



Respons og bortfald i ældredatabasen

AGNETE ASLAUG KJÆR
MARIA HEDEMARK POULSEN
ANU SIREN

KØBENHAVN 2016

RESPONS OG BORTFALD I
ÆLDRE DATABASEN
Afdelingsleder: Kræn Blume Jensen
Afdelingen for socialpolitik og velfærd

© 2016 SFI – Det Nationale Forskningscenter for Velfærd

SFI – Det Nationale Forskningscenter for Velfærd
Herluf Trolles Gade 11
1052 København K
Tlf. 33 48 08 00
sfi@sfi.dk
www.sfi.dk

SFI's publikationer kan frit citeres med tydelig angivelse af kilden.

SFI-notater skal danne grundlag for en faglig diskussion. SFI-notater er foreløbige resultater, og læseren bør derfor være opmærksom på, at de endelige resultater og fortolkninger fra projektet vil kunne afvige fra notatet.

INDHOLD

1	INTRODUKTION	5
	Analyse af respons og bortfald i Ældredatabasen	5
	Ældredatabasens design	6
	Analysens hovedresultater	9
	Læsevejledning	11
2	DATA OG METODE	13
	Analysens fokus	13
	Analysens datagrundlag	15
	Analysens metoder	17
3	SVARPROCENTER	21
	Ældredatabasens svarprocenter	21
4	NON-RESPONS	29
	Non-respons i Ældredatabasen	29
	Sammenligning af respondenter og ikke-respondenter	29
	Hvad prædikterer non-respons?	38
	Repræsentativitet	47

5	BORTFALD	51
	Bortfald mellem Ældredatabasens bølger	51
	Hvad prædikterer aktivt bortfald?	53
	Panelets repræsentativitet	61
6	OPSUMMERING OG KONKLUSION	65
	Respons og bortfald i Ældredatabasen	65
	BILAG	69
	Bilag 1 Kodning af registerbaserede helbredsindikatorer	70
	Bilag 2 Test af forskelle	72
	Bilag 3 Udvalgte socioøkonomiske Variables fordelinger i målpopulationen	79
	Bilag 4 Gennemsnitlige marginale effekter	81
	LITTERATUR	87

INTRODUKTION

ANALYSE AF RESPONS OG BORTFALD I ÆLDRE DATABASEN

Dette SFI-notat præsenterer resultaterne af en analyse af respons og bortfald i Ældredatabasen, gennemført af SFI i Afdelingen for Socialpolitik og Velfærd. Projektet er gennemført af ph.d.-studerende Agnete Aslaug Kjær, videnskabelig assistent Maria Hedemark Poulsen og seniorforsker Anu Siren.

Ældredatabasen blev grundlagt i 1997 med det formål at skabe en dansk forløbsdatabase til samfundsvidenskabelig og humanistisk forskning på ældreområdet samt til udredningsopgaver til brug for den offentlige administration. Ældredatabasen giver mulighed for at undersøge vigtige aspekter af de danske midaldrende og ældres levevilkår, herunder tilbage-trækning fra arbejdsmarkedet, helbred og førlighed, ressourcer og plejebenhov samt sociale relationer. Siden sin opstart har Ældredatabasen fundet bred anvendelse blandt forskere fra SFI og i mange andre forskningsmiljøer i Danmark til fremme af viden og dokumentation på ældreområdet.

Ældredatabasens primære styrke består i, at den er designet som en forløbsdatabase, hvor et repræsentativt udsnit af danske ældre geninterviewes hvert femte år. Dette giver blandt andet mulighed for at følge udviklingen over tid, ligesom forløbsdata giver en række metodiske fordele i sammenligning med tværsnitsdata (se fx Hsiao, 2014, om styrkerne ved paneldata). Der er dog også særligt to udfordringer, som kan have stor betydning for repræsentativiteten i spørgeskemabaserede forløbsdata. Det gælder henholdsvis *non-respons* i enkelte interviewbølger og *bortfald* mellem gentagne interviewbølger (se fx Winkels & Withers, 2000). Man bør særligt være opmærksom på omfanget af disse, når der, som i Ældredatabasen, er tale om ældre respondenter, hvor non-respons og bortfald

kan være sammenhængende med aldersrelateret kognitiv og fysisk funktionsnedsættelse. For at forskningsmiljøer i Danmark skal kunne udnytte Ældredatabasens fulde potentiale, er det derfor nødvendigt, at der foreligger dokumentation af respons og bortfald i Ældredatabasen.

NOTATETS FORMÅL

Det er første gang i Ældredatabasens historie, at omfanget og karakteren af non-respons og bortfald i databasen dokumenteres systematisk. Notatet har til formål at besvare følgende spørgsmål:

- Hvor stort er omfanget af non-respons og bortfald i Ældredatabasens første fire bølger (1997, 2002, 2007 og 2012)?
- Er non-respons og bortfald i Ældredatabasen systematisk korelateret med henholdsvis socioøkonomisk position og helbred?
- Hvilke andre faktorer forklarer non-respons og bortfald i Ældredatabasen?

