

Nr. 1 ■ 2013

# Dataindsamlingsmetodens betydning

Befolkningsundersøgelser om sundhed og sygelighed baseret på henholdsvis personligt interview og selvadministreret spørgeskema

**Anne Illemann Christensen**

Statens Institut for Folkesundhed, Syddansk Universitet

## Resumé

Personligt interview anses sædvanligvis som den gyldne standard når det gælder dataindsamlingsmetode, men selvadministrerede spørgeskemaer benyttes dog i større og større omfang på grund af ressourcemæssige begrænsninger. Det er derfor væsentligt at undersøge, om der er forskelle i deltagelsesprocent, repræsentativitet og svarfordeling i undersøgelser hvor der er anvendt forskellig dataindsamlingsmetode. I denne artikel sammenlignes to befolkningsundersøgelser om sundhed og sygelighed, med hensyn til bortfald og svarfordeling for centrale indikatorer. Resultaterne er baseret på sundheds- og sygelighedsundersøgelsen 2010 (selvadministreret spørgeskemaundersøgelse) og en interviewundersøgelse i Region Syddanmark (personligt interviewundersøgelse). Sociodemografiske oplysninger for alle inviterede personer blev indhentet fra administrative registre, og der blev anvendt logistisk regression til at undersøge sammenhængen mellem bortfald og sociodemografiske faktorer i de to undersøgelser. Logistisk regressionsanalyse blev endvidere anvendt til at undersøge, hvorvidt der var forskel på svarfordelingen for centrale indikatorer, fx daglig rygning, svær overvægt og brug af alternativ behandling, mellem de to undersøgelser.

Overordnet var bortfaldet større i den selvadministrerede undersøgelse (37,9 %) end i interviewundersøgelsen (27,3 %), men bortfaldsmønstret i forskellige sociodemografiske grupper var generelt det samme i begge undersøgelser. Dataindsamlingsmetoden havde signifikant betydning for forekomsten af nogle indikatorer, fx indikatorer relateret til helbredsrelateret livskvalitet, sundhedsadfærd og sociale relationer. For andre indikatorer blev der ikke fundet nogen forskel på forekomsten mellem de to undersøgelser, fx indikatorer relateret til brug af sundhedsvæsenet. Valg af dataindsamlingsmetode kan derfor have stor betydning hvis man skal sammenligne resultater fra undersøgelser, hvor der er anvendt forskellig dataindsamlingsmetode. Anvendelse af forskellige dataindsamlingsmetoder kan endvidere vanskeliggøre sammenligninger over tid. Ud fra nærværende undersøgelse er det dog ikke muligt at vurdere hvilken dataindsamlingsmetode der giver de mest korrekte svar. Andre forhold såsom økonomiske og tidsmæssige ressourcer må ligeledes tages i betragtning når man skal vælge den mest hensigtsmæssige form for dataindsamling.

## Indledning

I Danmark findes en række administrative registre, der kan bruges til kortlægning og monitorering af befolkningens sundhedstilstand. Disse registre omfatter sygdomssymptomer/tilstande, der har ført til kontakt med sundhedsvæsenet, men ikke oplysninger om gener og symptomer i dagligdagen, der ikke medfører lægebesøg, og heller ikke oplysninger om helbredsrelateret livskvalitet, sundhedsadfærd og funktionen af sociale relationer. Registerdata suppleres derfor ofte med data fra spørgeskemabaserede undersøgelser, hvor det er muligt at få disse oplysninger. I forbindelse med spørgeskemabaserede undersøgelser kan der anvendes flere forskellige dataindsamlingsmetoder. Sædvanligvis anses personligt interview som den gyldne standard, når der gælder dataindsamlingsmetode, men selvadministrerede spørgeskemaer anvendes dog i større og større omfang på grund af ressourcemæssige begrænsninger (Bowling, 2005; de Leeuw, Hox & Dillmann, 2008). Flere undersøgelser har vist, at dataindsamlingsmetoden kan have betydning for deltagelsesprocent, repræsentativitet og svarfordelingen for centrale indikatorer.

