eligible: I	-0.81[-1.23,-0.39]	-4.28[-5.86,-2.7]	-3.47[-5.11, -1.83]	<0.0001
eGFR	All	-2.50[-2.75,-2.26]	-1.72[-2.62,-0.83]	0.78[-0.15, 1.71]	0.0998061
eGFR	Eligible: E	-2.00[-2.53,-1.47]	-2.33[-4.44,-0.22]	-0.33[-2.41, 1.74]	0.7520056
eGFR	Non-eligible: I	-2.65[-2.94,-2.35]	-1.52[-2.5,-0.53]	1.13[0.09, 2.16]	0.0325984



Figur 1 Propensity-score-justerade minstakvadrat skattningar av förändring av A) HbA1c, B) vikt, C) systoliskt blodtryck och D) eGFR samt skillnad i förändring mellan liraglutide och SU med 95% konfidensintervall efter 6 månaders uppföljning



Figur 2 Propensity-score-justerade minstakvadrat skattningar av förändring av A) HbA1c, B) vikt, C) systoliskt blodtryck och D) eGFR samt skillnad i förändring mellan dapagliflozin och SU med 95% konfidensintervall efter 6 månaders uppföljning

Heterogenitet i effekt med Causal Forest

Vi inleder med övergripande analyser för att undersöka om effekten är heterogen genom ett hypotestest av nollhypotesen att effekten är homogen för varje utfall (Tabell 6).

Analysen med causal forest ger stöd för att effekten av liraglutid jämfört med SU efter 6 månaders uppföljning är heterogen i den undersökta populationen med avseende på HbA1c ($p < 0.001$) och systoliskt blodtryck ($p = 0.01026$), men inte med avseende på vikt ($p = 0.3710$) eller eGFR ($p = 0.6800$)

Analysen med causal forest ger stöd för att det effekten av dapagliflozin jämfört med SU efter 6 månaders uppföljning är heterogen i den undersökta populationen med avseende på HbA1c ($p = 0.003$), men inte med avseende på systoliskt blodtryck ($p = 1.000$), vikt ($p = 0.5256$) eller eGFR ($p = 0.9487$)

Tabell 6 Skattning och test av heterogenitet i effekten av liraglutid jämfört med SU och dapagliflozin jämfört med SU med avseende på förändring i HbA1c, vikt, systoliskt blodtryck och eGFR efter 6 månaders uppföljning med p-värde

Effektvariabel	Liraglutid vs SU		Dapagliflozin vs SU	
	Skattning	p-värde	Skattning	p-värde
HbA1c	1.26943	<0.0001	1.13624	0.003280
Vikt	0.184009	0.3710	-611.29	1.000
SBP	0.79175	0.01026	0.057854	0.5256
eGFR	-0.22404	0.6800	-1.08614	0.9487

Analyserna av generalitetsbias med causal forest (tabell 7) ger över lag liknande resultat som Prezlers metod med skillnaden att effekten av dapagliflozin jämfört med SU är olika för personer som kunde respektive inte kunde varit med i den kliniska prövningen såtillvida att effekten bland de inkluderbara personerna var 2.47 [-3.83, -1.12] och bland de icke inkluderbara var -0.66 [-0.65, 0.33].

Tabell 7 av Skattning av skillnad i förändring av HbA1c, vikt, systoliskt blodtryck och eGFR mellan dapagliflozin och SU med 95% konfidensintervall efter 6 månaders uppföljning baserad på causal forest uppdelat efter inkluderbarhet i respektive klinisk prövning

Variabel	Population	Liraglutide vs SU	Dapagliflozin vs SU
HbA1c	All	-2.57 [-3.20, -1.94]	-1.10 [-1.92, -0.27]
HbA1c	Eligible: E	-3.19 [-4.31, -2.09]	-2.47 [-3.83, -1.12]
HbA1c	Non-eligible: I	-2.39 [-3.14, -1.64]	-0.66 [-0.65, 0.33]
Weight	All	-2.57 [-2.84, -2.30]	-2.98 [-3.42, -2.54]
Weight	Eligible: E	-2.91 [-3.46, -2.36]	-3.39 [-4.47, -2.31]
Weight	Non-eligible: I	-2.47 [-2.78, -2.15]	-2.86 [-3.33, -2.39]
SBP	All	-1.66 [-2.47, -0.78]	-3.87 [-5.08, -2.63]
SBP	Eligible: E	-2.14 [-4.21, -0.08]	-3.81 [-5.56, -2.05]
SBP	Non-eligible: I	-1.38 [-2.31, -0.46]	-3.82 [-5.31, -2.34]
eGFR	All	0.19 [-0.42, 0.81]	0.27 [-0.54, 1.09]
eGFR	Eligible: E	-0.56 [-1.88, 0.77]	-1.34 [-3.14, 0.46]
eGFR	Non-eligible: I	0.44 [-0.26, 1.14]	0.73 [-0.18, 1.65]

Analys av heterogenitet i betingade effekter med hjälp maskininlärning ger en möjlighet att helt icke parametriskt utvärdera vilka variabler som är associerade med den betingade kausala effekten.

Gradient boosting används för varje effektvariabel med följande resultat:

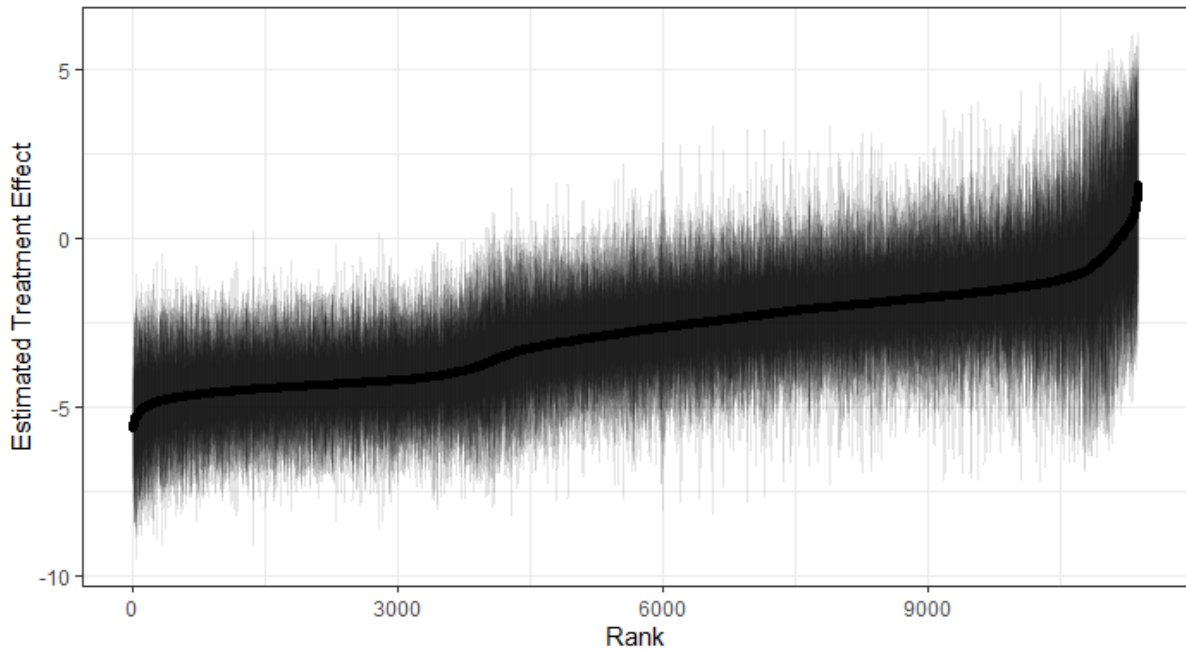
- Den skattade betingade effekten av liraglutide jämfört SU med avseende på förändring i HbA1c efter 6 månaders uppföljning är främst associerad med pre index HbA1c, kön, pre index vikt och pre index eGFR (figur 4).
- Den skattade betingade effekten av liraglutide jämfört SU med avseende på förändring i SBP efter 6 månaders uppföljning är främst associerad med pre index eGFR (figur 7).
- Den skattade betingade effekten av dapagliflozin jämfört SU med avseende på förändring i HbA1c efter 6 månaders uppföljning är främst associerad med pre index HbA1c, och pre indexålder.