ÆLDREDATABASENS DESIGN

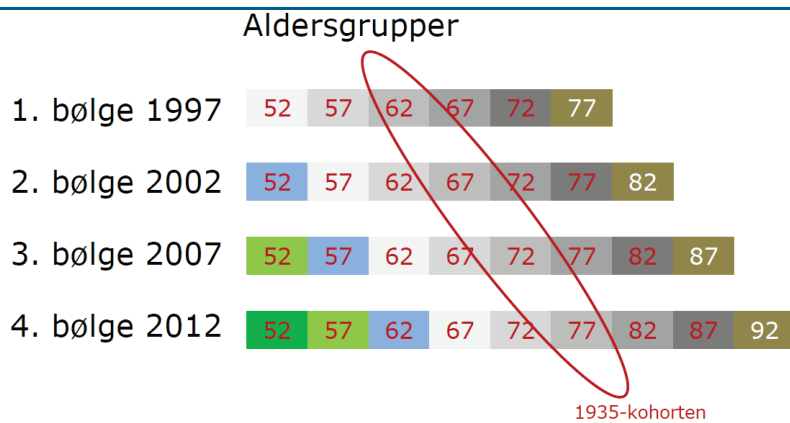
Ældredatabasen er en forløbsdatabase, der indeholder oplysninger om de samme personer over tid. Denne struktur muliggør analyser om følgende:

1. Aldersforandringer, altså forandringer, der sker over tid inden for en given fødselsårgang (fx personer født i 1935, som følges, fra de er henholdsvis 62 til 77 år gamle)
2. Kohorteforskelle, altså forskelle mellem mennesker på samme alderstrin, som er fra forskellige fødselsårgange (fx 52-årige fra førkrigsgenerationen sammenlignet med 52-årige fra efterkrigsgenerationen)
3. Tværsnitsforskelle, hvilket vil sige forskelle mellem forskellige fødselsårgange på et givent tidspunkt (fx 57-årige i 2012 sammenlignet med 92-årige i samme år).

Figur 1.1 illustrerer Ældredatabasens design med hensyn til kohorter og alder for databasens fire første bølger.

FIGUR 1.1

Aldersgrupper i Ældredatabasens fire første bølger.



Kilde: Figuren er fra slides af Kræn Blume Jensen.

Databasens udtræksmetode og design er beskrevet nærmere i Lauritzen (2012), og boks 1.1 indeholder en opsummering heraf.

BOKS 1.1

Ældredatabasens udtræksmetode og design (fra Lauritzen, 2012).

Første bølge, 1997:

I forbindelse med etableringen af Ældredatabasen i 1997 blev en stikprøve af personer født i 1920, 1925, 1930, 1935, 1940 og 1945 tilfældigt udtrukket fra CPR-registret. Hver kohorte er repræsenteret med samme vægt i stikprøven som i den danske befolkning. Bruttostikprøven bestod af 8.390 personer, hvoraf der blev opnået interview med 5.864 personer i alderen 52, 57, 62, 67, 72 og 77 år. Det svarer til en svarprocent på 70. Dog var svarprocenten noget lavere blandt de ældste kohorter. Samtlige interview blev gennemført som besøgsinterview i interviewpersonernes hjem.

Anden bølge, 2002:

I forbindelse med den anden interviewrunde i 2002 blev 4.634 personer, som havde deltaget i 1997, geninterviewet. Blandt de geninterviewede gav det en svarprocent på hele 88. En kohorte af personer født i 1950 (ny årgang 52-årige) blev tilføjet. Ligeledes blev den oprindelige nettostikprøve fra 1997 suppleret for at kompensere for bortfald i de øvrige kohorter og for at sikre et tilstrækkeligt antal personer til geninterview ved fremtidige opdateringer. Bruttostikprøven bestod af 9.836 personer, og der blev gennemført interview med 8.207 personer, hvor den ældste kohorte nu var 82 år. Det gav en samlet svarprocent på 83, og igen var svarprocenten lavest blandt de ældste kohorter. I modsætning til i 1997 blev interviewene primært gennemført som telefoninterview, og hvor dette ikke var muligt, blev besøgsinterview foretaget.

Tredje bølge, 2007:

Med henblik på at opretholde og fortsætte panelstrukturen i Ældredatabasen blev personer, der havde deltaget i de tidligere interviewrunder, forsøgt geninterviewet i 2007. Samlet set blev der opnået interview med i alt 7.010 personer fra interviewrunderne i 1997 og/eller 2002. En kohorte født i 1955 blev tilføjet for at tilføje databasen en ny årgang 52-årige, mens de øvrige kohorter i tråd med interviewrunden i 2002 blev suppleret med en ny stikprøve. Ud af en bruttostikprøve på 13.059 personer blev der gennemført 9.633 interview, hvilket svarer til en svarprocent på 74. Som ved de to forrige runder var svarprocenten lavere blandt de ældste kohorter, dvs. blandt de 82- og 87-årige. Svarprocenten var endvidere lav blandt de 52-årige. Interviewene blev som i 2002 foretaget som telefoninterview, og hvor dette ikke var muligt, blev besøgsinterview foretaget.

Fjerde bølge, 2012:

I 2012 blev den fjerde og indtil videre seneste interviewrunde gennemført. I tråd med de tidligere to interviewrunder blev en ny årgang 52-årige tilføjet ved at inkludere en kohorte født i 1960, samtidig med at de øvrige kohorter blev suppleret med en ny stikprøve. Den ældste kohorte består i denne runde af de 92-årige. På baggrund af en bruttostikprøve på 13.434 personer blev der gennemført 9.965 interview, hvilket giver en opnået svarprocent på 74. Svarprocenten var også i 2012 en smule lavere blandt de ældste som de yngste kohorter. Ud af de 9.965 interview er de 7.067 interview foretaget med personer, der ligeledes har deltaget i én eller flere af de tidligere interviewrunder. Gennemførelsen af interviewene blev som i de forrige runder foretaget over telefon, og igen hvor dette ikke var muligt, blev besøgsinterview foretaget.

Kilde: Lauritzen, 2012: 15f.

Ud af 15.973 personer, som har besvaret eller delvist besvaret spørgeskemaet til SFI's Ældre database, har 6.004 deltaget i én bølge, 4.648 har deltaget i to bølger, 2.915 har deltaget i tre bølger, og 2.406 har deltaget i alle fire bølger og udgør således Ældre databases fulde panel (tabel 1.1).

