

Datainsamling och analyser av deltagandet i mammografiscreening 2014-2018

Bilaga till rapport 2020:12 Avgiftsfri mammografi – en uppföljning av satsningen på kvinnors hälsa

INLEDNING

Myndigheten för vård- och omsorgsanalys har följt upp och analyserat införandet av avgiftsfri mammografiscreening. Ett viktigt underlag till rapportens innehåll och slutsatser är en forskningsstudie genomförd av professor Sophia Zachrisson och dr. Magdalena Lagerlund vid Lunds universitet. I den här bilagan ges en detaljerad redogörelse för genomförandet av forskningsstudien samt studiens styrkor och svagheter. Resultaten från studien presenteras i Myndigheten för vård- och omsorgsanalys rapport Avgiftsfri mammografi – en uppföljning av satsningen på kvinnors hälsa (rapport 2020:12) och i den tillhörande statistikfilen. Den här bilagan har författats av Magdalena Lagerlund.

SYFTE OCH FRÅGESTÄLLNINGAR

Studien syftade till att undersöka effekterna av införandet av den nationella avgiftsfriheten för mammografiscreening 2016-07-01 (reformen) på deltagandet i screeningen. Följande övergripande frågeställningar undersöktes:

- Har deltagandet i mammografiscreening på nationell nivå förändrats sedan införandet av reformen?
- Har deltagandet i mammografiscreening i vissa socioekonomiskt utsatta grupper förändrats sedan införandet av reformen?
- Finns det skillnader i deltagande i mammografiscreening mellan regionerna och har det skett någon förändring av deltagande i de olika regionerna?

MATERIAL OCH METODER

En forskningsplan upprättades under maj – juni 2018 och ansökan till regionala etikprövningsnämnden i Lund godkändes i augusti 2018. Kompletteringsansökan godkändes i november 2018.

Avsikten var att basera studien på screeningregister i majoriteten av Sveriges regioner. Stockholms läns landsting införde avgiftsfrihet redan 2012 och Östergötland redan under 2014, men vi önskade ändå inkludera dem i studien för att använda dem som tänkbara kontrollgrupper, och för att uppnå en så heltäckande rapportering som möjligt av deltagande i mammografiscreening i landet i stort, samt i specifika sociodemografiska grupper. Stockholm bedömdes dock inte utgöra en lämplig kontrollgrupp eftersom deltagarfrekvensen är så mycket lägre där än i resten av landet. Eftersom det i dagsläget inte finns något nationellt mammografiregister undersökte vi initialt möjligheten att samla in individbaserade deltagardata från samtliga 25 enskilda mammografiverksamheter. De flesta av Sveriges regioner använder i dagsläget Sectras IT-system RIS (Radiology Information System), medan Uppsala använder Carestream och Kronoberg använder GE-RIS. På grund av bristande möjligheter att verifiera jämförbarheten mellan de olika systemen begränsade vi oss till att använda data från verksamheter som använde Sectras RIS.

Datainsamlingsprocess

Under förberedelsearbetet för datainsamlingen framkom det att verksamheterna i Jönköping och Norrbotten övergick till Sectra för sina kallelser i början av 2017, dvs under den definierade studieperioden 2014–2018. Vi beslutade att utesluta de två verksamheter på grund av att arbetet med dessa datauttag riskerade att bli omfattande och kostnadskrävande och att det inte fanns garantier för att importerade data från tidigare system skulle vara eller skulle kunna göras jämförbara med dem som Sectra registrerat efter att de tog över. Initial kontakt med samtliga 21 aktuella mammografiverksamheter togs i oktober 2018 och av dessa gav följande 18 sitt uttryckliga tillstånd till datauttag:

1. Blekinge
2. Dalarna
3. Gotland
4. Gävleborg
5. Halland
6. Jämtland/Härjedalen
7. Kalmar
8. Skåne
9. Stockholm – Capio St Görän
10. Stockholm – Sös
11. VGR – Unilabs
12. VGR – Nu-sjukvården
13. VGR (Sahlgrenska)

14. Värmland
15. Västerbotten
16. Västernorrland
17. Västmanland
18. Örebro

Tillstånd uteblev från Karolinska Sjukhuset, Sörmland och Östergötland.