Detta ger en vägledning om vilka variabler som det kan löna sig att undersöka vidare för att hitta grupper av personer där den betingade kausala effekten avviker. Effekten plottas som funktion av de identifierade variabler för att identifiera möjliga cutoffs för subgruppindelning.

Liraglutide vs SU, effekt på HbA1c

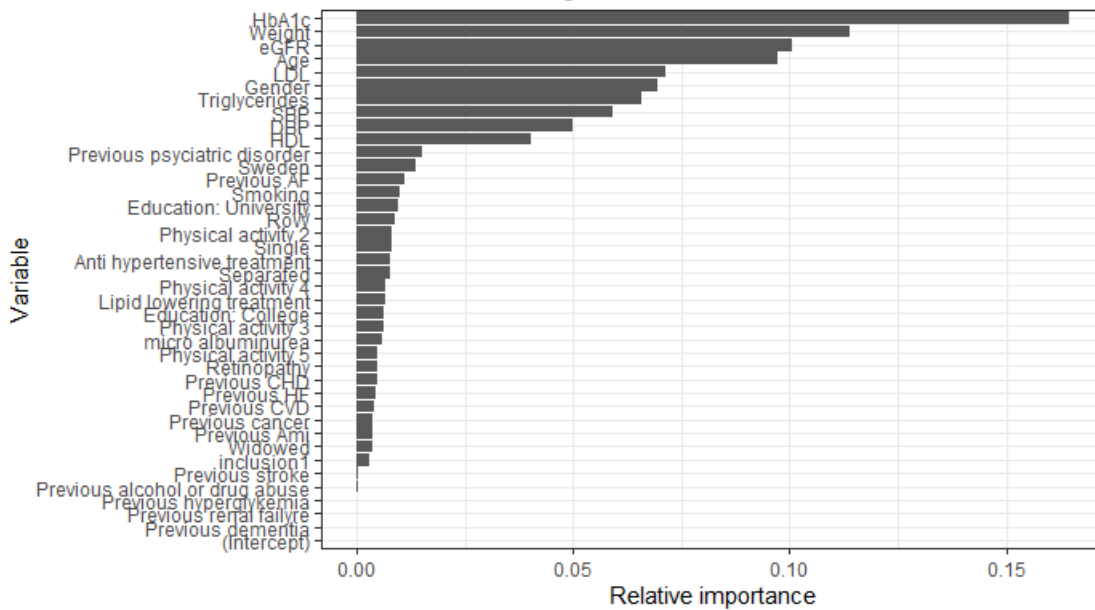
Heterogeniteten hos de skattade betingade kausala effekterna kan illustreras genom att plotta dessa i rangordning (figur 3) vilket bekräftar intrycket av att de betingade kausala effekterna förefaller variera mellan individer i relation till konfidensintervallen. HbA1c har högst relative importance (figur 4) som prediktor av betingad effekt följt av vikt, eGFR, ålder, LDL och kön.

Individual conditional expected treatment effects of liraglutide compared to SU on change in HbA1c (mmol/mol) with 95% Confidence intervals

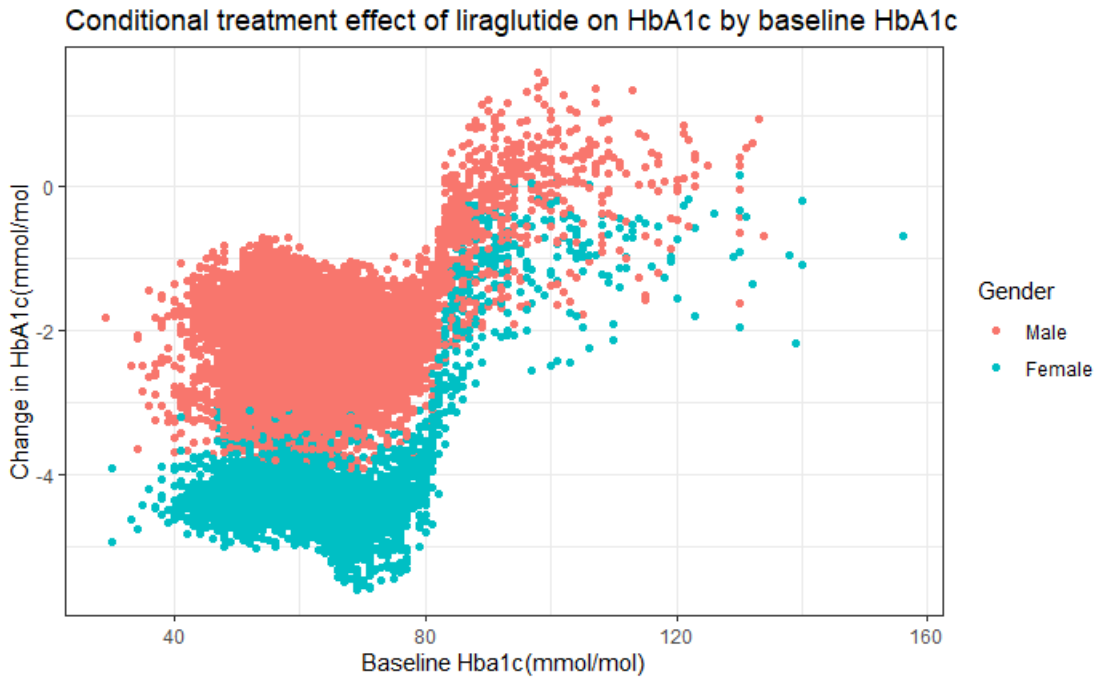


Figur 3 Rangordnade betingade kausala effekter av liraglutide jämfört med SU med avseende på förändring i HbA1c.

Relative importance as determinants of the average conditional causal effect of liraglutide on HbA1c



Figur 4 Relativ betydelse för heterogenitet i effekten av liraglutide jämfört med SU med avseende på förändring i HbA1c



Figur 5 Betingad effekt av liraglutid jämfört med SU med avseende på förändring i HbA1c gentemot pre index HbA1c uppdelat på män och kvinnor. Plot utan konfidensintervall.