TABEL 1.1

Respondenter i én, to, tre og fire bølger. Antal

Antal bølger besvaret	Antal respondenter
En bølge	6.004
To bølger	4.648
Tre bølger	2.915
Fire bølger	2.406
Respondenter mindst én runde, total	15.973

Kilde: Egne beregninger på baggrund af Ældre databasen.

ANALYSENS HOVEDRESULTATER

Analysen af respons og bortfald i Ældre databasen giver anledning til følgende hovedresultater:

- **Svarprocenter:** Vi udregner tre forskellige typer af svarprocenter for Ældre databasen og sammenligner disse med resultater fra lignende internationale databaser. Den *simple svarprocent*, som udtrykker noget om kvaliteten af dataindsamlingen, viser, at 7 ud af 10 adspurgte deltagere i Ældre databases første bølge, mens 8 ud af 10 respondenter fastholdes fra én bølge til den næste. På dette område er Ældre databases kvalitet på niveau med lignende internationale databaser på ældreområdet, såsom det engelske English Longitudinal Study of Ageing (ELSA), det amerikanske Health and Retirement Study (HRS) og det europæiske Survey on Health, Ageing and Retirement in Europe (SHARE). *Tværsnittets og panelets svarprocenter*, som udtrykker, hvor godt det er lykkedes at fastholde den oprindelige stikprøve gennem interviewstudiets bølger, viser, at Ældre databasen har lavere svarprocenter end eksempelvis HRS. En naturlig årsag hertil er, at Ældre databasen med sine 5 år mellem hver bølge har større intervaller end lignende undersøgelser, som typisk gennemføres hvert andet år. En anden årsag er, at Ældre databases dataindsamling ikke forsøger at overtale 'nægttere' fra én bølge til at deltage i senere bølger. En sådan, om end omkostningstung, strategi kunne forbedre Ældre databases resultater på disse typer af svarprocenter.
- **Non-respons:** Vi gennemfører en sammenlignende karakteristik af respondenter og ikke-respondenter på baggrund af udvalgte registervariable. Resultaterne viser som forventet, at ikke-respondenter ad-

skiller sig fra respondenter på en række karakteristika. De ældste aldersgrupper (87- og 92-årige) og de yngste aldersgrupper (52-årige) er eksempelvis mindre tilbøjelige til at deltage, end de mellemliggende aldersgrupper (62- og 67-årige). Ligeledes er personer med kort uddannelsesbaggrund og personer i de laveste indkomstkvarterer mindre tilbøjelige til at deltage end personer med en lang uddannelse og personer i de højeste indkomstkvarterer. Disse forskelle genfinder vi i et multivariat design, hvor vi undersøger, hvad der prædikterer non-respons i Ældredatabasens fire bølger. Resultaterne viser, at modellens observerbare karakteristika kun forklarer en ganske lille del af variansen i respons/non-respons. De primære årsager til non-respons kan vi således ikke observere. Følgende observerbare faktorer er dog systematisk forbundet med en højere sandsynlighed for non-respons: Mand (vs. kvinde); Ældste og yngste aldersgrupper (vs. mellemliggende); Indvandrer (vs. dansk eller efterkommer); Laveste indkomstkvarterer (vs. højeste); Forladt arbejdsmarkedet (vs. i beskæftigelse); Lejere (vs. boligejere); Alvorlig komorbiditet (vs. ingen indlæggelser).

- **Bortfald:** Vi undersøger omfanget og karakteren af bortfald mellem Ældredatabasens bølger. Resultaterne viser, at 35 pct. af Ældredatabasens respondenter fra første bølge var afgået ved døden før kontakttidspunktet for den fjerde bølge ('passivt bortfald'). I samme 15-årige periode var 23 pct. faldet fra pga. andre årsager ('aktivt bortfald'). Ældredatabasen er designet til at tage hånd om begge typer af bortfald ved at tilføre et opfyldnings-sample i hver bølge som erstatning for bortfaldne. Systematisk bortfald fra én bølge til den næste er imidlertid stadig en udfordring, hvis man ønsker at undersøge udviklingen over tid. Vi undersøger derfor årsagerne til aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge i et multivariat design. Analysen ekskluderer passivt bortfald (det vil sige bortfald grundet dødelighed). Også her finder vi, at modellen har en lav forklaringskraft. Resultaterne viser dog, at følgende observerbare karakteristika er systematisk korreleret med en højere sandsynlighed for aktivt bortfald mellem bølger: Ældste kohorter (vs. yngste); Laveste indkomstkvarterer (vs. højeste); Lejere (vs. boligejere); Alvorlig komorbiditet (vs. ingen indlæggelser); Psykiatrisk diagnose (vs. ingen indlæggelser); samt at have flyttet adresse siden seneste bølge.
- **Repræsentativitet:** Vi sammenligner fordelingen blandt respondenter i Ældredatabasen med fordelingen i målpopulationen på en række indikatorer: alder, køn, herkomst, familietype og arbejdsmarkedstilknytning. Sammenligningen viser, at der er signifikante forskelle i fordelingen blandt henholdsvis respondenter og målpopulation for de fleste af disse indikatorer. Imidlertid er der tale om størrelsesmæssigt ganske små forskelle, hvorfor vi ikke ser grund til at være bekymrede om Ældredatabasens repræsentativitet.

Ud fra de undersøgte parametre – svarprocent, non-respons og bortfald – konkluderer vi, at Ældredatabasen har en fornuftig kvalitet som forløbsundersøgelse. Ligesom i andre survey-undersøgelser er nogle segmenter i målgruppen svære at få fat i (fx de allersvageste grupper), og dette har konsekvenser i forhold til repræsentativitet. Ikke desto mindre, er det i Ældredatabasen lykkedes at få langt de fleste adspurgte respondenter til at deltage, og typisk gentagne gange.