De 18 separata datauttagen som Sectra gjorde under våren 2019 lades ihop till en fil som sedan avidentifierades på SCB och samkördes med befolkningsdata med hjälp av löpnummer. Årsdatafilerna för LISA (2014–2017), DeSO (2014–2018) och Hushåll (2014–2018) matchades på för respektive undersökningsår. Senast tillgängliga data används för undersökningstillfällena under 2018 och 2019.

DATARENSNING

Datarensning och samkörning av filer gjordes i följande steg:

1. Icke-träffar i födelseregistret (N=1054 observationer) Dessa kan röra sig om personnummer som skapats (hittats på) inom verksamheterna för att göra systemkontroller.
2. Återanvända personnummer (N=6001 observationer)
3. Dubletter i SCB:s filer (N=20 observationer)
4. Kön=man (N=154 observationer)
5. Kanceleringskoder som inte är screeningrelaterade har exkluderats (N=7994).
6. Undersökningskoder som inte är screeningrelaterade eller som rör tekniska omtag har exkluderats (N=36 225 observationer). Det finns dock fortfarande en risk att det döljer sig vissa cancerkontroller i observationer med screeningkod, men dessa har vi inte möjlighet att identifiera.
7. Dubletter, dvs löpnummer som har mer än en observation för samma datum
8. Dubletter, dvs löpnummer som har mer än en observation för samma år

Totalt minskade urvalet med 238 154 observationer (5,2 %), från 4 582 477 till 4 344 323 observationer och fördelade sig mellan verksamheterna enligt tabell 1.

Tabell 1 Datarensning

SITE	Förerensning		Efterrensning		Minskning	
	Antal	%	Antal	%	Antal	%
Blekinge	102071	2,2	98306	2,3	3765	3,7
Capio St Göran	425294	9,3	418582	9,6	6712	1,6
Dalarna	196788	4,3	191711	4,4	5077	2,6
Gävleborg	179061	3,9	170736	3,9	8325	4,6
Gotland	38187	0,8	37618	0,9	569	1,5
Halland	190931	4,2	187593	4,3	3338	1,7
Jämtland/Härjedalen	72749	1,6	72075	1,7	674	0,9
Kalmar	155215	3,4	153282	3,5	1933	1,2
Västmanland	149604	3,3	139782	3,2	9822	6,6
Örebro	148240	3,2	145829	3,4	2411	1,6
Skåne	832899	18,2	823010	18,9	9889	1,2
Södersjukhuset	447149	9,8	440695	10,1	6454	1,4
Värmland	183193	4	177488	4,1	5705	3,1
Västerbotten	137253	3	136285	3,1	968	0,7
Västernorrland	135708	3	135018	3,1	690	0,5
VGR NU-Sjukvården	204968	4,5	202616	4,7	2352	1,1
VGR-Sahlgrenska	218591	4,8	56910	1,3	161681	74,0
VGR-Unilabs	764576	16,7	756787	17,4	7789	1,0
Total	4582477	100	4344323	100	238154	5,2

Den kraftiga minskningen av observationer i Sahlgrenska förklaras nästan uteslutande (97%) av exkludering av en stor mängd rena kopior som registrerats både i Sahlgrenska och VGR-Unilabs. En detaljerad redovisning för antal som exkluderats i varje steg per verksamhet finns i excel Tabellen 'Datarensning'.

VARIABLER

Av de olika insamlade typer av variabler som presenteras nedan har vi i analyserna valt att lägga huvudvikten vid dem som redovisas i fetstil.

Variabler som kan användas som proxy för socioekonomisk status:

- **Utbildning** – generellt inte föränderlig över tid, speciellt inte i den här åldersgruppen
- **Disponibel inkomst (individualiserat från familj)** – känslig för t.ex. sjukdom och ålderspension som kan gälla en stor del av denna åldersgrupp.
- **Huvudsaklig inkomstkälla** enligt mikrodata för registerbaserad arbetsmarknadsstatistik (RAKS) (t.ex. förvärvsarbete, studerande, arbetslös). Med denna variabel kan vi identifiera socialt/ekonomiskt utsatta grupper av individer som under året haft sin huvudsakliga inkomstkälla från sociala ersättningar och bidrag.
- Arbete/yrke – information om tidigare yrke saknas för dem som inte längre är yrkesverksamma eller förvärvsarbetande. Det finns fortfarande inga registerdata som motsvarar socioekonomiskt index (SEI) som samlades in i folk- och bostadsräkningarna.
- **Boendeform** (ägende vs. övriga)
- Boendeara – enligt SCB finns inget definierat mått på trångboddhet enbart utifrån bostadsarean. Vi valde att dela in boendearan (antal kvadratmeter per person) i percentiler enligt fördelningen under hela studieperioden (2014, 2015, 2017 och 2018) bland kvinnor 40–75 år, inklusive Stockholm.