Den betingade effekten av liraglutide jämfört med SU med avseende på förändring i HbA1c som funktion av pre index HbA1c och kön återges grafiskt i figur 5 som antyder att effekten är större bland kvinnor än bland män och att personer med högt HbA1c har mindre effekt, även om dessa är mycket få. Motsvarande samband mellan effekt och pre index vikt och eGFR finns figur 5 som inte ger något vidare stöd för att hitta några bra cutoff även om inflytande av pre index eGFR förefaller skilja något mellan män och kvinnor. Den skattade genomsnittliga effekten uppdelat på män och kvinnor samt på personer med pre under HbA1c under och över 85 finns i tabell 8.

Conditional treatment effect of liraglutide on HbA1c by baseline weight and eGFR

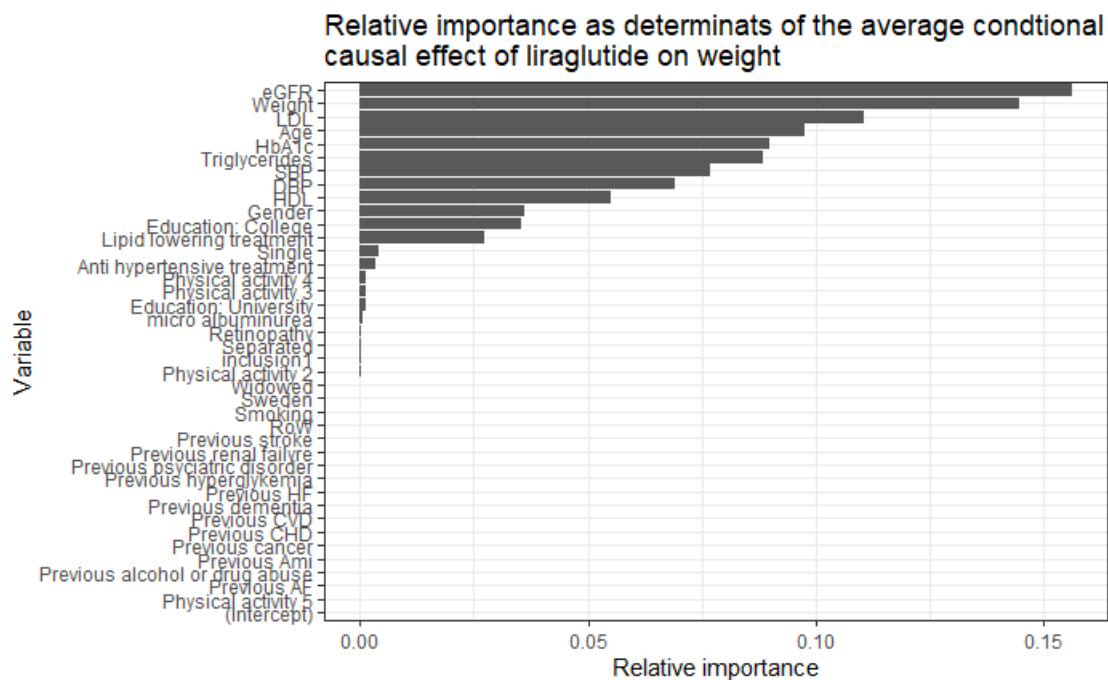


Figur 6 Betingad effekt av liraglutid jämfört med SU med avseende på förändring i HbA1c gentemot pre index vikt och eGFR uppdelat på män och kvinnor. Plot utan individuella konfidensintervall

Tabell 8 Betingad kausal effekt av liraglutide jämfört med SU bland män och kvinnor och personer med HbA1c under och över 85.

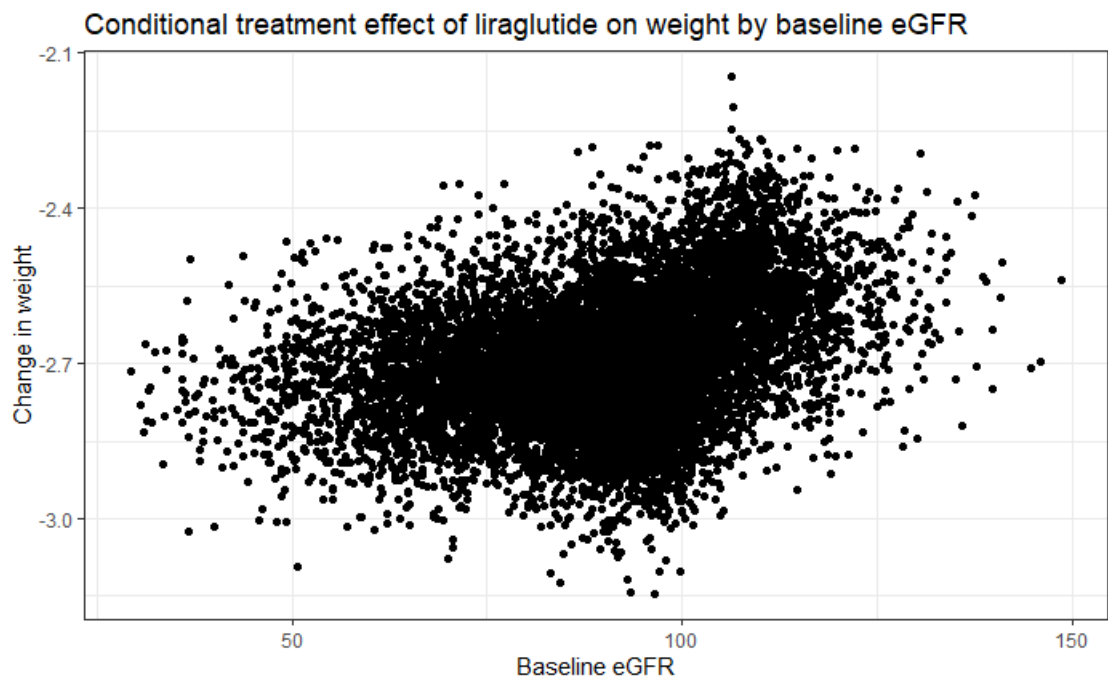
Jämförelse	Utfallsvariabel	Grupp	Skattad effekt
Liraglutide vs SU	HbA1c	HbA1c<85	-2.81[-3.43, -2.19]
Liraglutide vs SU	HbA1c	HbA1c ≥ 85	1.10 [-1.92, 4.11]
Liraglutide vs SU	HbA1c	Män	-1.15 [-2.00, -0.39]
Liraglutide vs SU	HbA1c	Kvinnor	-4.68 [-5.56, -3.79]

Liraglutide vs SU, effekt på vikt



Figur 7 Relativ betydelse för heterogenitet i effekten av liraglutide jämfört med SU med avseende på förändring i vikt