LÆSEVEJLEDNING

Dette notat indeholder resultaterne af en analyse af respons og bortfald i Ældredatabasen. I kapitel 2 præsenteres analysens fokus, datagrundlag og metoder. Herefter følger analysens resultater. Kapitel 3 præsenterer Ældredatabasens svarprocenter og sammenligner disse med svarprocenter i lignende databaser. Kapitel 4 indeholder en sammenlignende karakteristik af respondenter og ikke-respondenter samt en analyse af, hvilke faktorer der prædikterer non-respons. Kapitel 5 indeholder en analyse af, hvad der prædikterer bortfald mellem Ældredatabasens bølger. I kapitel 6 følger notatets konklusioner angående respons og bortfald i Ældredatabasen.

DATA OG METODE

ANALYSENS FOKUS

Et primært formål med Ældredatabasen er at undersøge og kortlægge sociale forskelle i alderdommen. Bortfald er derfor særligt problematisk, hvis det er systematisk korreleret med henholdsvis helbred eller socioøkonomisk position. I sådanne tilfælde vil bortfald udgøre en alvorlig kilde til bias. Samtidig er et vist bortfald forventeligt i Ældredatabasens stikprøvepopulation, som består af ældre danskere i alderen 52 til 92 år, hvor bortfald grundet aldersrelateret fysisk og kognitiv funktionsnedsættelse eller død er sandsynligt. I analysen fokuserer vi på 1) svarprocenter, 2) non-respons og 3) bortfald. I det følgende forklares disse begreber nærmere.

SVARPROCENTER

En svarprocent (response-rate) indikerer, hvor stor en andel af målgruppen eller af de adspurgte der har deltaget i undersøgelsen. Svarprocenten kan i paneldata udregnes på flere forskellige måder. For eksempel kan man udregne en svarprocent for hver enkelt interviewbølge, eller man kan udregne den på baggrund af en serie af bølger eller endelig på baggrund af alle bølgerne samlet (se eksempelvis Cheshire m.fl., 2011). De forskellige typer af svarprocenter kan anvendes til at sammenligne kvaliteten på tværs af databaser, som har samme målgruppe og design. Vi sammenligner i dette notat svarprocenterne fra Ældredatabasen med svarprocenter i det europæiske Survey on Health, Ageing and Retirement in Europe (SHARE), det amerikanske Health and Retirement Study (HRS) og det engelske Longitudinal Survey on Ageing (ELSA). Det skal dog bemærkes, at der er forskelle på Ældredatabasen og disse databaser,

hvad angår dataindsamlingens design og metode. En direkte sammenligning er derfor ikke mulig. Disse forskelle vil vi komme ind på løbende.

Når man udregner svarprocenter, er det vigtigt at skelne mellem personer, der er 'egnede' (eligible) henholdsvis 'uegnede' (ineligible) (se eksempelvis Cheshire m.fl., 2011). Vi betragter i dette studie personer, der er døde før kontaktdatoen¹, eller som er udvandret til et andet land siden sidste interview-bølge, som uegnede. For at være egnet til deltagelse (og dermed indgå i beregningerne) skal en person således være i live på interviewtidspunktet samt have folkeregisteradresse i Danmark.

NON-RESPONS

Non-respons beskriver det fænomen, at en person, der er udtrukket til deltagelse i undersøgelsen, af forskellige årsager ikke deltager. Til Ældre-databasens første bølge blev en repræsentativ stikprøve af ældre danskere tilfældigt udtrukket på baggrund af cpr-registret. Personer fra stikprøven, som deltager, altså fuldfører eller delvist fuldfører et interview, kaldes for 'respondenter', mens personer, der ikke deltager, kaldes for 'ikke-respondenter'. For hver af de efterfølgende bølger blev respondenter fra forrige bølge udtrukket til deltagelse igen, og der blev ligeledes tilføjet en ny fødselsårgang af 52-årige i hver bølge, ligesom der blev tilføjet en opfyldnings-stikprøve for at tage højde for tidligere bølgers non-respons. Det samlede stikprøve-udtræk kalder vi for 'bruttostikprøven'.

BORTFALD

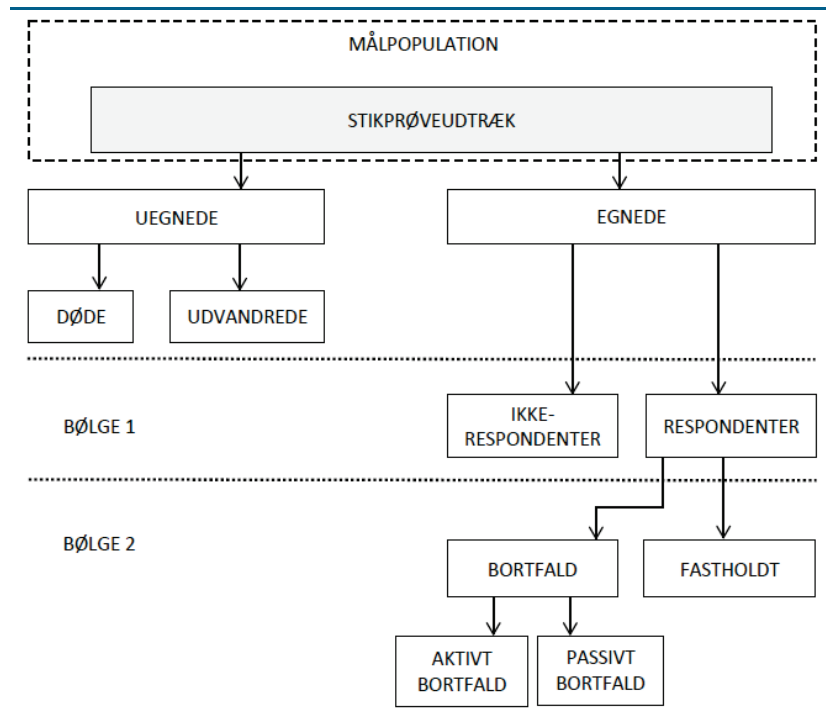
Et formål med forløbsdata er blandt andet at skabe et panel af respondenter, som deltager i alle bølger, hvorved de samme individer kan følges over tid. En udfordring opstår derfor, når personer, der deltog i panelets første bølge, falder fra i den næste bølge. Bortfald betegner netop den non-respons, som sker mellem forløbsdatabasens bølger. Eksempelvis, hvis personer, der deltog i undersøgelsens første bølge, ikke deltog igen i undersøgelsens anden bølge. En udfordring i stikprøvepopulationer bestående af ældre mennesker er det naturlige bortfald, som skyldes tidens gang. Der er tale om naturligt bortfald, hvis en respondent, der har nået en vis alder, dør. Man kan derfor, i ældre stikprøver, med fordel skelne mellem to former for bortfald: 'Passivt bortfald' opstår, når en person bortfalder fra stikprøven pga. død, og 'aktivt bortfald' opstår, når en person bortfalder fra stikprøven af andre årsager (se eksempelvis Cao & Hill, 2005).