Eksempelvis viser nogle studie, at undersøgelser baseret på personligt interview giver en højere deltagelsesprocent end papirbaserede undersøgelser, der til gengæld giver en højere deltagelsesprocent end internetbaserede undersøgelser. Dette kan muligvis tilskrives metodens evne til at formidle undersøgelsens vigtighed og troværdighed. Fx kan tilstedeværelsen af en interviewer være med til at motivere respondenterne og understrege vigtigheden af respondentens deltagelse. I litteraturen er der dog ikke enighed om dette og nogle studier viser, at dataindsamlingsmetoden ikke har nogen betydning for deltagelsesprocenten (Mindell et al, 2010; Chittleborough et al, 2010; Ekholm et al, 2010; de Leeuw, Hox & Dillmann, 2008). Befolkningsgrupper kan desuden have forskellige præferencer eller opfattelse af metoderne. Eksempelvis er der studier der viser, at internetbaserede undersøgelser i højere grad henvender sig til den yngre del af befolkningen end papirbaserede undersøgelser (Beebe et al, 2005; de Leeuw, Hox & Dillmann, 2008). Dataindsamlingsmetoden kan ligeledes have betydning for hvordan respondenterne svarer på spørgsmålene. Eksempelvis viser litteraturen, at hvis der er en interviewer tilstede kan respondenterne være mindre tilbøjelige til at videregive følsomme oplysninger, og rapportere social uacceptabelt adfærd sammenlignet med undersøgelser, hvor der anvendes selvadministrerede spørgeskemaer. Dette kan muligvis tilskrives, at respondenterne har en større følelse af anonymitet i selvadministrerede undersøgelser (Bowling, 2005; de Leeuw, Hox & Dillmann, 2008). I litteraturen er der dog ikke enighed om, hvilken betydning dataindsamlingsmetoden har for besvarelsen af følsomme emner. En interviewer kan fx være med til at præcisere, uddybe og forklare evt. tvivls-

spørgsmål (de Leeuw, Hox & Dillmann, 2008). Udover tilstedeværelse eller fravær af en interviewer har det betydning, i hvilken form spørgsmål og information bliver kommunikeret og præsenteret dvs. visuelt, verbalt eller begge dele. Hvis spørgsmål og information kun bliver kommunikeret verbalt (som i nogle interviewundersøgelser) læges der en større byrde på respondentens hukommelse, og der som følge heraf vil have en tendens til oftere at vælge den først- eller sidstnævnte svarkategori frem for de øvrige svarkategorier. Visuel præsentation eller en kombination af visuel og verbal præsentation vil derfor mindske byrden på respondentens hukommelse. I selvadministrerede internetbaserede spørgeskemaundersøgelser er situationen dog anderledes, da internetbrugere oftere er mere utålmodige, og vælger som konsekvens ofte en af de første svarkategorier, da de ikke læser hele listen af svarkategorier (de Leeuw, Hox & Dillmann, 2008).

På baggrund af ovenstående må det forventes, at dataindsamlingsmetoden i større eller mindre omfang har betydning for repræsentativitet og fordeling af data og dermed for undersøgelsens resultater. Det er derfor vigtigt, at man overvejer hvilken betydning dataindsamlingsmetoden kan have for ens resultater. I sær hvis man sammenligner resultater fra undersøgelser, der har anvendt forskellig dataindsamlingsmetode.

Statens Institut for Folkesundhed (SIF), Syddansk Universitet, har siden 1987 med jævne mellemrum gennemført nationalt repræsentative undersøgelser af den voksne (16 år eller derover) befolknings sundhed og sygelighed (SUSY-undersøgelserne) (Ekholm et al, 2009). Formålet med undersøgelserne er dels at beskrive forekomsten og fordelingen af sundhed og sygelighed i befolkningen, og dels at opstille tidsserier, der kan beskrive udviklingen. I undersøgelserne indsamles data om sundhed og sygelighed og forhold af betydning herfor, som ikke er tilgængelige i den del af de nationale danske informationssystemer, der dækker den institutionsbehandlede sygelighed, dødelighed og sociale begivenheder. Undersøgelserne har alle været gennemført som personlige interviews i respondentens hjem.

I 2010 blev der gennemført en ny SUSY-undersøgelse. Her blev dataindsamlingsmetoden ændret fra personlige interviews til selvadministrerede spørgeskemaer. Det kan derfor være svært at afgøre, om en udvikling over tid kan tilskrives en reel ændring eller en ændring i dataindsamlingsmetoden. For at afdække SUSY-2010 undersøgelsens sammenlignelighed på dette område med de tidligere SUSY-undersøgelser har SIF gennemført personligt interview på en mindre stikprøve i samme tidsperiode og med identiske spørgsmål som i SUSY-2010 undersøgelsen. Formålet med denne undersøgelse var at beskrive forskelle i deltagelsesprocenter, bortfald og svarfordelingen for centrale indikatorer mellem undersøgelser baseret på to forskellige spørgeskemabaserede dataindsamlingsmetoder.