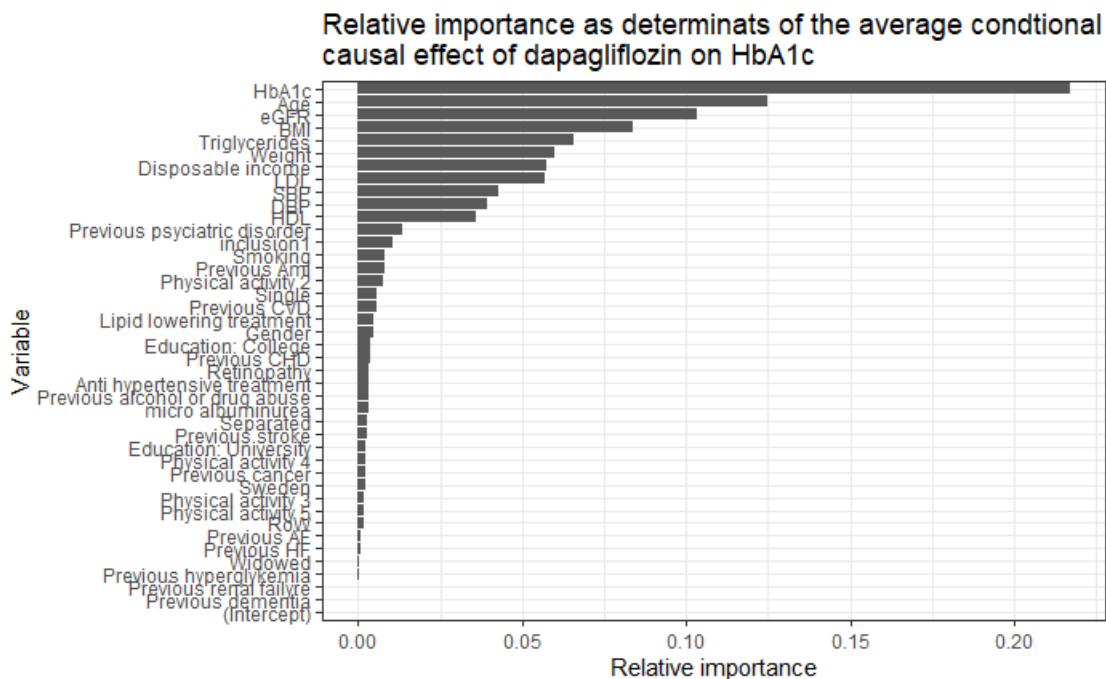
Även om analysen med gbm (figur 7) identifierar pre index eGRF som den viktigaste variabeln för heterogenitet i effekten av liraglutide på förändring av vikt så visar det sig i figur 8 att associationen mellan pre index eGRF och effekten är liten och att det inte lönar sig att leta efter någon subgrupp definierad utifrån eGRF.



Figur 8 Betingad effekt av liraglutid jämfört med SU med avseende på förändring i HbA1c gentemot pre index HbA1c. Plot utan konfidensintervall.

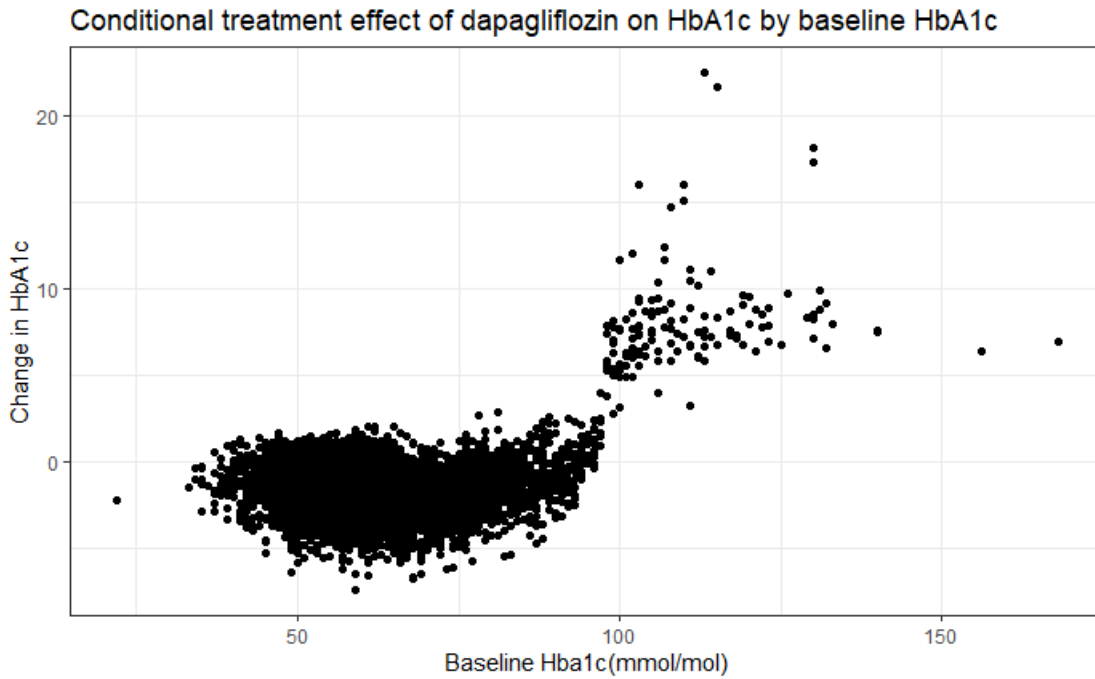
Dapagliflozin vs SU, effekt på HbA1c

Den betingade kausala effekten av dapagliflozin jämfört med SU med avseende på förändring av HbA1c är främst associerad med pre index HbA1bc följt av ålder och eGFR (figur 9).



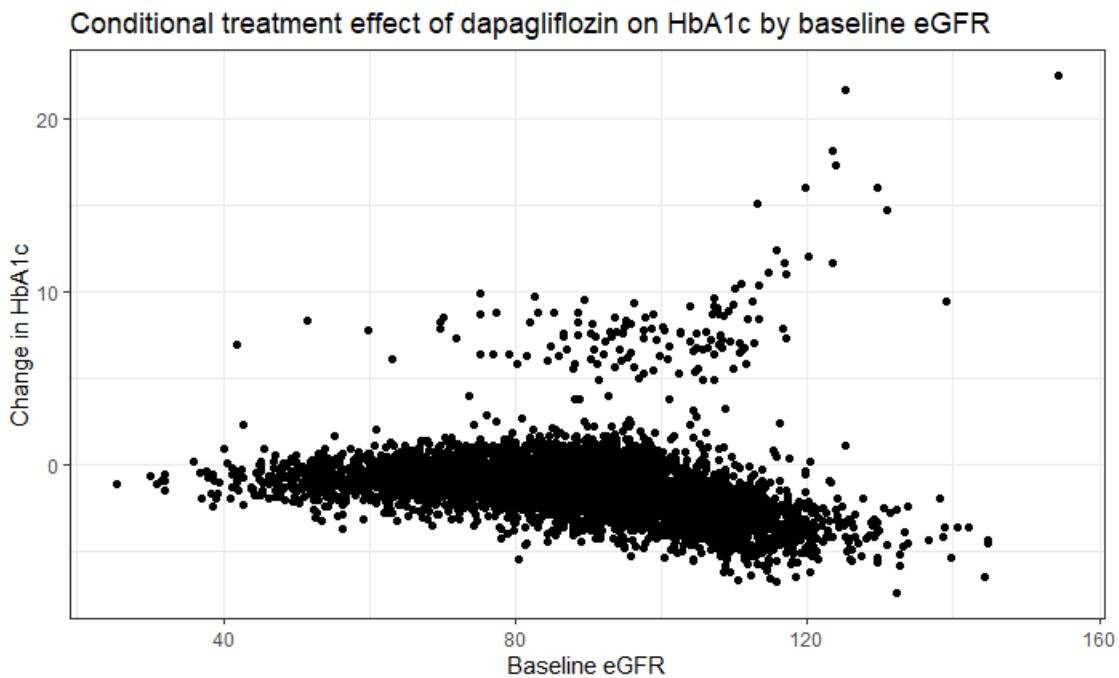
Figur 9 Relativt betydelse för heterogenitet i effekten av dapagliflozin jämfört med SU med avseende på förändring i HbA1c