1. Vi kender kun interview-datoen for personer, der har deltaget i undersøgelsen, derimod kender vi ikke kontakt-datoen for ikke-respondenter. For at undersøge, hvorvidt en ikke-respondent er død eller udvandret før kontakt-datoen, anvender vi derfor den sidst mulige interview-dato i hver bølge som kontakt-dato for den pågældende bølge.

Figur 2.1 illustrerer population, stikprøve, egnede, non-respons og bortfald i Ældredata-basen.

FIGUR 2.1

Illustration af population, stikprøve, egnede, non-respons og bortfald.



Kilde: Egen illustration.

ANALYSENS DATAGRUNDLAG

Analysen af respons og bortfald i Ældredata-basen er baseret på data fra to kilder, henholdsvis spørgeskemabaserede informationer fra Ældredata-basens fire første bølger (1997, 2002, 2007 og 2012) og registerbaserede oplysninger for de samme fire år.² Spørgeskemadata findes for samtlige personer, der har besvaret én eller flere af ældredata-basens fire første bølger (respondenter), mens registerdata er udtrukket for Ældredata-basens bruttostikprøve, det vil sige alle personer, der nogensinde er udtrukket til at deltage – både de, der har besvaret spørgeskemaet (respondenter), og de, der ikke har besvaret spørgeskemaet (ikke-respondenter). Da-

2. I 2007-bølgen blev interviewene gennemført i både 2007 og 2008. Vi anvender i dette år som udgangspunkt registerdata fra 2007, men opdaterer eventuelle missings med data fra 2008. I 2012-bølgen blev interviewene gennemført i både 2012 og 2013. Vi anvender i dette år som udgangspunkt registerdata fra 2012, men opdaterer eventuelle missings med data fra 2013.

ta fra registrene linkes til data fra Ældredatabasen på individniveau, og undersøgelsen er i den forbindelse anmeldt til Datatilsynet.

REGISTERDATA OG VARIABLE

TABEL 2.1

Anvendte variable og deres registre samt hvilken årrække, de er til rådighed i analysen.

Variabel	Register	År
<i>Sociodemografi</i>		
Køn	Registerbaserede arbejdsstyrke statistik (RAS)	1997, 2002, 2007, 2012
Alder	Befolkningen (BEF)	1997, 2002, 2007, 2012
Etnisk herkomst	Indvandrere og Efterkommere (IEPE)	1997, 2002, 2007, 2012
Civilstatus	Befolkningen (BEF)	1997, 2002, 2007, 2012
Har en partner	Befolkningen (BEF)	1997, 2002, 2007, 2012
Partner er død	Befolkningen (BEF), Døde i Danmark (DOD)	1997, 2002, 2007, 2012
Flyttet	Befolkningen (BEF)	1997, 2002, 2007, 2012
Egen bolig	Indkomst (IND)	1997, 2002, 2007, 2012
<i>Socioøkonomisk position</i>		
Disponibel indkomst	Indkomst (IND)	1997, 2002, 2007, 2012
Ækvivaleret familie indkomst ³	Indkomst (IND)	1997, 2002, 2007, 2012
Modtager førtidspension og folkepension	Indkomst (IND)	1997, 2002, 2007, 2012
Uddannelse	Uddannelser (BUE) (UDDA)	1997, 2002, 2007, 2012
Arbejdsmarkedsstatus	Registerbaserede arbejdsstyrke statistik (RAS)	1997, 2002, 2007, 2012
<i>Helbred</i>		
Charlson Comorbidity Indeks (seneste 5 år)	Landspatientregistret diagnoser (LPR-DIAG)	1997, 2002, 2007, 2012
Psykisk diagnose	Landspatientregistret diagnoser (LPR-DIAG)	1997, 2002, 2007, 2012
Patienttype	Landspatientregistret diagnoser (LPR-DIAG)	1997, 2002, 2007, 2012
Antal kontakter til læge	Sygesikringsregistret (SSSI) (SSSY)	1997, 2002, 2007, 2012

En analyse af non-respons og bortfald er per definition en analyse af personer, som ikke har besvaret Ældredatabasens spørgeskema. Vi må derfor benytte registerdata, når vi skal forsøge at opklare, hvad der karakteriserer ikke-respondenter, og hvorvidt disse adskiller sig systematisk fra respondenter. De faktorer og variable, som er i fokus i analysen, er beskrevet i tabel 2.1. Vi inddrager i denne analyse sociodemografiske faktorer som alder, køn, etnisk herkomst, civilstatus og enkestatus samt socioøkonomiske faktorer som indkomst, uddannelse, sociale ydelser og arbejdsmarkedsstatus. For også at afdække helbred inddrager vi desuden en række registerbaserede indikatorer angående forbrug af sundhedsydelse.