Andra sociodemografiska variabler:

- **Ålder** (kategoriserat i 5- och 10-årsgrupper)
- **Samboende** (som estimerats utifrån familjeställning), civilstånd, samt en kombination av dessa två. Beroende på vad syftet med användningen av denna sociodemografiska information är finns olika problem med dessa variabler. Problemet med att använda civilstånd är att alla som är gifta inte bor tillsammans och problemet med att använda samboende utifrån familjeställning är att samboende som inte har barn tillsammans blir kategoriserade som ensamstående.
- **Födelseland**
- Medborgarskap
- Utländsk bakgrund (tar hänsyn till både det egna och föräldrarnas födelseland)
- Etableringsersättning – ersättning som utbetalas till nyanlända flyktingar och deras anhöriga. Bidraget riktar sig enbart till personer i arbetsför ålder som fått uppehållstillstånd som flykting, kvotflykting eller skyddsbehövande. Här tittar vi inte på storleken

på ersättningen utan delar in populationen i dem som fått etableringsersättning under året och dem som inte fått det.

- Tid sen senaste invandring till Sverige (uträknad vid respektive undersökningstillfälle)
- Antal migrationer över svenska gränsen (sammanlagt antal in- och utvandringsposter före respektive undersökningstillfälle). Viktigt att tänka på när man använder den här variabeln är att noll migrationer sammanfaller med de flesta som fötts i Sverige och att en migration sammanfaller med kvinnor som fötts utomlands och invandrat till Sverige.
- **Bostadsort indelat i tre respektive nio typer av kommuntyper** (utifrån kriterier som tätortsstorlek, närhet till större tätort och pendlingsmönster) konstruerat utifrån kommunkod (fyra första siffrorna) i DeSO-koden (demografiska statistikområden) vid respektive undersökningstillfälle. Enligt SKL:s kommungruppsindelning 2017: <https://skr.se/tjanster/kommunerochregioner/faktakommunerochregioner/kommungruppsindelning.2051.html>

Screeningvariabler:

- **Årtal för undersökningstillfälle** (2014, 2015, 2017 och 2018)
- Årstid (höst, vinter, vår, sommar)
- **Region** (15 st) och mammografiverksamhet (18 st)
- Screeningomgång
- **Deltagarstatus** (deltagande, aktivt och passivt icke-deltagande) – vi redovisar andelen passivt deltagande totalt och per verksamhet och region, men använder sedan enbart deltagande och icke-deltagande i våra jämförande analyser.

DEFINITION AV DELTAGANDE

Socialstyrelsen har tagit fram 11 indikatorer för att följa upp och utvärdera screening för bröstcancer (Socialstyrelsen 2014b). För denna forskningsstudie fokuserar vi på den tredje indikatorn – ”andel kvinnor som efter kallelse till screening för bröstcancer genomfört mammografiundersökning, i procent”. Enligt deras tekniska beskrivning skall deltagarfrekvensen räknas ut enligt följande:

- Täljare: antal kvinnor som mammografiundersökts efter kallelse till screening för bröstcancer.
- Nämnare: totala antalet kvinnor som kallats till mammografiundersökning under ett kalenderår.
- Uppföljningsperiod: årsvis uppföljning. Socialstyrelsen rekommenderar att kvinnor kallas till screening för bröstcancer med 18–24 månaders intervall. Uppföljning bör ändå

genomföras årsvis då man kan anta att minst halva målpopulationen har deltagit i screeningprogrammet.

- Definition kalenderår: datum på kallelsen till mammografiundersökning utgör underlag.