## Materiale og metode

### Interviewundersøgelsen

Til interviewundersøgelsen blev der udtrukket 1.500 personer i alderen 16 år eller derover bosiddende i Region Syddanmark pr. 1. januar 2010. Der blev udtrukket en simpel tilfældig stikprøve fra CPR-registeret (Pedersen, 2011). Alle udtrukne personer modtog et introduktionsbrev, der kort beskrev formålet og indholdet af undersøgelsen, og det blev understreget, at deltagelsen var frivillig. Data blev indsamlet ved personligt interview i respondentens hjem (mindst fire kontakt forsøg) og gennemført af professionelle interviewere fra Det Nationale Forskningscenter for Velfærd. Data blev indsamlet ved Computer-Assisted Personal Interview (CAPI). Invitationerne blev udsendt løbende, og data blev indsamlet kontinuerligt fra februar 2010 til april 2010.

### SUSY-2010 (den selvadministrerede undersøgelse)

Til SUSY-2010 blev der udtrukket en national repræsentativ stikprøve på 25.000 personer i alderen 16 år eller derover bosiddende i Danmark pr. 1. januar 2010. Stikprøven er baseret på et simpelt tilfældigt udtræk fra CPR-registeret (Pedersen, 2011). Alle udtrukne personer modtog et introduktionsbrev, der kort beskrev formålet og indholdet af undersøgelsen, og det blev understreget, at deltagelse var frivillig. I brevet blev de udtrukne personer opfordret til enten at besvare spørgeskemaet online eller besvare det vedlagte papir spørgeskema. Brevet indeholdt et unikt brugernavn og en adgangskode, som gav adgang til online spørgeskemaet. Invitationerne blev udsendt 1. februar, og data blev indsamlet kontinuerligt fra februar 2010 til april 2010. Der blev udsendt to påmindelsesbreve hvor den sidste påmindelse, foruden brevet, indeholdt et nyt papir spørgeskema. Da interviewundersøgelsen alene er gennemført i Region Syddanmark anvendes kun den del af SUSY-2010 der er gennemført i Region Syddanmark (n=5.502). Denne del af SUSY-2010 undersøgelsen kaldes i det efterfølgende for 'den selvadministrerede undersøgelse'.

### Spørgeskemadata

I begge undersøgelser blev der brugt identiske spørgsmål inden for emnerne socio-demografiske karakteristika, helbredsrelateret livskvalitet, sundhedsadfærd, sygelighed, brug af sundhedsvæsenet, konsekvenser af sygdom og sociale relationer. I denne undersøgelse sammenlignede vi 14 centrale indikatorer, som beskrives i det følgende. Helbredsrelateret livskvalitet blev belyst ved spørgsmålet: 'Føler du dig frisk nok til at gennemføre det, som du har lyst til at gøre', med svarkategorierne: 'Ja, det meste af tiden', 'Ja, lejlighedsvis', 'Nej (næsten aldrig)' og 'Ved ikke'. Information om