Effekten av dapagliflozin jämfört med SU med avseende på förändring av HbA1c ser ut att vara relativt konstant till dapagliflozins fördel med en kraftig avvikelse för personer med mycket höga HbA1c där effekten reverseras (figur 10). Gruppen där detta sker är mycket liten och möjligen en artefakt.



Figur 10 Betingad effekt av dapagliflozin jämfört med SU med avseende på förändring i HbA1c gentemot pre index HbA1c. Plot utan konfidensintervall

Effekten av dapagloflozin jämfört med SU med avseende på förändring av HbA1c ser ut öka med ökade njurfunktion (Figur 11).



Figur 11 Betingad effekt av dapagliflozin jämfört med SU med avseende på förändring i HbA1c gentemot pre index eGFR. Plot utan konfidensintervall

Effekten av dapagliflozin jämfört med SU med avseende på förändring i HbA1c kan exempelvis delas upp i personer med låga eller höga pre index HbA1c eller eGFR (tabell 9)

Tabell 9 Skattad effekt av dapagliflozin jämfört med SU med avseende på förändring i HbA1c efter 6 månader uppdelat på HbA1c och eGFR.

Jämförelse	Utfallsvariabel	Grupp	Skattad effekt
dapagliflozin vs SU	HbA1c	HbA1c<90	-1.43 [-2.25, -0.61]
dapagliflozin vs SU	HbA1c	HbA1c ≥ 90	6.94 [2.37, 11.51]
dapagliflozin vs SU	HbA1c	eGFR<80	0.24 [-1.73, 2.20]
dapagliflozin vs SU	HbA1c	eGFR ≥ 80	-1.53 [-2.41, -0.64]

Diskussion och slutsats

De redovisade analyserna adresserar frågan om generaliserbarhet utanför den population som undersöks i en RCT samt den mer generella frågan om effekten är homogen eller heterogen i den population som exponerad för interventionen i klinisk praxis. Vi har använt två metoder där båda kan utvärdera generaliserbarheten men bara den ena kan undersöka graden av heterogenitet på ett mer detaljerat sätt.

Effekten av de två läkemedel som undersöks visar sig vara åtminstone delvis heterogen vilket innebär att alla personer i populationen som undersöks inte har samma effekt av läkemedlen men att det inte i någon större utsträckning hänger ihop med om de kunde varit med i respektive klinisk prövning, om de fått chansen, eller inte. Detta tyder på att resultaten från de prövningar vi refererar till här så vitt vi kan se kan generaliseras till alla personer som tar respektive läkemedel i praktiken även om dessa inte faller inom ramen för den prövning som låg till grund för läkemedlets godkännande. Huruvida detta resultat kan generaliseras till andra liknande prövningar av orala diabetesläkemedel eller utanför den population som ingår i våra analyser är naturligtvis oklart.

Heterogenitet i effekt identifieras i första hand för effekt med avseende på förändring av HbA1c och heterogeniteten kopplas i första hand till pre index HbA1c men även kön för liraglutid så till vida att effekten jämfört med SU förfaller vara mindre för de som har högst pre index HbA1c och för liraglutid även mindre för män än för kvinnor.

Både typerna av analyser bygger helt på registerdata och är bara möjliga för interventioner som används i tillräcklig omfattning och där man via registerdata kan fånga exponering, utfall och potentiella störfaktorer. Antagandet om att det inte får finnas ytterligare störfaktorer ställer stora krav på datatillgänglighet och även i våra analyser bör man fråga sig om vi verkligen har data på alla variabler som påverkar exponering, utfall och inkluderbarhet i en klinisk prövning.

Analyser som syftar till att undersöka heterogenitet i behandlingseffekt kan bli en viktig del i strävan att individualisera behandling och kan på sikt utgöra grund för val eller subvention av behandling, men antagandet om att det inte får finnas några oobserverade störfaktorer och kravet på tillförlitliga effekt data i relevanta tidsperioder gör att metoderna inte kan användas rutinmässigt utan kräver kvalificerade bedömningar av förutsättningar i varje enskilt fall.

Vi har visat att det är möjligt att utvärdera skillnader i behandlingseffekt av orala diabetesläkemedel baserat på registerdata och finner att effekten av både liraglutid och dapagliflozin är delvis heterogen men att denna heterogenitet inte hindrar resultaten från de två kliniska studierna NN2211-1575 och D1690C00004 att vara generaliserbart utanför den population som undersöks i respektive studie.

Referenser

Athey, Susan; Tibshirani, Julie; Wager, Stefan. Generalized random forests. *Ann. Statist.* 47 (2019), no. 2, 1148-1178.

Holland, P.W. 1986. Statistics and causal inference. *Journal of the American Statistical Association* 81, 945–70.

Pressler TR, Kaizar EE. The use of propensity scores and observational data to estimate randomized controlled trial generalizability bias. *Stat Med.* 2013;32(20):3552-3568.

Rubin, D.B. 1973a. Matching to remove bias in observational studies. *Biometrics* 29, 159–83. Correction note: 1974. *Biometrics* 30, 728.

Rubin, D.B. 1973b. The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics* 29, 185–203.

Rubin, D.B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66, 688–701.

Rubin, D.B. 1975. Bayesian inference for causality: the importance of randomization. *Proceedings of the Social Statistics Section of the American Statistical Association*, 233–9.

Rubin, D.B. 1976. Inference and missing data. *Biometrika* 63, 581–92.

Rubin, D.B. 1977. Assignment of treatment group on the basis of a covariate. *Journal of Educational Statistics* 2, 1–26.

Rubin, D.B. 1978. Bayesian inference for causal effects: the role of randomization. *Annals of Statistics* 6, 34–58.

Rubin, D.B. 1979. Discussion of 'Conditional independence in statistical theory' by A.P. Dawid. *Journal of the Royal Statistical Society, Series B* 41, 27–8.

Rubin, D.B. 1980. Discussion of 'Randomization analysis of experimental data in the Fisher randomization test'. *Journal of the American Statistical Association* 75, 591–3.