Vår ursprungliga avsikt var att undersöka mammografideltagande inom 90 dagar från en primärbokad tid enligt överenskommelse inom en nationell arbetsgrupp inom Sveriges Kommuner och Landsting (Törnberg 2014). Om data redovisas efter för kort uppföljningstid finns det en risk att de kvinnor som kallats i slutet av ett kalenderår inte hunnit delta och att man då får en falskt låg deltagarfrekvens. Detta var anledningen till att vi väntade med att göra datauttagen till april 2019 så att 90 dagar skulle ha passerat efter slutet på 2018. När förarbetet till datauttagen påbörjades visade det sig dock att information om primärt kallelsedatum saknades helt före 2017 för samtliga verksamheter. Att återskapa denna information skulle inte ha varit omöjligt, men det merarbete som det skulle kräva överskred den tillgängliga projektbudgeten. En jämförelse av deltagande före och efter införandet av avgiftsfrihet måste baseras på samma definition av deltagande för respektive tidsperiod och vi har därför inte kunnat använda oss av den önskvärda 90-dagardefinitionen, utan har varit tvungna att begränsa oss till de faktiska undersökningsdatumen inom respektive kalenderår. Vi gör dock antagandet att den underskattning av deltagande som eventuellt blir konsekvensen är relativt konstant över tid.

Icke-deltagare kan delas upp i två grupper – aktiva icke-deltagare, som aktivt har lämnat återbud till sin kallelse, och passiva deltagare, som har uteblivit från sin tid utan att uppge skäl eller hört av sig. I våra analyser undersöker vi icke-deltagare som den kombinerade gruppen av aktiva och passiva icke-deltagare, och redovisar endast initialt förekomsten av de två olika typerna av icke-deltagande över tid och per region.

I vårt datamaterial har vi en observation för respektive undersökningstillfälle. Ibland sammanfaller undersökningsdatumet med det primära kallelsedatumet (detta blir fallet om kvinnan kommer på utsatt tid eller om hon uteblir från utsatt tid), men om kvinnan bokar om sin undersökningstid sammanfaller inte dessa två datum. När vi talar om ett undersökningstillfälle eller undersökningsdatum innebär det inte att mammografiundersökningen faktiskt ägde rum vid detta tillfälle, utan att detta var det tillfälle/datum då undersökningen var planerad att äga rum, vare sig kvinnan deltog eller inte, och vare sig om det var den ursprungliga undersökningstiden eller en ombokad tid.

POPULATIONEN

Studiepopulationen består av samtliga kvinnor mellan 40 och 75 år som kallats till mammografiscreening under åren 2014, 2015, 2017 och 2018 i de 18 inkluderade screeningverksamheterna och uppgår till 3 254 437 observationer och 1 447 299 individer. Vi analyserar även en population som exkluderar Stockholm och denna uppgår till 2 592 455 observationer och 1 126 669 individer. Eftersom det här är en undersökning av totalpopulationen i dessa screeningverksamheter, så har vi inte att göra med ett urval i vanlig statistisk mening.

Beskrivande statistik av populationen (både med och utan Stockholm) finns i exceltabellen 'Studiepopulationen'. Det är viktigt att poängtera att frekvenserna som redovisas här gäller undersökningstillfällena och inte individer, eftersom samma individ kan vara inkluderad upp till fyra gånger (dock inte mer än en gång per kalenderår) och nästan samtliga variabler (både beroende och oberoende) kan variera över tid. De regioner som bidrar med flest observationer är Skåne (18,9%), St Görans och Sös (20,3%) och VGR (23,3%). De minsta regionerna som finns representerade i populationen är Gotland (0,9%) och Jämtland/Härjedalen (1,6%).

STATISTISKA ANALYSER

I jämförande analyser har procentuellt deltagande och 95-procentiga konfidensintervall beräknats för den totala populationen, för enskilda regioner, för olika tidsperioder och för olika sociodemografiska grupper. Med hjälp av logistiska regressionsmodeller har oddskvoter (OR) och 95-procentiga konfidensintervall för icke-deltagande beräknats för olika regioner och sociodemografiska grupper. Vi gör dels bivariata analyser där vi undersöker effekten av en variabel på deltagandet och dels multivariabla analyser där vi tittar på effekten av flera variabler samtidigt i en modell för att kunna undersöka den oberoende effekten av varje variabel när man kontrollerar för övriga variabler i modellen. I de multivariabla analyserna kan något färre individer ingå i jämförelse med de bivariata analyserna på grund av att det sammanlagda bortfallet (missing values) blir större när flera variabler kombineras i en modell. Utöver oddskvoter redovisas även marginaleffekter, som är lättare att tolka och begripa.

Man bör vara försiktig med att fokusera för mycket på statistiskt signifikanta skillnader eftersom mycket små skillnader kan vara statistiskt signifikanta (dvs. konfidensintervall som inte inkluderar 0 när man undersöker skillnader, eller konfidensintervall som inte inkluderar OR=1 när man tolkar resultat av logistisk regression) i ett så stort datamaterial som vi har. Det är därför viktigare att fundera på vad man anser vara kliniskt relevanta skillnader.