daglig rygning og BMI blev belyst ud fra spørgsmål om rygeadfærd samt højde målt i meter og vægt målt i kg. Fysisk aktivitet blev belyst ved spørgsmålet: 'Hvis du ser på det seneste år, hvad ville du så sige passer bedst som beskrivelse af din fysiske aktivitet i fritiden?', med svarkategorierne: 'Træner hårdt og dyrker konkurrenceidræt regelmæssigt og flere gange om ugen', 'Dyrker motionsidræt eller udfører tungt havearbejde eller lignende mindst 4 timer om ugen', 'Spadsere, cykler eller har anden lettere motion mindst 4 timer om ugen (medregnet også søndagsture, lettere havearbejde og cykling eller gang til arbejde)', og 'Læser, ser fjernsyn eller har anden stillestående beskæftigelse'. Respondenterne blev ligeledes spurgt, om de har nogen langvarig sygdom, og efterfølgende om de på nuværende tidspunkt havde eller havde haft en række specifikke sygdomme (herunder kronisk lungesygdom og diskusprolaps). Brug af sundhedsvæsenet blev belyst ved spørgsmål om, hvorvidt respondenterne nogensinde havde gjort brug af behandlere uden for det almindelige sundhedsvæsen, om respondenterne i løbet af de seneste 5 år har gået regelmæssigt til tandlæge (mindst en gang om året) for eftersyn af tænderne, og hvornår de sidst havde fået målt deres kolesterol. Social relation blev målt ved spørgsmål om, hvorvidt de nogensinde var alene, selv om de ville foretrække at være sammen med andre mennesker ('Ja, ofte', 'Ja, en gang imellem', 'Ja, men sjældent', 'Nej'), om de har nogen at tale med, når de har problemer eller brug for støtte ('Ja, ofte', 'Ja, for det meste', 'Ja, nogen gange', 'Nej, aldrig eller næsten aldrig'), og hvis de fik syge, og havde brug for hjælp til praktiske problemer, om de så kunne regne med hjælp fra andre ('Ja, helt sikkert', 'Ja, måske', 'Nej'). Spørgsmålene har alle været anvendt i tidligere danske sundhedsundersøgelser. De dikotomiserede variable der blev anvendt i analyserne fremgår af tabel 2.

### **Registerdata**

Alle udtrukne personer til de to undersøgelser blev ved brug af deres cpr-nummer koblet til administrative registre. Derved er det muligt at gennemføre en detaljeret beskrivelse af deltagere og ikke-deltagere i de to undersøgelser. I begge undersøgelser blev oplysninger om køn, alder, civilstand, statsborgerskab, fødeland samt forældres fødeland indhentet fra CPR-registeret (Pedersen, 2011). Oplysninger omkring uddannelsesniveau blev indhentet fra uddannelsesregisteret (Jensen & Rasmussen, 2011).

### **Statistiske analyser**

Bortfaldet i de to undersøgelser blev undersøgt ved brug af multipel logistisk regressionsanalyse. Ud fra den multiple logistiske regressionsanalyse blev der beregnet odds ratio (OR) med 95 % CI. Formålet med den logistiske regressionsanalyse var at give en nærmere beskrivelse af hvilke sociodemografiske faktorer der havde betydning

for bortfaldet, når man samtidig kontrollerede for de øvrige sociodemografiske faktorer. Endvidere blev der gennemført multiple logistisk regressionsanalyser for at undersøge hvorvidt der var forskel på forekomsten af de dikotomiserede variable mellem de to undersøgelser når der samtidig blev kontrolleret for køn, alder, civilstand, etnisk baggrund og højeste fuldførte uddannelsesniveau. For hver variabel blev der endvidere undersøgt, om der var en vekselvirkning mellem dataindsamlingsmetode og henholdsvis køn, alder, civilstand, etnisk baggrund og højeste fuldførte uddannelsesniveau. Dette blev ligeledes gjort ved brug af multiple logistisk regressionsanalyser. I alle analyser blev der anvendt 5 % som signifikansniveau. De logistiske regressionsanalyser blev alene gennemført for aldersgruppen 25 år eller derover, for at sikre, at størstedelen havde afsluttet deres uddannelse.

## Resultater

Af tabel 1 fremgår for hver af de to undersøgelser dels den totale bortfaldsprocent, dels bortfaldsprocenten fordelt på køn, alder, civilstand, etnisk baggrund og højeste fuldførte uddannelsesniveau. I interviewundersøgelsen er der et bortfald på 23,7 %. Bortfaldet er større blandt mænd (25,9 %) end blandt kvinder (21,5 %). I forhold til civilstand er bortfaldet mindst blandt gifte (18,5 %) og størst blandt ugifte (32,6 %). Bortfaldet er endvidere stort blandt personer med ikke-vestlig (44,5 %) og anden vestlig (45,1 %) baggrund. Den logistiske regressionsanalyse bekræfter, at bortfaldet afhænger af civilstand ( $p < 0,0001$ ), etnisk baggrund ( $p < 0,0001$ ) og højeste fuldførte uddannelsesniveau ( $p = 0,0118$ ), men ikke af køn ( $p = 0,1860$ ) og alder ( $p = 0,2635$ ).

