

Enzalutamide mot abiraterone, en design för en komparativ effektjämförelse i klinisk praxis.

SAMMANFATTNING:

Vi utvärderar enzalutamide (EZ) mot abiraterone (AA) med hjälp av våra nationella populations- och kvalitetsregister. Utvärderingar som är baserade på observationsdata kritiserar ofta för deras brist på objektivitet. Vi vill visa att observationsstudier kan genomföras objektivt som vid en randomiserad studie. Objektiviteten beror på att vi, innan vi har utfallsdata, bestämmer upplägget för vår utvärdering, dvs en design.

I vårt fall inbegrips i designen: (1) att definiera *vem* som ingår i studien innan analys; (2) att specificerar *vad* vi ska göra, dvs vilka analyser som ska göras vilket inbegriper de utfall vi avser att analysera och de undergrupper vi avser är viktiga att studera; och (3) att precisera *hur* vi kan utvärdera känsliga antaganden för vår analys, vilket i vårt fall innebär att specificera känslighetsanalyser för potentiell utelämnade variabler, s.k. confounding. Vi har sammanfattat vår design i bifogat dokument.

Vi har än så länge valt att enbart arbeta med data från våra nationella register. Under hösten kommer vi även att beställa data ifrån Nationella Prostata Cancer Registret (NPCR). Vi kommer att använda NPCR för att avgöra om de data vi har använt från Nationella registren räcker för att designa vår studie. Dessutom använder vi NPCR för att bättre kunna studera fler effektskillnader mellan preparaten. NPCR torde fånga sjukdomsprogressionen bättre än de nationella registren.

HUVUDRAPPORT¹

INTRODUKTION

Dessa preparat har aldrig utvärderats mot varandra i en kontrollerad randomiserad studie (RCT). En RCT är kostsam och dessutom skulle det ta lång tid till det att analyserna vad gäller, primärt, mortalitet kan genomföras. Det är också svårt att göra en indirekt jämförelse av RCTer, eftersom EZ har utvärderats mot placebo, medan AA har utvärderats mot prednisone. Dessutom är patienterna inte helt jämförbara i RCTerna. När AA och EZ utvärderades före kemoterapi i COU-AA-302 och PREVAIL skiljde sig patienterna åt t.ex. beträffande andelen som har benmetastaser, 51% mot 85%. Oavsett detta, så är resultat från väldesignade observationsstudier av intresse. Det primära skälet till detta är att resultatet från RCT ofta är svåra att generalisera till klinisk praxis, detta eftersom patienter i RCT i regel är förhållandevis friska. Vad gäller dessa läkemedel har patienterna i COU-AA-302 och PREVAIL bättre allmäntillstånd än patienter i klinisk praxis.² De observationsstudier som har gjorts har visserligen jämfört kliniska (3 studier), och patientrapporterade (5 studier) utfall, men det är svårt att avgöra om grupperna var jämförbara vid dessa utvärderingen. Få eller inga observerade variabler, s.k. kovariter, används i analys och oftast finns inte heller en beskrivning av kovariater uppdelad på de två grupperna.

FRÅGESTÄLLNINGAR

Vi utvärderar två alternativa behandlingar för patienter med mCRPC, abiraterone acetate plus prednisone (AA) mot enzalutamide (EZ) i klinisk praxis. Vi vill utnyttja data från både populations- och kvalitetsregister, en unik källa till kunskap om de används på korrekt sätt. Till att börja med har vi enbart utnyttjat våra populationsregister.

METOD

Utvärderingsproblemet består i att man som individ inte kan vara i två tillstånd samtidigt. I detta fall innebär detta att en patient får antingen AA eller EZ. För den som får AA vet vi deras utfall under AA behandling men inte vad deras utfall skulle ha varit om de i stället givits EZ, och vise versa.

Innebörden av detta är de kontrafaktiska utfallen för de som får AA och EZ måste skattas. Detta görs i ett randomiserat experiment (RCT) genom att slumpmässigt allokera patienter till, i det här fallet, AA och EZ.

Resultat i form av skillnader i medelvärden mellan de som fick AA och de som fick EZ från ett idealt (t.ex. dubbelblindat) RCT kan sägas ha hög intern validitet. I och med att man känner tilldelningen till behandling (d.v.s. randomisering) är det lätt att testa (t.ex. via ett vanligt t-test) om det finns

¹ Denna rapport bygger på Johansson, Jonéus and Langenskiöld (2020).