För att undersöka deltagarfrekvensen över tid har vi dels tittat på varje enskilt kalenderår (2014, 2015, 2017 och 2018) och dels på en kombination av två år före avgiftsfrihetens införande (2014+2015) och efter (2017+2018). Att kombinera två år är viktigt för att fånga in alla individer som kallas under en hel undersökningscykel för ett geografiskt upptagningsområde, vilken kan sträcka sig över ett till två år (beroende på verksamhet). Genom denna sammanslagning av kalenderår jämnar man dessutom ut den fluktuation som förekommer naturligt för de flesta fenomen som studeras vid flera mättillfällen. Inom dessa tvåårsperioder kan samma individ finnas inkluderade med upp till två kallelsetillfällen. Vår resultatredovisning baseras dock på analyser där vi inkluderat enbart det senaste kallelsetillfälle per individ inom respektive period. På så sätt kan vi studera det senaste observerade beteendet i populationen inom respektive tidsperiod och vi kan använda perioden 2017+2018 för att beskriva nuläget i Sverige. Vi undviker därmed också att vissa

individer får större vikt i analysen (för att de inkluderats två gånger), vilket annars måste kontrolleras för.

I en särskild analys (paneldataanalys) inkluderar vi ett urval av individer som har kallats både före och efter 2016 och inkluderar även där den senaste observationen för respektive 24-månadersperiod. Totalt har vi då en population på 1 303 616 kvinnor (inklusive Stockholm). Denna analys är tänkt som en kontrollanalys av förändringar över tid. Om våra resultat kvarstår när den multivariabla analysen körs i denna panelpopulation kan man göra antagandet att eventuella förändringar över tid inte har orsakats av förändringar i befolkningssammansättningen, utan av andra faktorer (t.ex. strukturella/organisatoriska). Vissa sociodemografiska förändringar kan ändå ske för varje enskild individ.

Anledningen till att vi inkluderar Stockholm (St Görans och Sös) i analyser är att vi vill få en så heltäckande bild av deltagandet i Sverige som möjligt och då kan inte populationen i Stockholm utelämnas. Anledningen till att vi gör samma analyser även utan Stockholm är att deltagandet skiljer sig så starkt mellan Stockholm och resten av landet, samt att Stockholm införde avgiftsfrihet redan 2012, vilket gör att det inte är relevant att inkludera Stockholm vid utvärdering av en eventuell effekt av den reform som infördes 2016. I Stockholm genomgick mammografiscreeningen även en större omorganisation under perioden 2014-2016 som kan ha påverkat deltagandet och som gör Stockholm olämplig som kontrollgrupp. Deltagandet i screeningverksamheten vid Karolinska kan också eventuellt avvika från deltagandet vid St Görans sjukhus och Södersjukhuset och om så är fallet har vi inte ett helt representativt urval av kvinnor som kallas till mammografiscreening i Region Stockholm.

STUDIENS STYRKOR OCH SVAGHETER

Denna studie är den största som gjorts i Sverige och inkluderar totalbefolkningen i 14 regioner och i två av tre verksamheter i Region Stockholm av landets 21 regioner, och täcker ca 80% av Sveriges befolkning. Den viktigaste anledningen till att samla in ett så stort material är att vi kan uttala oss om deltagarfrekvensen och mönster i merparten av Sveriges mammografiverksamheter. En annan fördel är att man kan upptäcka och statistiskt säkerställa även mycket små skillnader. I våra tabeller och resultat har vi presenterat estimat med 95-procentiga konfidensintervall. Man måste dock hålla i åtanke att denna typ av statistisk metod har utvecklats för att hantera den osäkerhet som föreligger när man gör urval ur en population. I en totalbefolkning är statistisk osäkerhet inte lika relevant. Istället för att bedöma statistisk signifikans, bör man snarare avgöra vad som utgör kliniskt meningsfulla skillnader i mammografiprogrammet och i verksamheterna.

Sverige har en lång tradition av att hantera och kvalitetssäkra populationsbaserade registerdata på individnivå för både sociodemografisk och medicinsk information. Data för både den beroende variabeln (deltagande) och de oberoende variablerna (sociodemografisk information) är registerbaserade och av mycket hög kvalitet. Screeningdata har tagits fram

av samma företag och på samma sätt för alla 18 screeningverksamheter som ingår i vår studiepopulation.