I den selvadministrerede undersøgelse er der et bortfald på 37,9 % (tabel 1). Bortfaldet er større blandt mænd (42,5 %) end blandt kvinder (33,2 %). Bortfaldet er størst i aldersgruppen 16-24 år. I forhold til civilstand er bortfaldet mindst blandt gifte (29,5 %) og størst blandt ugifte (50,9 %). Bortfaldet er endvidere stort blandt personer med anden vestlig (55,7 %) og ikke-vestlig (64,4 %) baggrund. Den logistiske regressionsanalyse bekræfter, at bortfaldet afhænger af køn ( $p < 0,0001$ ), alder ( $p < 0,0001$ ), civilstand ( $p < 0,0001$ ), etnisk baggrund ( $p < 0,0001$ ) og højeste fuldførte uddannelsesniveau ( $p < 0,0001$ ). I forhold til bortfald blev der fundet en vekselvirkning mellem køn og alder ( $p = 0,009$ ). Analysen viste, at bortfaldet er større blandt mænd end blandt kvinder i aldersgruppen 25-64 år, mens der ikke er forskel i den ældste aldersgruppe. Endvidere blev der fundet en vekselvirkning mellem køn og højeste fuldførte uddannelsesniveau ( $p = 0,03$ ). I gruppen med grundskole som højeste fuldførte uddannelsesniveau var bortfaldet det samme blandt mænd og kvinder, mens det var større blandt mænd end blandt kvinder i grupperne med gymnasial eller erhvervsfaglig uddannelse samt videregående uddannelse.

Det er endvidere undersøgt, om bortfaldsmønsteret er det samme i interviewundersøgelsen og i den selvadministrerede undersøgelse. Analysen viste, at bortfaldet er større i den selvadministrerede undersøgelse end i interviewundersøgelsen ( $p < 0,0001$ ). Der blev ikke fundet nogen vekselvirkning mellem dataindsamlingsmetode og henholdsvis køn, alder, civilstand, etnicitet og højeste fuldførte uddannelsesniveau. Der er således ikke signifikant forskel på bortfaldsmønsteret i de to undersøgelser.

**Tabel 1: Bortfald i forskellige sociodemografiske grupper opdelt på dataindsamlingsmetode. Procent, OR og antal inviterede (N).**

	Interviewundersøgelsen				Den selvadministrerede undersøgelse			
	Bortfald (%)	OR	95% CI	N	Bortfald (%)	OR	95% CI	N
<b>Køn</b>		p=0,1860				p<0,0001		
Mænd	25,9	1,20	(0,92-1,58)	742	42,5	1,40	(1,23-1,59)	2800
Kvinder	21,5	1,00		758	33,2	1,00		2702
<b>Alder</b>		p=0,2635				p<0,0001		
16-24 år	29,2			192	53,0			729
25-44 år	27,1	1,27	(0,91-1,77)	406	43,5	1,54	(1,32-1,78)	1564
45-64 år	19,4	1,00		546	31,5	1,00		2020
≥65 år	23,3	1,26	(0,88-1,82)	356	32,3	0,85	(0,71-1,01)	1189
<b>Civilstand</b>		p<0,0001				p<0,0001		
Gift	18,5	1,00		802	29,5	1,00		2.946
Ugift	32,6	2,36	(1,60-3,21)	457	50,9	1,92	(1,63-2,26)	1.639
Fraskilt	21,8	1,28	(0,80-2,02)	133	42,7	1,75	(1,44-2,14)	504
Enkestand	26,9	1,44	(0,85-2,43)	108	41,2	1,97	(1,54-2,52)	413
<b>Etnisk baggrund</b>		p<0,0001				p<0,0001		
Dansk	21,9	1,00		1.384	36,0	1,00		5.068
Anden vestlig	44,4	2,09	(0,98-4,48)	45	55,7	1,91	(1,36-2,68)	201
Ikke-vestlig	45,1	3,35	(1,88-5,97)	71	64,4	2,85	(2,04-3,98)	233
<b>Uddannelse</b>		p=0,0118				p<0,0001		
Grundskole	25,5	1,29	(0,93-1,78)	538	44,0	1,42	(1,22-1,65)	1.927
Gymnasial eller erhvervsfaglig uddannelse	21,4	1,00		593	35,9	1,00		2.171
Videregående uddannelse	19,9	1,02	(0,71-1,45)	321	27,7	0,73	(0,62-0,85)	1.196
Ingen oplysninger	56,3	3,28	(1,55-6,95)	48	62,0	1,90	(1,34-2,72)	208
<b>Total</b>	<b>23,7</b>			<b>1.500</b>	<b>37,9</b>			<b>5.502</b>