Appendix

Beskrivande statistik presenteras för data som ingår i respektive analys

Tabell 10 Beskrivande statistik redovisat som medelvärden med standardavvikelse för kontinuerliga variabler eller antal med procent för kategoriska variabler uppdelat på exponering inom de två analyspopulationerna för personer med observerad förändring i HbA1c efter 6 månader

Variable	Liraglutide vs SU		Dapagliflozin vs SU	
	Liraglutide	SU	Dapagliflozin	SU
n	9541	1826	5884	605
Age (mean (SD))	62.84 (10.38)	56.45 (10.21)	63.62 (10.48)	59.23 (9.98)
Female gender (%)	3682 (38.6)	757 (41.5)	2223 (37.8)	227 (37.5)
Diabetes duration (mean	7.42 (5.59)	6.48 (5.38)	8.18 (5.95)	7.33 (5.45)
BMI (mean (SD))	30.58 (5.12)	35.75 (5.95)	30.18 (5.14)	32.48 (5.54)
Weight (mean (SD))	90.22 (17.51)	107.41 (20.18)	89.13 (17.60)	97.70
Smoking (%)	1355 (16.3)	232 (14.4)	794 (15.4)	77 (14.4)
HbA1c (mean (SD))	62.85 (12.50)	65.15 (14.24)	62.46 (12.76)	66.06
SBP (mean (SD))	135.52 (14.83)	135.21 (14.92)	135.41	135.61
DBP (mean (SD))	78.56 (9.43)	80.83 (9.45)	78.46 (9.45)	81.07 (9.39)
LDL (mean (SD))	2.57 (0.90)	2.62 (0.91)	2.50 (0.91)	2.59 (0.93)
HDL (mean (SD))	1.18 (0.33)	1.13 (0.32)	1.19 (0.34)	1.16 (0.32)
Triglycerides (mean (SD))	2.07 (1.36)	2.40 (1.73)	2.08 (1.49)	2.12 (1.20)
Retinopathy (%)	1386 (21.1)	273 (20.9)	1032 (22.7)	105 (22.3)
Micro albuminurea (%)	1267 (18.2)	244 (19.5)	821 (19.1)	87 (20.4)
Macro albuminurea (%)	269 (4.0)	66 (5.3)	135 (3.4)	15 (3.7)
Anti hypertensive	6863 (74.5)	1320 (75.0)	4213 (74.1)	470 (80.1)
Creatinine (mean (SD))	71.92 (16.95)	70.24 (16.65)	71.89 (17.21)	69.69
Physical Activity (%)			88.69 (16.51)	93.61
< time per week	915 (12.3)	262 (18.4)		
1-2 times per week	1663 (22.3)	317 (22.2)	575 (12.0)	74 (15.2)
3-5 times per week	1785 (24.0)	330 (23.1)	1045 (21.8)	106 (21.8)
Daily	2060 (27.6)	269 (18.9)	1120 (23.4)	102 (20.9)
Never	1030 (13.8)	249 (17.4)	1351 (28.2)	115 (23.6)
Lipid lowering treatment	6134 (66.9)	1108 (63.7)	698 (14.6)	90 (18.5)
Income (mean (SD))	2164.40	2453.16	3837 (67.6)	394 (67.2)
Education (%)				
Basic	3247 (35.3)	362 (21.0)	1910 (34.3)	130 (23.6)
College	4185 (45.5)	976 (56.7)	2588 (46.4)	298 (54.0)
University	1758 (19.1)	384 (22.3)	1074 (19.3)	124 (22.5)
Marital status (%)				
Married	5303 (56.7)	886 (51.1)	3176 (55.9)	289 (51.9)
Separated	1627 (17.4)	333 (19.2)	1007 (17.7)	103 (18.5)
Single	1633 (17.5)	430 (24.8)	999 (17.6)	139 (25.0)
Widowed	783 (8.4)	85 (4.9)	501 (8.8)	26 (4.7)
Origin (%)				
Europe except Sweden	493 (5.2)	88 (4.8)	294 (5.0)	28 (4.6)

RoW	1774 (18.6)	205 (11.2)	1262 (21.4)	75 (12.4)
Sweden	7274 (76.2)	1533 (84.0)	4328 (73.6)	502 (83.0)
CHD (%)	1375 (14.4)	199 (10.9)	914 (15.5)	81 (13.4)
AMI (%)	609 (6.4)	95 (5.2)	397 (6.7)	41 (6.8)
Stroke (%)	361 (3.8)	48 (2.6)	238 (4.0)	26 (4.3)
CVD (%)	934 (9.8)	142 (7.8)	609 (10.4)	65 (10.7)
HF (%)	377 (4.0)	76 (4.2)	238 (4.0)	16 (2.6)
AF (%)	608 (6.4)	119 (6.5)	426 (7.2)	34 (5.6)
Renal dysfunction (%)	76 (0.8)	17 (0.9)	53 (0.9)	4 (0.7)
Hyper glycemias (%)	73 (0.8)	24 (1.3)	50 (0.8)	4 (0.7)
Psychiatric disorder (%)	466 (4.9)	125 (6.8)	314 (5.3)	51 (8.4)
Alcohol or drug abuse (%)	156 (1.6)	52 (2.8)	108 (1.8)	14 (2.3)
Cancer (%)	957 (10.0)	151 (8.3)	695 (11.8)	50 (8.3)
Dementia (%)	16 (0.2)	1 (0.1)	21 (0.4)	1 (0.2)

Tabell 11 Beskrivande statistik redovisat som medelvärden med standardavvikelse för kontinuerliga variabler eller antal med procent för kategoriska variabler uppdelat på exponering inom de två analyspopulationerna för personer med observerad förändring i vikt efter 6 månader

Variable	Liraglutide vs SU		Dapagliflozin vs SU	
	Liraglutide	SU	Dapagliflozin	SU
n	8385	1659	5077	524
Age (mean (SD))	62.88 (10.33)	56.56 (10.19)	63.76 (10.40)	59.14 (9.67)
Female gender (%)	3230 (38.5)	684 (41.2)	1909 (37.6)	196 (37.4)
Diabetes duration (mean	7.41 (5.54)	6.54 (5.41)	8.25 (5.94)	7.37 (5.57)
BMI (mean (SD))	30.60 (5.12)	35.78 (5.97)	30.19 (5.14)	32.44 (5.60)
Weight (mean (SD))	90.34 (17.52)	107.50 (20.23)	89.18 (17.62)	97.76
Smoking (%)	1205 (15.8)	217 (14.5)	699 (15.2)	63 (13.2)
HbA1c (mean (SD))	62.70 (12.29)	65.10 (14.28)	62.30 (12.51)	66.07
SBP (mean (SD))	135.46 (14.88)	135.20 (14.78)	135.38	135.39
DBP (mean (SD))	78.53 (9.45)	80.70 (9.45)	78.38 (9.43)	81.18 (9.42)
LDL (mean (SD))	2.56 (0.90)	2.62 (0.91)	2.50 (0.91)	2.57 (0.93)
HDL (mean (SD))	1.19 (0.33)	1.13 (0.32)	1.19 (0.34)	1.16 (0.32)
Triglycerides (mean (SD))	2.06 (1.35)	2.41 (1.78)	2.07 (1.51)	2.16 (1.16)
Retinopathy (%)	1243 (20.8)	251 (20.6)	901 (22.2)	98 (23.5)
Micro albuminuria (%)	1126 (17.9)	226 (19.3)	719 (18.9)	74 (19.5)
Macro albuminuria (%)	244 (3.9)	61 (5.2)	119 (3.4)	11 (3.0)
Anti hypertensive	6091 (74.9)	1211 (75.5)	3656 (74.3)	410 (80.4)
Creatinine (mean (SD))	72.01 (16.82)	70.30 (16.64)	71.93 (16.98)	69.58
Physical Activity (%)			88.53 (16.35)	93.73
< time per week	869 (12.6)	245 (18.2)		
1-2 times per week	1550 (22.5)	298 (22.2)	526 (12.2)	60 (13.7)
3-5 times per week	1648 (23.9)	311 (23.1)	940 (21.8)	92 (21.0)
Daily	1910 (27.7)	252 (18.7)	1008 (23.4)	99 (22.6)
Never	923 (13.4)	239 (17.8)	1238 (28.7)	106 (24.1)
Lipid lowering treatment	5430 (66.9)	1010 (63.6)	603 (14.0)	82 (18.7)