Som vi beskrivit närmare i metoddelen hade vi inte möjlighet att undersöka deltagande inom en 90-dagarsperiod efter ursprungligt kallelsedatum. På grund av detta finns det en risk att vi har felskattat deltagarfrekvensen något för respektive undersökningsperiod. En eventuell sådan felskattning bör dock ha påverkat båda undersökningsperioderna på motsvarande sätt, vilket betyder att vi ändå kan göra en giltig jämförelse av deltagarfrekvenserna under respektive tidsperiod.

Vad gäller enskilda sociodemografiska variabler finns ett visst bortfall. Bland de variabler som vi har inkluderat i de multivariabla modellerna har vi högst bortfall för utbildningsvariabeln (1,3%), följt av disponibel inkomst, huvudsaklig inkomstkälla, samboende (samtliga 0,3%) och födelseland (0,006%). För samtliga variabler är bortfallet högre bland icke-deltagare: 3,8% för utbildning, 1,2% för disponibel inkomst, huvudsaklig inkomstkälla och samboende, och 0,013% för födelseland. I en tidigare svensk studie där samband mellan sociodemografiska faktorer och deltagande i mammografiscreening undersöktes fann man att kvinnor som saknade information för enstaka variabler generellt hade lägre socioekonomisk status, var äldre och födda utanför Norden, vilket pekade på att man skulle ha underskattat den verkliga skillnaden mellan deltagare och icke-deltagare.

Någon idealisk registerbaserad variabel för socioekonomisk status har vi tyvärr inte. I denna studie har vi valt utbildning, inkomst och huvudsaklig inkomstkälla som bästa tillgängliga indikatorer på socioekonomisk status. De visar samtliga något olika samband med deltagande, men påverkar alla deltagande i samma riktning. Det hade varit önskvärt att kunna inkludera data för socioekonomiskt index (SEI), en klassifikation som publicerades 1982 och i huvudsak baserades på uppgifter om individers yrken. SCB har inte samlat in och erbjudit denna typ av data sen Folk- och bostadsräkningen 1990. Även om denna indelning är förhållandevis föråldrad var den anpassad efter svenska förhållanden och någon motsvarighet har inte utvecklats. Det som kommer närmast vad gäller yrkesindelning är SSK, men för vår specifika åldersgrupp är det stora bortfallet ett avsevärt problem p.g.a. att yrkesdata inte finns för kvinnor som gått i pension. Indelningen i arbetar- och tjänsteklass är också relativt grov. För den nyare europeiska indelningen ESEG finns samma problem med bortfall bland pensionärer som för SSK och har kritiserats för att den inte är utvecklad i enlighet med unika svenska förhållanden.

Införandet av avgiftsfriheten kan betecknas som ett så kallat naturligt experiment. Det genomfördes i hela landet utan att välja ut kontrollgrupper där t.ex. avgiften behölls, eller att kunna kontrollera för övriga regionala åtgärder som kan ha genomförts under samma tidsperiod. Det enda mätinstrument vi har till vårt förfogande för att undersöka effekten av avgiftsfriheten på deltagande är att låta undersökningstillfällena före reformens införande vara kontrollgrupp till undersökningstillfällena efter, och att justera analyserna för region. Det finns alltså inget sätt att garantera att eventuella förändringar har orsakats av reformen utöver den information som samlats in från representanter för verksamheterna.

Vi samlade in data enbart för tidsperioden 2014–2018 och kunde alltså inte fastställa en längre trend bakåt i tiden eller undersöka livstidsdeltagande bland kvinnorna. På grund av tidsåtgången för att samla in tillstånd och göra datauttagen hade vi heller inte möjlighet att inkludera data för 2019. De antal mätpunkter vi har är för få för att kunna göra en meningsfull tidsserieanalys. Det kan dock vara så att effekterna av avgiftsfrihet tar längre tid än två år för att visa sig och att vi skulle behövt studera en längre tidsperiod efter reformens införande för att kunna påvisa detta. Alla kvinnor deltar inte vid varje efterföljande screeningtillfälle och det kan ta tid innan kunskapen om avgiftsfrihet blir manifest i målgruppen. Alla regioner har till exempel inte aktivt informerat om avgiftsfrihet utöver den information som ges i kallelsebrevet och då är det framför allt kvinnor som faktiskt kommer till screening som är medvetna om avgiftsfriheten. Med uppstart av det nationella kvalitetsregistret för mammografiscreening under 2020 finns på sikt möjligheter att göra längre tidsstudier. Det vore dock ändå svårt att frigöra en eventuell effekt av andra reformer och orsaker till förändrade deltagarmönster hos kvinnor från reformen för avgiftsfrihet, då variationen i deltagande över tid ser ut att vara liten.