² Alla patienter i kliniska studier har ECOG nivåerna 0 och 1 mot 85% i klinisk praxis.

skillnader i effekter mellan de två behandlingarna i själva experimentet. Om inte experiment population är dragen slumpmässigt från population av cancersjuka män är det svårt att uttala sig om det finns en effekt i denna population.

Vår metod tar sin utgångspunkt i det randomiserade experimentet. Innebörden är att vi fokuserar på att modellerar tilldelningen av behandlingen. Strategin skapar det slumpmässiga experimentet i efterhand genom att erhålla balans mellan kovariaterna hos de som erhållit AA och EZ i våra data, före en analys av effekten genomförs, och kallas av den anledningen en kvasi-experimentell design av statistiker och epidemiologer. En tydlig nackdel, och i motsats till det ideala slumpmässiga experimentet, är att vi inte är säkra på att ha balans bland icke observerade kovariater som är betydelsefulla för att förklara skillnader i utfall. Sådana variabler benämns confounders i litteraturen. Av denna anledning har vi föreslagit två känslighetsanalyser för att studera hur känsliga vår design är för "confounding". De slutgiltiga analyserna görs med regressionsanalys och Neyman-Perarson inferens (d.v.s. t-test).

I om med att det metodologiska fokuserat är på tilldelning av de två medicinerna började vi med ett kvalitativt metod arbete för att få en bättre förståelse av denna. Analys av de kvalitativa data visar att läkare inte förskriver dessa alternativ slumpmässigt. Många läkare väljer bort AA för patienter med osteoporos, diabetes, och hjärtsjukdomar, men väljer läkemedlet för dem med trötthet och dåligt allmäntillstånd. Analyserna visar också att inte alla läkare förskriver likadant. Vi fann bland annat att heterogeniteten beror på att läkare tolkar evidensen olika, och att läkare påverkas av de riktlinjer som gällde för landstingen och som varierade över landet, riktlinjer som läkarna tror speglar priserna som landstingen har förhandlat fram med företagen. Förskrivningen varierar betydligt mer över landstingen än vad som kan förklaras med olikheter i hälsostatus i populationerna. Därmed finns mycket goda förutsättningar att utvärdera AA mot EZ, dvs att hitta de patienter som får AA hos en läkare men skulle ha fått EZ hos en annan.

DATA

Datakällor

Vi använder register som tillhandahålls av Socialstyrelsen. Patientregistret (PAR) registrerar information kring besök i den öppna och slutna vården. Läkemedelsregistret innehåller information om läkemedelsanvändningen, dvs sådana läkemedel som förskrivs på recept. Cancerregistret innehåller information om alla cancerdiagnoser som ges.

Vi använder också information från SCB. Longitudinell integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier, LISA, från 1990-2015 integrerar befintliga data från arbetsmarknads-,

utbildnings- och socialförsäkringar . Folk- och Bostadsräkning som genomfördes vart femte år mellan 1960-1990, samlar in data om demografi, sysselsättning och boende.

Vi upplevde inga problem att få tag i vårt datamaterial. Både Socialstyrelsen och SCB var beredda att leverera snabbt (4 månader) när leveransen gällde ett regeringsuppdrag. Däremot upplevde vi till att börja med att vi inte fick ett så detaljerat datamaterial som vi önskade. Tack vara vårt kvalitativa arbete och vår systematiska litteraturgenomgång kunde vi däremot motivera vilka variabler vi behövde.

Vår stora utmaning, däremot, var att vi inte engagera en urolog eller onkolog till vårt projekt. Detta gjorde det extra svårt att definiera kovariaterna och utfallen. För detta arbete, hade vi återigen stor användning av det kvalitativa arbete och litteraturgenomgången. Faktum är att vi inte hade klarat uppdraget utan detta.

Kovariater

I och med registren från Socialstyrelsen och SCB kan vi skapa 1000-tals potentiella kovariater relaterade till historisk hälsa och socioekonomisk status (SES). Vi valde att definiera 162 variabler som används för att skapa vår design. Av dessa består mer än 50% av kovariaterna SES. De hälsorelaterade kovariaterna består av sjukhusbesök och tidigare diagnoser, både före och efter diagnosen men samtliga före behandlingen.