Income (mean (SD))	2163.25	2415.43	3335 (67.8)	349 (68.3)
Education (%)				
Basic	2908 (35.9)	338 (21.5)	1666 (34.5)	113 (23.4)
College	3686 (45.5)	889 (56.5)	2246 (46.5)	256 (53.1)
University	1516 (18.7)	347 (22.0)	923 (19.1)	113 (23.4)
Marital status (%)				
Married	4696 (57.0)	809 (51.1)	2786 (56.5)	258 (53.0)
Separated	1423 (17.3)	307 (19.4)	870 (17.7)	92 (18.9)
Single	1458 (17.7)	387 (24.4)	857 (17.4)	117 (24.0)
Widowed	665 (8.1)	80 (5.1)	415 (8.4)	20 (4.1)
Origin (%)				
Europe except Sweden	435 (5.2)	79 (4.8)	257 (5.1)	19 (3.6)
RoW	1476 (17.6)	190 (11.5)	1058 (20.8)	66 (12.6)
Sweden	6474 (77.2)	1390 (83.8)	3762 (74.1)	439 (83.8)
CHD (%)	1223 (14.6)	188 (11.3)	798 (15.7)	69 (13.2)
AMI (%)	549 (6.5)	88 (5.3)	347 (6.8)	34 (6.5)
Stroke (%)	308 (3.7)	42 (2.5)	207 (4.1)	24 (4.6)
CVD (%)	819 (9.8)	129 (7.8)	529 (10.4)	56 (10.7)
HF (%)	339 (4.0)	71 (4.3)	219 (4.3)	12 (2.3)
AF (%)	530 (6.3)	100 (6.0)	372 (7.3)	30 (5.7)
Renal dysfunction (%)	71 (0.8)	15 (0.9)	48 (0.9)	5 (1.0)
Hyper glycemias (%)	66 (0.8)	19 (1.1)	45 (0.9)	3 (0.6)
Psychiatric disorder (%)	410 (4.9)	116 (7.0)	274 (5.4)	42 (8.0)
Alcohol or drug abuse (%)	131 (1.6)	46 (2.8)	89 (1.8)	12 (2.3)
Cancer (%)	866 (10.3)	141 (8.5)	608 (12.0)	45 (8.6)
Dementia (%)	15 (0.2)	2 (0.1)	18 (0.4)	0 (0.0)

Tabell 12 Beskrivande statistik redovisat som medelvärden med standardavvikelse för kontinuerliga variabler eller antal med procent för kategoriska variabler uppdelat på exponering inom de två analyspopulationerna för personer med observerad förändring i SBP efter 6 månader

Variable	Liraglutide vs SU		Dapagliflozin vs SU	
	Liraglutide	SU	Dapagliflozin	SU
n	9077	1726	5632	572
Age (mean (SD))	62.90 (10.34)	56.57 (10.24)	63.76 (10.47)	59.09 (9.98)
Female gender (%)	3515 (38.7)	711 (41.2)	2138 (38.0)	219 (38.3)
Diabetes duration (mean	7.44 (5.56)	6.54 (5.36)	8.27 (5.93)	7.35 (5.48)
BMI (mean (SD))	30.58 (5.12)	35.70 (5.92)	30.17 (5.14)	32.48 (5.57)
Weight (mean (SD))	90.22 (17.51)	107.31 (20.17)	89.12 (17.57)	97.67
Smoking (%)	1284 (15.9)	219 (14.1)	766 (15.3)	72 (14.1)
HbA1c (mean (SD))	62.60 (12.23)	65.03 (14.06)	62.23 (12.54)	65.82
SBP (mean (SD))	135.57 (14.90)	135.17 (14.88)	135.51	135.54
DBP (mean (SD))	78.56 (9.44)	80.73 (9.49)	78.46 (9.43)	81.03 (9.55)
LDL (mean (SD))	2.56 (0.90)	2.61 (0.91)	2.49 (0.91)	2.59 (0.93)
HDL (mean (SD))	1.18 (0.33)	1.13 (0.32)	1.19 (0.34)	1.16 (0.32)
Triglycerides (mean (SD))	2.07 (1.35)	2.40 (1.75)	2.07 (1.49)	2.18 (1.32)
Retinopathy (%)	1323 (20.7)	255 (20.4)	990 (22.4)	105 (23.4)
Micro albuminuria (%)	1214 (18.0)	234 (19.4)	794 (19.1)	83 (20.4)
Macro albuminuria (%)	264 (4.0)	58 (4.9)	132 (3.4)	15 (3.9)