Tabel 2 viser forekomsten for centrale indikatorer i interviewundersøgelsen og i den selvadministrerede undersøgelse. Det fremgår, at svarfordelingen for de centrale indikatorer er signifikant forskellige i de to undersøgelser når det gælder spørgsmål omkring helbredsrelateret livskvalitet, sundhedsadfærd, sociale relationer og delvis

spørgsmål omkring sygelighed. Sammenlignet med den selvadministrerede undersøgelse er der i interviewundersøgelsen således en større andel, der føler sig frisk nok til at gøre, hvad man har lyst til, der er svært overvægtige, der ryger dagligt, der er fysisk aktive i fritiden, der har en langvarig sygdom, der ikke regner med at kunne få hjælp fra andre i tilfælde af sygdom, og en mindre andel der aldrig eller næsten aldrig har nogle at tale med hvis man har problemer eller brug for støtte. Derimod er der ikke forskel på forekomsten mellem de to undersøgelser når det gælder spørgsmål omkring brug af sundhedsvæsenet. Det samme mønster ses når der justeres for køn, alder, civilstand, etnisk baggrund og højeste fuldførte uddannelsesniveau.

**Tabel 2: Forekomsten af udvalgte variable opdelt på dataindsamlingsmetode. Procent og OR**

	Interview	Selvadministreret	OR (selvadministreret vs. interview)		OR* (selvadministreret vs. interview)	
			95% CI	95% CI		
<b>Helbredsrelateret livskvalitet</b>						
Føler sig frisk nok til at gøre, hvad man har lyst til	80,1	74,3	0,76	(0,64-0,90)	0,71	(0,59-0,85)
<b>Sundhedsadfærd</b>						
Svær overvægt (BMI≥30)	17,5	13,5	0,77	(0,63-0,93)	0,78	(0,64-0,95)
Daglig rygning	25,3	20,6	0,77	(0,65-0,91)	0,81	(0,68-0,96)
Moderat eller hård fysisk aktivitet i fritiden	32,3	27,2	0,80	(0,68-0,93)	0,76	(0,64-0,89)
<b>Sygelighed</b>						
Langvarig sygdom	41,2	35,0	0,79	(0,68-0,91)	0,78	(0,67-0,91)
Kronisk lungesygdom	6,7	5,0	0,71	(0,54-0,95)	0,78	(0,58-1,06)
Diskusprolaps eller andre ryg sygdomme	13,4	15,5	0,19	(0,97-1,45)	1,21	(0,99-1,48)
<b>Brug af sundhedsvæsenet</b>						
Brugt alternative behandling inden for det seneste år	28,4	28,0	1,00	(0,86-1,18)	0,96	(0,81-1,13)
Har fået målt kolesterol inden for de seneste tre år	54,8	54,4	0,97	(0,84-1,12)	0,97	(0,83-1,13)
Regelmæssig brug af tandlæge i løbet af de seneste fem år	79,3	78,2	0,96	(0,81-1,14)	0,89	(0,74-1,07)
<b>Sociale relationer</b>						
Regner ikke med at kunne få hjælp fra andre i tilfælde af sygdom	5,4	3,3	0,56	(0,41-0,77)	0,56	(0,40-0,78)
Ofte uønsket alene	3,1	4,4	1,37	(0,91-2,05)	1,40	(0,93-2,13)
Aldrig eller næsten aldrig nogle at tale med, hvis man har problemer eller brug for støtte	2,2	3,7	1,64	(1,05-2,57)	1,67	(1,06-2,62)

\*Kontrolleret for køn, alder, civilstand, etnisk baggrund og højeste fuldførte uddannelsesniveau (25 år eller derover)