Population

Från Cancerregistret finner vi 243 535 unika patienter med prostatacancer. Diagnosåret sträcker sig mellan 1986 och 2016. Läkemedlen subventionerades i juni (EZ) och juli (AA) 2015. Innan dess var det i princip ingen som fick dessa behandlingar. Population begränsas därför till alla män med recept på AA eller EZ under perioden 1 juni 2015 till 15 juni 2018 (n = 4 601).

Design

Utifrån vårt kvalitativa arbete gjordes bedömning om vilka variabler som var de mest viktiga att ta hänsyn till i vår design. Medelvärde av dessa variabler för de båda grupperna presenteras i tabell 1, liksom resultatet av det test vi använt oss av för att avgöra om det finns statistiska säkerställda skillnader mellan de två grupperna bland dessa kovariater. Det som är mest slående är att de två populationerna är så lika vad gäller dessa kovariater. Det är få kovariater som skiljer sig åt över två grupperna, och skillnaderna är så små som man ofta finner efter det att man genomfört ett RCT. Med en risknivå om 10% eller mindre kan vi t.ex. se att väntetiden till behandling för EZ är ca 0.34 år kortare än för AA. Vidare kan vi se en tre procentenheter lägre förekomst av en akut hjärtinfarkt (*Acute myocardial infarction before treatment*), en fyra procentenheters högre förekomst av diabetes, och en fem procentenheters lägre förekomst av malign cancer (Secondary malignant

neoplasms) för EZ än för AA. Utifrån denna beskrivning så är det inte uppenbart att man kan tro att en grupp i genomsnitt har sämre hälsa än den andra.

Description	enzalutamide	abiraterone	Diff
Age at treatment	75.27 (7.85)	75.29 (7.7)	-0.02
Years to treatment from diagnosis	6.95 (5.00)	7.29 (5.31)	-0.34*
Less than secondary school education	0.36 (0.48)	0.35 (0.48)	0.01
Secondary school education	0.39 (0.49)	0.39 (0.49)	-0.00
More than secondary school education	0.25 (0.43)	0.26 (0.44)	-0.01
Living with a partner	0.66 (0.47)	0.67 (0.47)	-0.01
Cardiovascular disease before treatment	0.35 (0.48)	0.34 (0.47)	0.01
<i>Atrial fibrillation and flutter before treatment (I48)</i>	0.16 (0.36)	0.18 (0.38)	-0.02
<i>Acute myocardial infarction before treatment (I21)</i>	0.10 (0.30)	0.07 (0.25)	0.03***
<i>Other cardiovascular diseases before treatment</i>	0.27 (0.45)	0.26 (0.44)	0.01
Diabetes before treatment	0.16 (0.37)	0.12 (0.32)	0.04***
Osteoporosis before treatment	0.01 (0.11)	0.01 (0.10)	0.00
Secondary malignant neoplasms	0.70 (0.46)	0.75 (0.43)	-0.05***
Malaise and fatigue	0.05 (0.22)	0.04 (0.20)	0.01

*Standard deviations within parentheses. *p < 0.1; **p < 0.05; ***p < 0.01.*

Tabell 1: Skillnad i kritiska variabler.

Vi antar att sannolikheten att få AA kan beskrivas som en funktion av alla potentiella kovariater (inklusive interaktioner och polynom för de kontinuerliga. Denna sannolikhet benämns Propensity Score (PS) i litteraturen, och vid både analys och design bör denna skattas. Vi vet från tidigare forskning att det inte är bra att ha med för många kovariater vid skattning, se tex de Luna m.fl. (2010). Det viktiga att ha med de kovariater som avgör behandling av AA och EZ och även kan påverka utfall, sk. confounders. Är inte de grupper som jämförs lika när det gäller dessa kovariater, så vet vi inte om eventuella skillnader beror på preparaten eller på andra skillnader över grupperna. Utmaningen är att identifiera dessa kovariater. Tack vara vårt kvalitativa arbete har vi dock viss kunskap. Vi väljer även att använda en maskininlärningsalgoritm för att plocka ut de mest betydelsefulla. Om den skattade modellen är en korrekt beskrivning för att en man får AA istället för EZ räcker det med att kontrollera för denna skattade PS, dvs se till att de grupper som jämförs är lika med avseende på PS.