3. De disponible indkomster er korrigeret med ækvivalensfaktoren, der grundlæggende er et mål for, hvor mange "enlige" voksne en familie indkomstmæssigt svarer til. Ækvivalensfaktoren beregnes som $(\text{antal voksne} + \text{antal børn})^{0,6}$, som svarer til den ækvivaleringsskala, som Finansministeriet benytter (jf. Arbejderbevægelsens Erhvervsråd, 2011). Herefter er indkomsten korrigeret for inflation.

ser og diagnoser stillet i forbindelse med hospitalsindlæggelser (Bilagstabel B1.1 indeholder en oversigt over kodningen af de registerbaserede helbredsindikatorer). Vi benytter konkret registerdata fra den Registerbaserede Arbejdsstyrke Statistik (RAS), Befolkningen (BEF), Indvandrere og Efterkommere (IEPE), Døde i Danmark (DOD), Indkomst (IND), Uddannelser (BUE), Landspatientregistret Diagnoser (LPRDIAG) og Sygesikringsregistret (SSSI/SSSY).

ANALYSENS METODER

Analysens primære formål er at dokumentere to udfordringer, som er typiske for forløbsdatabaser, henholdsvis non-respons og bortfald, samt hvorvidt disse er systematisk korelateret med socioøkonomisk status og helbred.

BEREGNING AF SVARPROCENTER

Der er tre hovedtyper af svarprocenter, som vi i dette notat vil beskæftige os med: simple svarprocenter (i litteraturen *field response-rates*), tværsnitets svarprocenter (*cross-sectional response-rates*) og panelets svarprocenter (*longitudinal response-rates*). Disse typer af svarprocenter kan desuden enten udregnes betinget af tidligere deltagelse (*conditional*) eller ubetinget (*unconditional*). Det er disse forskellige typer af svarprocenter, som beregnes i analysen og rapporteres senere i notatet.

TABEL 2.2

Typer af svarprocenter og deres beregningsmetode: simpel svarprocent, tværsnitets svarprocent og panelets svarprocent.

Svarprocent	Beskrivelse	Tæller	Nævner
SSP _i	Simpel ubetinget svarprocent blandt udtrukne i bølge <i>i</i>	Deltagelse i bølge <i>i</i> (<i>i=1,...,k</i>)	Udtrukket og egnet i bølge <i>i</i> (<i>i=1,...,k</i>)
SSP _{ii-1}	Simpel betinget svarprocent blandt udtrukne, givet deltagelse i den tidligere bølge (<i>i-1</i>)	Deltagelse i bølge <i>i</i> (<i>i=1,...,k</i>)	Udtrukket og egnet i bølge <i>i</i> (<i>i=1,...,k</i>) og deltagelse i bølge <i>i-1</i>
TSP _i	Ubetinget tværsnit-svarprocent i bølge <i>i</i>	Deltagelse i bølge <i>i</i> (<i>i=1,...,k</i>)	Egnet i bølge <i>i</i> (<i>i=1,...,k</i>)
TSP _{ii-1}	Betinget tværsnit-svarprocent i bølge <i>i</i> , givet deltagelse i den tidligere bølge (<i>i-1</i>)	Deltagelse i bølge <i>i</i> (<i>i=1,...,k</i>)	Egnet i bølge <i>i</i> (<i>i=1,...,k</i>) og deltagelse i bølge <i>i-1</i>
PSP _{1,2,...,k}	Ubetinget panelsvarprocent for en kombination af bølger <i>1, 2, ..., k</i>	Deltagelse i en serie af bølger (<i>1, 2, ..., k</i>)	Egnet i en serie af bølger (<i>1, 2, ..., k</i>)
PSP _{{1,2,...,k} 1}	Betinget panelsvarprocent for en kombination af bølger <i>1, 2, ..., k</i> , givet deltagelse i bølge 1	Deltagelse i en serie af bølger (<i>1, 2, ..., k</i>)	Egnet i en serie af bølger (<i>1, 2, ..., k</i>) og deltagelse i bølge <i>i-1</i>

Anm.: Egnede betegner i denne analyse personer, som ikke er døde før den senest mulige interviewdato, og som ikke er udvandret til udlandet.

Kilde: Inspireret af Cheshire m.fl., 2011 (oprindeligt Lynn, 2005).

Tabel 2.2 viser, hvordan de tre typer af svarprocenter beregnes (inspireret af Cheshire m.fl., 2011; oprindeligt Lynn, 2005).

SIMPLE SVARPROCENTER (SSP)

Den simple svarprocent (SSP_i) beregnes i dette notat ved andelen af personer, der har deltaget i én bestemt interviewbølge (i) ud af alle personer, der var udtrukket til netop denne bølge, og som desuden var egnede på interviewtidspunktet (dvs. ikke døde eller udvandret før interviewets kontaktdato). Den simple betingede svarprocent ($SSP_{i|i-1}$) findes ved andelen af personer, der deltog i én bestemt interviewbølge (i) ud af alle udtrukne og egnede til denne bølge, som også deltog i den tidligere interviewbølge ($i-1$). Den simple svarprocent kan også kaldes felt-svarprocenten (*field response rate*), idet nævneren udelukkende inkluderer personer, som var udtrukket til den pågældende runde. På dette punkt adskiller denne type svarprocent sig fra de to øvrige typer af svarprocenter.