I vårt förarbete med granskning och rensning av datafiler fann vi dubbla undersökningstillfällen för samma datum inom och över verksamhetsgränserna. Inom samma verksamhet var det inte något problem att rensa bort dessa dubletter (N=76), men när sådana dubletterna förekom över verksamhetsgränserna (N=1622) beslöt vi att behålla den observation som hade lägst värde på variabeln screeningomgång. Eftersom denna variabel inte är tillförlitlig när man studerar observationer över verksamhetsgränserna är det osäkert om rätt observation behövs. Alternativet hade varit att slumpmässigt välja vilken observation som skulle behållas. Ingen av metoderna är idealisk, men att behålla dessa dubletter hade varit det sämsta alternativet. Vi tog även beslutet att exkludera dubletter som inföll inom samma kalenderår (N=28 344) för att undvika dubbla observationer för samma individ (i praktiken ska screeningprogrammet inte kalla kvinnor oftare än var 18:e månad). När kanceringskoderna var densamma valde vi att behålla det undersökningstillfälle som inträffade först (N=5347). Hittills i rensningsarbetet har vi alltså inte påverkat deltagarfrekvensen utan bara antal undersökningstillfällen inom respektive verksamhet. För resterande dubletter inom samma kalenderår (N=22 997) där undersökningskod, kanceringskod och verksamhet kan ha varierat valde vi i första hand den observation där kvinnan deltagit och i andra hand valde vid det tidigare undersökningsdatumet. Genom att exkludera observationer där kvinnorna inte deltagit kan vi alltså ha överskattat deltagarfrekvensen något. Det rör sig endast om 0,5% av vårt slutgiltiga urval och det bör ha påverkat alla undersökningsperioder i lika stor utsträckning.

Det finns en risk att det döljer sig cancerkontroller i våra screeningdata, dvs de kvinnor som haft bröstcancer och sedan kontrolleras för återfall och cancer i andra bröstet en gång per år under 5–10 år efter diagnos. En del landsting använder screeningsystemet för att kalla dessa kvinnor. I vårt datarensningarbete såg vi en misstänkt hög förekomst av tätare kontroller i Dalarna, Gävleborg, Halland, Värmland och Västmanland. Det är sannolikt att kvinnor som haft bröstcancer och går på cancerkontroller har högre deltagarfrekvens än kvinnor som

aldrig haft bröstcancer. Dessa misstankar bekräftas också när vi undersöker deltagarfrekvensen bland misstänkta cancerkontroller (data ej inkluderade). Under perioden 2014–15 var 2922 (0,2%) av undersökningstillfällena misstänkta cancerkontroller och deltagandet i denna grupp var 86,3% jämfört med 81,5% bland dem som inte var misstänkta cancerkontroller. Under perioden 2017–18 var 1571 (0,1%) av undersökningstillfällena misstänkta cancerkontroller och deltagandet i denna grupp var 89,6% jämfört med 81,3% bland dem som inte var misstänkta cancerkontroller. Dessa misstänkta cancerkontroller handlar alltså om ett litet antal observationer och de påverkar inte utfallet varken totalt eller i enskilda regioner.

Vi hade inte möjlighet att undersöka effekten av screeningomgång eftersom denna variabel inte var tillförlitlig över verksamhetsgränserna. Screeningomgång är en systemvariabel som varken Sectra eller verksamheterna administrerar och som börjar på 1 för varje specifik individ i specifik databas (verksamhet). Tidigare svensk forskning (Törnberg 2005) har t.ex. visat att den första inbjudan och undersökningen är särskilt viktiga för framtida deltagande. En dansk studie visade att deltagarfrekvensen sjönk med antal inbjudningar och drog slutsatsen att deltagarfrekvens på populationsnivå tenderar att överdriva täckningsgraden för enskilda individer (von Euler-Chelpin 2006). Det hade därför varit önskvärt att kunna studera denna variabel i detalj och att kunna justera våra analyser för screeningomgång.