## Diskussion

Ud fra denne undersøgelse er det ikke muligt at forklare årsagerne til de fundne forskelle mellem de to undersøgelsesmetoder. Resultater fra andre studier peger på, at en mulig forklaring kan være tilstedeværelsen af en interviewer. Intervieweren kan være med til at motivere de inviterede, og endvidere kan det personlige interviews interaktive form virke mere tiltalende for de inviterede frem for at skulle besvare spørgeskemaet på en mere eksamensagtig form (Mindell et al, 2010). I begge undersøgelser var bortfaldet større blandt mænd end blandt kvinder (dog ikke signifikant i interviewundersøgelsen). Tidligere undersøgelser har fundet, at bortfaldet er større blandt mænd end blandt kvinder i selvadministrerede undersøgelser, mens der ikke blev fundet forskel i bortfaldet blandt mænd og kvinder i personlige interviews (Chittleborough et al, 2010; Ekholm et al, 2010). I den selvadministrerede undersøgelse var bortfaldet relativt højt i aldersgruppen 25-44 år og relativt lavt i aldersgruppen 65 år eller derover. Dette er i overensstemmelse med en svensk undersøgelse baseret på selvadministrerede spørgeskemaer der fandt, at unge var mindre tilbøjelige til at deltage (Eaker et al, 1998). Bortfaldet var lavest blandt gift i begge undersøgelser. Dette er i overensstemmelse resultater fra andre studier (Ekholm et al, 2009; Tolonen et al, 2006). I overensstemmelse med andre studier fandt vi endvidere, at personer med anden etnisk baggrund end dansk var underrepræsenteret i begge undersøgelser (Strandhagen et al, 2010; Søgaard et al, 2004). I den selvadministrerede undersøgelse fandt vi, at personer med en lang uddannelse var mere tilbøjelige til at deltage sammenlignet med personer med en kort uddannelse, hvilket også svarer til resultater fra andre undersøgelser (Søgaard et al, 2004; Tjønneland et al, 2007).

Der blev ikke fundet signifikante forskelle mellem interviewundersøgelse og den selvadministrerede undersøgelse for indikatorer relateret til brug af sundhedsvæsenet, hvilket kan forklares ved at alle disse indikatorer er baseret på faktuelle spørgsmål med enkle svarkategorier. Der blev dog konstateret væsentlige forskelle mellem de to undersøgelser for en række andre indikatorer. Eksempelvis fandt vi en signifikant forskel i forekomsten for en indikator relateret til helbredsrelateret livskvalitet ('Hvor ofte føler du dig frisk nok til at gennemføre det, som du har lyst til at gøre?'). Her var der en større andel blandt respondenterne i interviewundersøgelsen end i den selvadministrerede undersøgelse, der rapporterede, at de for de meste følte sig frisk nok til at gennemføre det som de havde lyst til at gøre. Dette er i overensstemmelse med resultaterne fra en tidligere undersøgelse der viste, at respondenterne i en interviewundersøgelse i højere grad end respondenterne i en

selvadministreret undersøgelse vurderede deres helbred som meget godt (Tripping et al, 2010). Årsagen kan være, at den visuelle fremstilling af skalaen kan have påvirket respondenterne i den selvadministrerede undersøgelse til at vælge de mindre ekstreme kategorier. Alternativt kan respondenterne i interviewundersøgelsen have været mere tilbøjelige til at give et socialt acceptabelt svar (dvs. 'Ja, det meste af tiden') på grund af tilstedeværelsen af en interviewer. Undersøgelser har vist, at respondenter er mere tilbøjelige til at give positive vurderinger af deres helbred i interviewundersøgelser hvor der er en interviewer er til stede, sammenlignet med selvadministrerede undersøgelser (Bowling, 2005). Forekomsten af personer, der ryger dagligt, og forekomsten af svært overvægtige var højere i interviewundersøgelsen end i den selvadministrerede undersøgelse. Hvis det antages, at den højere forekomst i vores undersøgelse kan tilskrives mere ærlige svar, så er dette i modstrid med tidligere undersøgelser, der viser, at selvadministrerede undersøgelser generelt er bedre til at opnå ærlige svar når det gælder følsomme emner (Tripping et al, 2010; Bowling, 2005; Tourangeau & Smith, 1996; Beebe et al, 2005; de Leeuw, Hox & Dillmann, 2008). Alternativet er, at det er vanskeligere for respondenterne at lyve om rygevaner, højde og vægt i et personligt interview end i en selvadministreret undersøgelse. Forekomsten af personer, der dyrker hård eller moderat fysisk aktivitet i fritiden, var ligeledes højere i interviewundersøgelsen end i den selvadministrerede undersøgelse. Årsagen kan være, at respondenterne i den selvadministrerede undersøgelse undervurderer deres fysiske aktivitets niveau, når der ikke er en interviewer til stede som kan præcisere eller uddybe spørgsmålet således at respondenterne husker at indregne alle de påtænkte former for fysisk aktivitet. En anden forklaring kan være en højere grad af socialt acceptable svar blandt respondenterne i interviewundersøgelsen sammenlignet med den selvadministrerede undersøgelse. Dette kan ligeledes være tilfældet i forbindelse med spørgsmålene om sociale relationer, da stærke sociale relationer fremfor svage sociale relationer er det mest socialt acceptable. Resultaterne vedrørende de sociale relationer er dog ikke entydige. Interviewerens præcisering eller uddybning kan også være forklaringen på den højere forekomst af langvarig sygdom i interviewundersøgelsen sammenlignet med den selvadministreret undersøgelse. Dette underbygges af, at der ikke er nogen forskel i forekomsten mellem de to dataindsamlingsmetoder når det gælder specifikke sygdomme eller lidelser. Konklusionen på nærværende undersøgelse er derfor, at dataindsamlingsmetoden kan have stor betydning hvis man skal beskrive udviklingen eller sammenligne resultater der er baseret på undersøgelser, hvor der er anvendt forskellig dataindsamlingsmetode. Andre forhold såsom økonomiske og tidsmæssige ressourcer må ligeledes tages i betragtning når man skal vælge den mest hensigtsmæssige form for dataindsamling.