Eftersom vi inte kan veta om så är fallet väljer vi att inte bara använda denna skattade PS för att hitta jämförbara män utan att också använda en delmängd av kovariater som vi bedömer som extra viktiga. Sedan använder vi en genetisk matchningsalgoritm för att välja ut män som är jämförbara i de två grupperna med avseende på dessa kovariater och PS. Algoritmen kan beskrivas som att för en man som förskrivs AA hitta en jämförbar man bland dem som förskrivs EZ så att skillnaden i dessa kovariater och PS, mätt som det kvadratiska avståndet,³ är minimalt. Detta upprepas för alla män som förskrivits AA och sedan på samma sätt för de som förskrivs EZ.

³ Mer precist så används ett viktat avstånd, Mahalanobis distance metrix.

Vi anser att sådana variabler som läkarna har beskrivit som påverkar deras förskrivning är extra viktiga. Resten av de 130 variablerna sammanfattar vi med 9 faktorer, Dessa faktorer skapas vi faktoranalys. De nio faktorerna förklarar 42 procent av variansen bland de inkluderade kovariaterna⁴. För en beskrivning av de nio faktorerna se tabell 2. En fullständig beskrivning av vilka kovariater som ingick ges i Tabell 3.

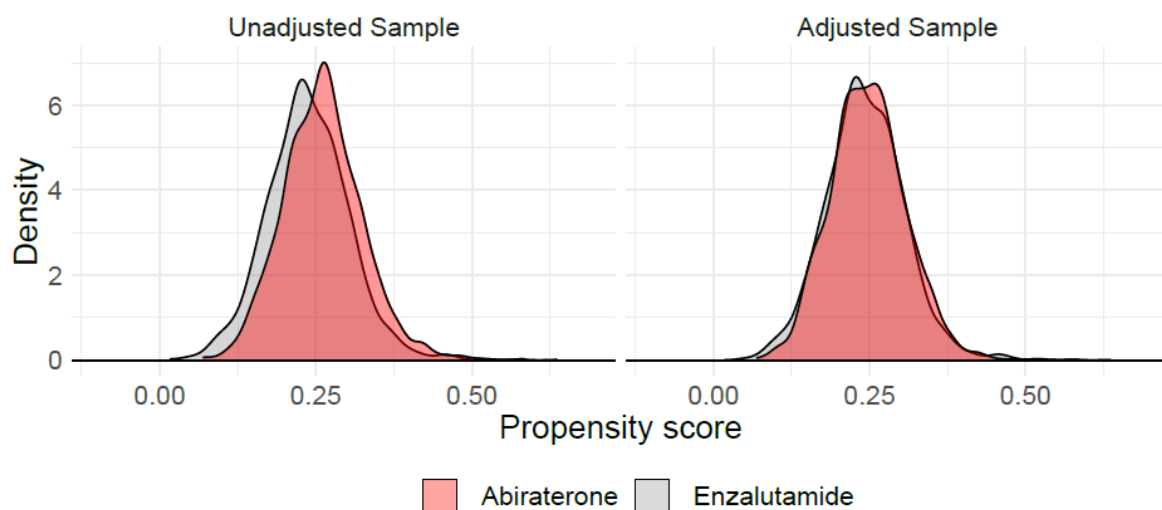
Factor	Interpretation	Cumulative Variance
Factor 1	Welfare and social security benefits	0.07
Factor 2	Wages and disposable income	0.13
Factor 3	Occupational pensions	0.18
Factor 4	Early retirement benefits and welfare	0.22
Factor 5	Sickness and unemployment benefits	0.27
Factor 6	Private pensions	0.31
Factor 7	Health before treatment	0.34
Factor 8	Income at diagnosis	0.38
Factor 9	Income from business	0.42

Tabell 2: Resultat från faktoranalysen, ett sätt att reducera antal variabler.