TVÆRSNITTETS SVARPROCENTER (TSP)

Tværsnittets svarprocenter adskiller sig fra den simple svarprocent, idet nævneren udelukkende omfatter alle, der var udtrukket til Ældredatabasens første bølge, uanset om de blev udtrukket igen til senere deltagelse. Den vil derfor per definition være noget lavere end den simple svarprocent. En ubetinget tværsnits-svarprocent (TSP_i) beregnes i dette notat ved andelen af personer, der har deltaget i én bestemt interviewbølge (i) ud af alle dem, der var egnede i den interviewrunde (dvs. ikke døde eller udvandret i den pågældende runde), og som desuden var udtrukket til den første interviewrunde. Man kan altså udregne en svarprocent for hver enkelt interviewrunde. En betinget tværsnits-svarprocent ($TSP_{i|i-1}$) findes ved andelen af personer, der deltog i én bestemt interviewrunde (i), ud af alle egnede, som også deltog i den tidligere interviewbølge ($i-1$). Sidstnævnte giver et overblik over, hvor godt respondenter fastholdes fra én interviewrunde til den næste interviewrunde (jf. Cheshire m.fl., 2011).

PANELETS SVARPROCENTER (PSP)

En ubetinget panelsvarprocent ($PSP_{\{1,2,\dots,k\}}$) findes ved andelen af personer, der deltager i en serie af interviewbølger (eks. 1, 2, 3 og 4) ud af alle personer, der er egnede i alle disse bølger (dvs. ikke døde eller udvandret i nogle af de efterfølgende bølger). Denne procent indikerer, hvor stor en andel af den oprindelige stikprøve der er repræsenteret i en sekvens af interviewbølger. En betinget panelsvarprocent ($SP_{\{1,2,\dots,k\}|1}$) udregnes ved andelen af personer, der deltager i en serie af interviewbølger ud af dem, der var egnede i disse bølger, og som desuden deltog i den første interviewrunde. Denne type frafaldsprocent giver et overblik over, hvor godt det lykkes at fastholde respondenter fra det oprindelige panel af deltagere i de efterfølgende interviewrunder. Ligesom i tværsnittets svarprocenter omfatter nævneren her alle, der var udtrukket til Ældredatabasens første bølge, uanset om de blev udtrukket til deltagelse senere (jf. Cheshire m.fl., 2011).

DESKRIPTIV STATISTIK

Vi foretager i dette notat en sammenlignende karakteristik af respondenter og ikke-respondenter ud fra udvalgte registerbaserede statistiske indikatorer. For at teste forskelle på fordelinger på disse variable mellem de to grupper af henholdsvis respondenter og ikke-respondenter anvender vi henholdsvis χ^2 -test og pr-test.⁴

MULTIVARIAT ANALYSE

Vi foretager desuden en analyse af, hvad der prædikterer henholdsvis non-respons og bortfald i Ældredatabasen. Til dette formål gennemfører vi en række multivariate analyser. Analyserne af årsager til non-respons og bortfald falder i to dele. Først undersøges den afhængige variabel non-respons for alle fire bølger separat. Vi måler non-respons for hver enkelt bølge af egnede til deltagelse, hvor 1 indikerer non-respons, og 0 indikerer respons. Dernæst undersøges Ældredatabasens bortfald. Vi undersøger bortfaldet mellem to bølger i Ældredatabasen, nemlig den tredje og fjerde bølge. Bortfald måles gennem en dikotom variabel, hvor 1 indikerer bortfald, og 0 indikerer fastholdelse.

Vi anvender i begge analyser multivariat logistisk regression, som netop antager en binær afhængig variabel.⁵ Koefficienterne repræsenterer derved sandsynligheden for henholdsvis non-respons eller bortfald. Da de umiddelbare koefficienter fra en logistisk regression kan være vanskelige at fortolke, afrapporterer vi i stedet de estimerede sandsynligheder i form af de uafhængige variables gennemsnitlige marginale effekter på den estimerede sandsynlighed for henholdsvis non-respons eller bortfald.

4. Både χ^2 -test og pr-test bruges til at teste forskellene i andele, altså to eller flere gruppers fordeling på kategoriske variable, fx højeste fuldførte uddannelse. Til at teste om to gruppers fordeling på en dikotom variabel er signifikant forskellig har vi benyttet pr-test. Til at teste om to grupper eller fleres fordeling på en kategorisk variabel er signifikant forskellig har vi benyttet χ^2 -test.

5. Resultaterne estimeres ved hjælp af en logit transformation og en generaliseret lineær model.

SVARPROCENTER

ÆLDREDATABASENS SVARPROCENTER

I de følgende afsnit præsenteres Ældredatabasens svarprocenter. Som beskrevet i kapitel 1 beregner vi tre typer af svarprocenter, som gør det muligt at sammenligne databasens kvalitet med lignende internationale databaser.

SIMPLE SVARPROCENTER (SSP)

Den første type svarprocent, som beregnes, er den simple svarprocent (SSP), som baserer sig på personer, der er udtrukket til deltagelse i den pågældende bølge, og som tilmed er 'egnede' til deltagelse på kontakttidspunktet.