Anti hypertensive	6601 (74.9)	1264 (75.6)	4076 (74.5)	448 (80.4)
Creatinine (mean (SD))	71.94 (16.97)	70.40 (16.67)	71.93 (17.19)	69.55
Physical Activity (%)			88.52 (16.50)	93.88
< time per week	898 (12.3)	250 (18.2)		
1-2 times per week	1632 (22.4)	311 (22.6)	559 (12.0)	69 (14.7)
3-5 times per week	1732 (23.8)	315 (22.9)	1025 (22.0)	98 (20.9)
Daily	2018 (27.7)	262 (19.1)	1084 (23.2)	101 (21.6)
Never	1001 (13.7)	237 (17.2)	1315 (28.2)	110 (23.5)
Lipid lowering treatment	5887 (67.0)	1053 (63.6)	681 (14.6)	90 (19.2)
Income (mean (SD))	2160.59	2446.92	3705 (67.9)	370 (66.5)
Education (%)				
Basic	3104 (35.5)	349 (21.5)	1818 (34.0)	128 (24.2)
College	3998 (45.7)	915 (56.2)	2502 (46.8)	280 (53.0)
University	1649 (18.8)	363 (22.3)	1022 (19.1)	120 (22.7)
Marital status (%)				
Married	5074 (57.0)	839 (51.2)	3059 (56.2)	281 (52.8)
Separated	1525 (17.1)	314 (19.2)	957 (17.6)	94 (17.7)
Single	1555 (17.5)	401 (24.5)	944 (17.3)	132 (24.8)
Widowed	741 (8.3)	85 (5.2)	483 (8.9)	25 (4.7)
Origin (%)				
Europe except Sweden	472 (5.2)	85 (4.9)	288 (5.1)	27 (4.7)
RoW	1685 (18.6)	196 (11.4)	1213 (21.5)	73 (12.8)
Sweden	6920 (76.2)	1445 (83.7)	4131 (73.3)	472 (82.5)
CHD (%)	1324 (14.6)	185 (10.7)	891 (15.8)	80 (14.0)
AMI (%)	594 (6.5)	88 (5.1)	388 (6.9)	41 (7.2)
Stroke (%)	346 (3.8)	42 (2.4)	233 (4.1)	25 (4.4)
CVD (%)	899 (9.9)	129 (7.5)	592 (10.5)	64 (11.2)
HF (%)	370 (4.1)	72 (4.2)	241 (4.3)	14 (2.4)
AF (%)	581 (6.4)	110 (6.4)	411 (7.3)	31 (5.4)
Renal dysfunction (%)	71 (0.8)	17 (1.0)	49 (0.9)	6 (1.0)
Hyper glycemias (%)	67 (0.7)	21 (1.2)	48 (0.9)	3 (0.5)
Psychiatric disorder (%)	444 (4.9)	118 (6.8)	303 (5.4)	46 (8.0)
Alcohol or drug abuse (%)	155 (1.7)	51 (3.0)	108 (1.9)	14 (2.4)
Cancer (%)	919 (10.1)	146 (8.5)	668 (11.9)	50 (8.7)
Dementia (%)	14 (0.2)	2 (0.1)	20 (0.4)	1 (0.2)

Tabell 13 Beskrivande statistik redovisat som medelvärden med standardavvikelse för kontinuerliga variabler eller antal med procent för kategoriska variabler uppdelat på exponering inom de två analyspopulationerna för personer med observerad förändring i eGFR efter 6 månader

Variable	Liraglutide vs SU		Dapagliflozin vs SU	
	Liraglutide	SU	Dapagliflozin	SU
n	7762	1472	4678	503
Age (mean (SD))	63.21 (10.35)	56.55 (10.25)	64.03 (10.44)	59.17
Female gender (%)	3020 (38.9)	604 (41.0)	1786 (38.2)	184 (36.6)
Diabetes duration (mean	7.57 (5.70)	6.55 (5.46)	8.34 (6.04)	7.34 (5.49)

BMI (mean (SD))	30.57 (5.10)	35.77 (6.01)	30.14 (5.10)	32.52 (5.53)
Weight (mean (SD))	90.15 (17.47)	107.65 (20.56)	88.96 (17.39)	98.03
Smoking (%)	1107 (16.2)	191 (14.6)	625 (15.1)	61 (13.5)
HbA1c (mean (SD))	62.65 (12.36)	64.73 (14.32)	62.15 (12.65)	66.16
SBP (mean (SD))	135.62 (15.04)	135.20 (14.52)	135.42	135.41
DBP (mean (SD))	78.52 (9.50)	80.79 (9.48)	78.27 (9.54)	80.99 (9.28)
LDL (mean (SD))	2.56 (0.90)	2.62 (0.91)	2.49 (0.91)	2.60 (0.93)
HDL (mean (SD))	1.19 (0.33)	1.13 (0.32)	1.20 (0.35)	1.16 (0.32)
Triglycerides (mean (SD))	2.05 (1.32)	2.43 (1.78)	2.05 (1.46)	2.10 (1.15)
Retinopathy (%)	1155 (21.2)	217 (20.3)	832 (22.5)	88 (22.7)
Micro albuminurea (%)	1098 (18.1)	206 (19.2)	713 (19.5)	77 (20.5)
Macro albuminurea (%)	239 (4.0)	51 (4.8)	119 (3.5)	14 (3.9)
Anti hypertensive	5689 (75.2)	1078 (75.3)	3407 (74.7)	394 (80.4)
Creatinine (mean (SD))	72.09 (17.04)	70.43 (16.58)	72.01 (17.23)	69.76
Physical Activity (%)			88.43 (16.55)	93.66
< time per week	755 (12.4)	213 (18.4)		
1-2 times per week	1379 (22.6)	258 (22.3)	475 (12.3)	64 (15.5)
3-5 times per week	1428 (23.4)	261 (22.6)	849 (22.0)	91 (22.0)
Daily	1694 (27.7)	214 (18.5)	884 (22.9)	88 (21.3)
Never	851 (13.9)	209 (18.1)	1098 (28.4)	96 (23.2)
Lipid lowering treatment	5086 (67.2)	909 (63.8)	558 (14.4)	74 (17.9)
Income (mean (SD))	2163.44	2349.80	3116 (68.2)	329 (67.0)
Education (%)				
Basic	2636 (35.2)	299 (21.4)	1503 (33.8)	111 (23.9)
College	3418 (45.6)	788 (56.5)	2096 (47.1)	246 (52.9)
University	1442 (19.2)	308 (22.1)	854 (19.2)	108 (23.2)
Marital status (%)				
Married	4357 (57.1)	720 (51.2)	2568 (56.5)	246 (52.5)
Separated	1307 (17.1)	270 (19.2)	785 (17.3)	87 (18.6)
Single	1330 (17.4)	349 (24.8)	790 (17.4)	114 (24.3)
Widowed	636 (8.3)	66 (4.7)	401 (8.8)	22 (4.7)
Origin (%)				
Europe except Sweden	399 (5.1)	66 (4.5)	239 (5.1)	21 (4.2)
RoW	1432 (18.4)	171 (11.6)	987 (21.1)	58 (11.5)
Sweden	5931 (76.4)	1235 (83.9)	3452 (73.8)	424 (84.3)
CHD (%)	1183 (15.2)	169 (11.5)	760 (16.2)	69 (13.7)
AMI (%)	510 (6.6)	83 (5.6)	317 (6.8)	34 (6.8)
Stroke (%)	307 (4.0)	39 (2.6)	196 (4.2)	23 (4.6)
CVD (%)	781 (10.1)	121 (8.2)	488 (10.4)	57 (11.3)
HF (%)	322 (4.1)	67 (4.6)	195 (4.2)	13 (2.6)
AF (%)	520 (6.7)	95 (6.5)	347 (7.4)	28 (5.6)
Renal dysfunction (%)	67 (0.9)	13 (0.9)	46 (1.0)	5 (1.0)
Hyper glycemias (%)	62 (0.8)	19 (1.3)	42 (0.9)	2 (0.4)
Psychiatric disorder (%)	373 (4.8)	107 (7.3)	245 (5.2)	44 (8.7)
Alcohol or drug abuse (%)	132 (1.7)	44 (3.0)	84 (1.8)	11 (2.2)
Cancer (%)	823 (10.6)	117 (7.9)	577 (12.3)	43 (8.5)
Dementia (%)	14 (0.2)	1 (0.1)	19 (0.4)	1 (0.2)