I figur 1 beskrivs fördelningarna för den skattade PS för de två grupperna före och efter matchning. På x-axeln ges den skattade sannolikheten att bli förskriven AA för de två grupperna och på y-axelns frekvens av patienter med den skattade sannolikheten. Från den vänstra figuren kan vi se att fördelningarna ser likartade ut redan före matching, men att populationen som får AA har, som förväntat, en högre sannolikt att få AA. Skillnaden är dock inte dramatisk. Efter vår matchning är fördelningarna i det närmsta identiska, så som skulle varit fallet om man genomfört ett grupperat randomiserat experiment, där gruppen bestämts av den skattade PS.⁵

⁴ De extra viktiga kovariaterna ingår inte bland dessa.

⁵ Ett grupperat randomiserat experiment görs ofta för att erhålla exakt balans för vissa kovariater, tex att man vill ha lika många kvinnor och män bland de som får behandling och kontroll. Ett randomiserat experiment där en av två individer med PS=x randomiseras att få AA medan den andre får EZ skulle se ut enligt den högra figuren.



Figur 1: Fördelningarna för den skattade sannolikheten att be förskrivna AA för de två populationerna som blir förskrivna AA och EZ, före och efter matchning.

Optimalt sett ska hela fördelningarna för kovariaterna vara lika för de som får AA och EZ. Av den anledningen presenteras i Tabell 3 standardiserat bias (SB) och varianskvot (VR) före och efter matchning för de kovariater som användes för att hitta matcher. SB är standardiserade medelvärdeskillnader och VR är kvoten för spridningen för kovariaterna. En tumregel är att SB ska vara mindre än 25% och att VR ska vara nära ett. Från tabellen kan vi se att före matchningen så har den skattade PS ett SB som överstiger 25% och de andra kovariaterna ett VR som många gånger är mycket lägre än ett (tex VR =0.13 för faktor sex) och även mycket högre än 1 (tex VR =1.42 för Faktor 4). Innebörden är att medelvärdeskillnaden är ganska små men att fördelningarna skiljer sig åt ganska substantiellt. Efter Matchning är PS fortsatt störst men nu är SB = 9.10. Den minsta VR är 0.76 (för faktor sex) och den största VR är 1.03. Således matchningen har åstadkommit två grupper med snarlika fördelningar av de kovariater som använts i designen.

	SB before	SB after	VR before	VR after
Age (Years)	0.28	1.10	0.96	0.90
Diabetes (Y/N)	12.48	0.70	0.77	0.99
Other CVD (Y/N)	3.65	1.77	0.96	0.98
Acute myocardial infarction, I21 (Y/N)	10.99	0.00	0.72	1.00
Atrial fibrillation and flutter, I48 (Y/N)	5.43	0.91	1.10	0.98
Osteoporosis (Y/N)	1.79	0.00	0.85	1.00
Less than secondary school education	3.15	5.08	0.98	1.03
Secondary school education	0.45	2.32	1.00	0.99
Time to treatment (Years)	6.58	0.55	1.13	0.94
Secondary malignant neoplasms (Y/N)	11.73	1.48	0.89	0.99
Time to treatment squared (Years)	7.75	1.26	1.23	0.88
Age squared (Years)	0.07	0.58	0.97	0.91
Age and Diabetes	12.41	0.45	0.77	1.00
Age and CVD	3.44	1.43	0.97	0.99
Age and Secondary malignant neoplasms	12.06	1.59	0.90	0.98
Diabetes and CVD	6.77	0.00	0.77	1.00
County specific mortality	6.05	0.23	0.99	0.90
Partner (Y/N)	4.01	6.33	0.94	0.90
Factor 1	1.03	1.48	1.08	0.72
Factor 2	1.29	0.54	0.36	0.94
Factor 3	0.66	5.03	0.68	0.81
Factor 4	5.07	2.39	1.42	0.94
Factor 5	4.37	1.48	0.71	0.95
Factor 6	5.54	3.35	0.13	0.76
Factor 7	1.30	1.90	0.53	0.78
Factor 8	0.15	0.43	0.90	0.90
Factor 9	0.21	0.35	0.48	0.97
Propensity score	37.03	9.10	0.76	0.83
Fatigue (Y/N)	4.56	0.00	0.82	1.00

Tabell 3: Standardiserat bias och varianskvot före och efter matchning.