## Litteraturreferencer

Beebe TJ, McRae JA, Jr., Harrison PA, Davern ME, Quinlan KB. Mail surveys resulted in more reports of substance use than telephone surveys. *J.Clin.Epidemiol.* 2005;58(4):421-24.

Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J.Public Health (Oxf)* 2005;27(3):281-91.

Chittleborough CR, Taylor AW, Baum FE, Hiller JE. Non-response to a life course socioeconomic position indicator in surveillance: comparison of telephone and face-to-face modes. *BMC Med.Res.Methodol.* 2008;8:54.

Eaker S, Bergstrom R, Bergstrom A, Adami HO, Nyren O. Response rate to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. *Am.J.Epidemiol.* 1998;147(1):74-82.

Ekholm O, Hesse U, Davidsen M & Kjølner M: The study design and characteristics of the Danish national health interview surveys. *Scand J Pub Health* 2009; 37; 758-65

Ekholm O, Gundgaard J, Rasmussen NK, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. *Scand.J.Public Health* 2010;38(7):699-706.

de Leeuw ED, Hox J, Dillmann DA. *International Handbook of Survey Methodology*. New York: Taylor & Francis Group, 2008.

Jensen V, Rasmussen A. Danish education registers. *Scand.J.Public Health* 2011;39(suppl 7):91-94.

Mindell JS, Tipping S, Pickering K, Hope S, Roth MA, Erens B. The effect of survey method on survey participation: analysis of data from the Health Survey for England 2006 and the Boost Survey for London. *BMC Med.Res.Methodol.* 2010;10:83.

Pedersen C. The danish civil registration system. *Scand.J.Public Health* 2011;39(suppl 7):22-25.

Strandhagen E, Berg C, Lissner L et al. Selection bias in a population survey with registry linkage: potential effect on socioeconomic gradient in cardiovascular risk. *Eur.J.Epidemiol.* 2010;25(3):163-72.

Søgaard AJ, Selmer R, Bjertness E, Thelle D. The Oslo Health Study: The impact of self-selection in a large, population-based survey. *Int J Equity Health* 2004;6(3):3.

Tipping S, Hope S, Pickering K, Erens B, Roth MA, Mindell JS. The effect of mode and context on survey results: analysis of data from the Health Survey for England 2006 and the Boost Survey for London. *BMC.Med.Res.Methodol.* 2010;10:84.

Tjønneland A, Olsen A, Boll K et al. Study design, exposure variables, and socioeconomic determinants of participation in Diet, Cancer and Health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scand.J.Public Health* 2007;35(4):432-41.

Tolonen H, Helakorpi S, Talala K, Helasoja V, Martelin T, Prattala R. 25-year trends and socio-demographic differences in response rates: Finnish adult health behaviour survey. *Eur.J Epidemiol* 2006;21(6):409-15.

Tourangeau R, Smith TW. Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *Public Opinion Quarterly* 1996;60:275-304.