Utfall

Vi väljer att jämföra grupperna primärt med avseende på dödlighet. Sekundärt önskar vi studera skillnader i sluten- eller öppenvårdsbesök för grav smärta eller för skelettrelaterade händelser (SRH). Smärta uppstår ofta i samband med de benmetastaser som i princip alla patienter upplever förr eller senare. Även SRH är en komplikation till dessa benmetastaser. Vi väljer också att studera skillnaden i tiden som patienter står kvar på behandling, eftersom detta är ett mått på hur länge behandlingen är effektiv eller saknar biverkningar. Enligt det kvalitativa arbetet avbryts behandling vid sådan behandlingssvikt. Vi väljer också att studera och jämföra dessa primära och sekundära utfall efter varje 30 dagars period. Det innebär att vi som mest observerar 24 perioder. Vi anser att sådana periodvisa utfall ger bäst förståelse för och kunskap om behandlingens effekt över tid. Vi får dessutom en möjlighet att studera hur robusta våra resultat är över tid. Eftersom vi har fyra utfall använder vi oss av Bonferroni justerade standardfel för respektive variabel, dvs en signifikans uppstår först då den är mindre än 1,25% risk att skillnaden skulle uppstå slumpmässigt för respektive variabel.

Undergrupper

Vi vill studera om de komparativa effekterna skiljer sig åt hos de som har en snabbare progression, eller ett mer elakartat förlopp. Vi definierar en kortare tid än mediantiden mellan diagnos av prostata

cancer och behandling med AA/EZ som ett snabbt förlopp. Vidare anser vi att patienter som har tecken på visceral metastaser har ett mer elakartat förlopp. Slutligen vill vi även förstå om de komparativa effekterna skiljer sig åt mellan de som anses ha störst chans att svara på behandlingen med AA/EZ, dvs de som tidigare har svarat väl på hormonbehandling. De som har svarat längre än 12 månader på någon av behandlingarna Leuprolin, Goserelin, Triptorelim, Histrelin, eller Degarelix anses vara s.k. high-responders.

Test och Känslighetsanalyser

Metoden som beskrivits ovan är principiellt enkel: Vi letar efter individer som får AA som är jämförbara på observerbara karaktäristika, eller kovariater, med de som får EZ och vice versa. Vi har stor variation i förskrivning mellan landsting (se tabell 4). Innebörden är att jämförbara individer hittas i stor utsträckning från olika län. Detta kan vara ett problem om det finns annan systematisk variation över landsting som kan påverka utfallet av cancerbehandling, t.ex. skillnader i vårdkvalitet. Om det skulle vara så att det finns skillnader i tid mellan diagnos och behandling mellan landsting efter det att vi kontrollerat för patienternas hälso- och socioekonomiska status skulle detta kunna tyda på att det finns kvalitetsskillnader i vården som kan snedvrider resultat från en analys. I och med att vi observera tiden mellan diagnos och behandling kan denna hypotes testas. Vi finner inga sådana skillnader mellan landsting, vilket således stödjer vår design att hitta jämförelser över landsting.

Vi har också förslagit två olika känslighetsanalyser. I den första används skillnaden i förskrivning mellan landsting för skattning, en skillnad som beskrivs i Tabell 4, inom ramen för en instrumentalvariabel analys. I den andra studeras hur känsliga resultaten är för om det skulle finnas en utelämnad variabel som både påverkar förskrivning och utfall, dvs icke observerbar confounder (se Rosenbaum 2002). För detaljer se Avsnitt 6 i Johansson m.fl. (2020)

	Enzalutamide				Abiraterone				Enzalutamide, total	Abiraterone, total
	2015	2016	2017	2018	2015	2016	2017	2018		
Blekinge	0.84	0.91	0.84	0.80	0.16	0.09	0.16	0.20	0.85	0.15
Dalarna	0.86	0.65	0.55	0.62	0.14	0.35	0.45	0.38	0.69	0.31
Gavleborg	0.60	0.92	0.93	0.96	0.40	0.08	0.07	0.04	0.85	0.15
Gotland	0.75	0.92	1.00	1.00	0.25	0.08			0.89	0.11
Halland	0.75	0.89	0.88	0.76	0.25	0.11	0.12	0.24	0.82	0.18
Jamtland	0.47	0.81	0.77	0.67	0.53	0.19	0.23	0.33	0.70	0.30
Jonkopings lan	0.62	0.91	0.89	0.76	0.38	0.09	0.11	0.24	0.80	0.20
Kalmar	0.74	0.89	0.83	0.93	0.26	0.11	0.17	0.07	0.84	0.16
Kronoberg	0.48	0.43	0.21	0.20	0.52	0.57	0.79	0.80	0.37	0.63
Norrbottnen	0.83	0.88	0.64	0.65	0.17	0.12	0.36	0.35	0.74	0.26
Orebro	0.77	0.95	0.68	0.38	0.23	0.05	0.32	0.62	0.74	0.26
Ostergotlands lan	0.12	0.90	0.88	0.90	0.88	0.10	0.12	0.10	0.83	0.17
Skane	0.87	0.96	0.96	0.85	0.13	0.04	0.04	0.15	0.92	0.08
Sodermanland	0.77	0.98	1.00	0.92	0.23	0.02		0.08	0.91	0.09
Stockholm	0.71	0.82	0.80	0.79	0.29	0.18	0.20	0.21	0.78	0.22
Uppsala	0.62	0.56	0.25	0.36	0.38	0.44	0.75	0.64	0.46	0.54
Varnmland	0.87	0.96	0.67	0.75	0.13	0.04	0.33	0.25	0.83	0.17
Vasterbotten	0.63	0.70	0.62	0.76	0.37	0.30	0.38	0.24	0.67	0.33
Vasternorrland	0.47	0.71	0.73	0.86	0.53	0.29	0.27	0.14	0.64	0.36
Vastmanland	0.73	0.79	0.80	0.55	0.27	0.21	0.20	0.45	0.74	0.26
Vastra gotalands lan	0.51	0.78	0.69	0.65	0.49	0.22	0.31	0.35	0.66	0.34
Total	0.68	0.83	0.76	0.74	0.32	0.17	0.24	0.26	0.76	0.24

Tabell 4: Variation i förskrivning av EZ och AA över landstingen.

Sammanfattning

Huvudanalys:

- Ta reda på hur läkarna förskriver gärna genom att intervjua 7-8 läkare. Eftersom det är den här selektionen du ska kontrollera för, så måste du förstå den!
- Se till att du har så mycket information som möjligt kring patientens hälso- och socioekonomiska status för tiden fram till behandling med de läkemedel som ska utvärderas.
- Målet är att hitta två jämförbara grupper m.h.a. denna information, dvs grupper som över stämmer med dessa variabler och deras interaktioner.
- Reducera en stor mängd information via faktoranalys. Presentera hur stor del av variation i informationen som fångar upp m.h.a. de faktorerna du väljer
- Skatta sannolikheten att få den ena jämfört den andra behandlingen med tex en logit modell (skattad propensity score (PS)). Använd de kovaritar som bäst predicerar sannolikheten. Detta uppnås t.ex. med LASSO regression.
- Matcha patienter på de båda behandlingarna med varandra, och tvinga match med avseende på kritiska faktorer och PS. Mahalanobis distance matchning är en metod att uppnå detta.
- Studera hur de matchade grupperna skiljer sig åt med avseende på standardiserade medelvärde och varianskvoter för alla kovariter

Känslighetsanalys:

Detta kan göras på många olika sätt men beror av vilka data som är tillgängliga och regelverk etc. Vi har föreslagit två: (1) I vårt fall så varierar förskrivningen över landsting. Detta blir vår källa till exogen information, dvs instrument som kan användas vid en instrumentalvariabel analys.; (2) Vi tillämpar också Rosenbaum (2002) känslighetsanalys för dold bias och anser att sådan analys alltid kan/ska göras.

Referenser

de Luna, Xavier; Waernbaum, Ingeborg; Richardson, Thomas (2011) Covariate selection for the non-parametric estimation of an average treatment effect, *Biometrika*, Oxford University Press 2011, Vol. 98, (4) : 861-875

Johansson Per, Paulina Jonéus and Sophie Langenskiöld (2020) Study protocol for a comparative effectiveness evaluation of abiraterone against enzalutamide. Mimeo Uppsala University.

Rosenbaum, P. (2002), *Observational Studies*, Springer Series in Statistics, Springer. URL: <https://books.google.se/books?id=K0OglGXtpGMC>