Tabel 3.1 danner først et overblik over, hvor mange personer der er blevet tilfældigt udtrukket til hver af ældredatabasens fire bølger, og hvor mange af disse, der er egnede til deltagelse. Udtræk til Ældredatabasen er baseret på CPR-registret og tager således i udgangspunktet højde for eventuel migration eller dødsfald i populationen. Imidlertid vil en mindre andel af udtrækket alligevel bestå af personer, der er udvandret til udlandet, eller som er afgået ved døden på kontakttidspunktet. Disse personer er således uegnede til deltagelse fra start og bør ikke indgå i beregninger af svarprocenter. Tabellen viser derfor ligeledes antallet af egnede (dvs. i live på kontakttidspunktet og har folkeregisteradresse i Danmark). Det fremgår af tabel 3.1, at udtrækket øges bølge for bølge, idet der tilføjes en ny fødselsårgang af 52-årige i hver bølge, ligesom bortfald erstattes med en opfrisknings-stikprøve. Det fremgår ligeledes, at der i hver bølge er en mindre andel af de udtrukne, som reelt ikke er egnede til deltagelse, fordi de enten er udvandret til udlandet siden sidste

interviewbølge, eller fordi de er døde før kontakttidspunktet. I første bølge omfatter udtrækket således 8.390 personer, hvoraf 8.340 reelt er egnede til deltagelse og dermed indgår i vores efterfølgende beregninger af svarprocenter. I anden bølge omfatter udtrækket 9.863 personer, heraf 9.820 egnede. I tredje bølge er 13.150 personer udtrukket, heraf er 12.936 egnede, og i fjerde bølge er 13.558 personer udtrukket, og heraf er 13.337 egnede.

TABEL 3.1

Udtræk og 'egnede', fordelt på Ældredatabasens fire bølger. Antal.

	Bølge 1	Bølge 2	Bølge 3	Bølge 4
Udtræk	8.390	9.863	13.150	13.558
Egnede	8.340	9.820	12.936	13.337
Respondenter	5.864	8.207	9.633	9.965

Anm.: 'Egnede' omfatter personer, der er udtrukket til den pågældende bølge i Ældredatabasen, og som er i live på kontakttidspunktet og ikke er udvandret til et andet land siden Ældredatabasens seneste interview bølge.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og dødsdato fra DODE-registret samt migrationsdata fra VNDS-registret.

Den simple ubetingede svarprocent (SSP_i) udregnes herefter for hver af Ældredatabasens fire bølger (tabel 3.2). Tælleren omfatter bølgens respondenter, mens nævneren omfatter alle, der var udtrukket til den pågældende bølge, og som desuden var egnede på kontakttidspunktet. Den simple ubetingede svarprocent giver således et overblik over dataindsamlingens succesrate i hver enkelt runde, idet den viser, hvor stor en andel af de kontaktede der deltog, samtidig med at der tages højde for, at ikke alle var egnede til deltagelse. Det fremgår af tabel 3.2, at svarprocenten blandt egnede var 70 pct. i Ældredatabasens første bølge, 84 pct. i anden bølge, og 75 pct. i henholdsvis tredje og fjerde bølge.

TABEL 3.2

Simpel ubetinget svarprocent (SSP_i) for Ældredatabasens fire bølger. Procent.

Ubetinget	SSP_1	SSP_2	SSP_3	SSP_4
	70,3	83,6	74,5	74,7

Anm.: Nævneren omfatter personer, der er udtrukket til deltagelse i den pågældende bølge, og som desuden er egnede til deltagelse, dvs. i live på kontakttidspunktet og ikke udvandret til et andet land siden Ældredatabasens seneste interviewbølge.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og dødsdato fra DODE-registret samt migrationsdata fra VNDS-registret.

Med en simpel svarprocent i første bølge på 70 pct. er Ældredatabasens kvalitet i sit udgangspunkt på niveau med lignende internationale databaser på ældreområdet. Eksempelvis er Ældredatabasens svarprocent på niveau med den engelske ækvivalent English Longitudinal Study on Ageing (ELSA), som har en individuel svarprocent på 67 pct. i første bølge, som blev gennemført i 2002. Ældredatabasens svarprocent ligger samti-

dig markant højere end i den europæiske Survey on Health Ageing and Retirement in Europe (SHARE), hvor svarprocenten på tværs af lande ligger på 62 pct. og for Danmark specifikt på 63 pct. i første bølge, som blev gennemført i 2004. Dog er Ældredatabasens svarprocent lavere end i det amerikanske Health and Retirement Study, som har en svarprocent på 82 pct. i første bølge i 1992 (De Lucca & Paracchi, 2005; Sonnega m.fl., 2014; Steptoe m.fl., 2012).

Dataindsamlingens interviewere er blevet bedt om at notere en årsag til non-respons ved kontakter, som ikke resulterede i et interview. Tabel 3.3 viser resultatet af den første kontakt til hver person i udtrukket til Ældredatabasens fire bølger blandt egnede ikke-respondenter. Det fremgår af tabel 3.3, at den hyppigste årsag til non-respons er, at respondenter ikke ønsker at deltage ('nægtere'). Nægtere udgør 83 pct. af ikke-respondenter i Ældredatabasens første bølge. Den næst-hyppigste årsag er sygdom og handicap. Andelen, der angiver sygdom eller handicap som årsag til non-respons, stiger efter første bølge, i takt med at personer, som indgår i stikprøven, ældes. Dette kan give anledning til bekymring angående databasens repræsentativitet, såfremt Ældredatabasens respondenter består af en selekteret gruppe med et bedre helbred end gennemsnittet. Dette undersøger vi nærmere i kapitel 4 om årsager til non-respons. Sprogvanskeligheder optræder først som en hyppig årsag til non-respons fra og med tredje bølge, hvilket skyldes, at personer med ikke-vestlig baggrund samples med dobbelt sandsynlighed fra 2007.

TABEL 3.3

Årsager til non-respons blandt egnede, fordelt på Ældredatabasens fire første bølger. Antal og procent.

	Bølge 1, 1997		Bølge 2, 2002		Bølge 3, 2007		Bølge 4, 2012	
	Antal	Pct.	Antal	Pct.	Antal	Pct.	Antal	Pct.
Nægter	2.052	82,9	1.107	68,7	2.086	63,2	2.283	67,7
Sygdom eller handicap	178	7,3	214	13,2	292	8,8	390	11,6
Ikke truffet	158	6,4	208	12,9	564	17,1	343	10,2
Bortrejst, på hospitalet	45	1,8	53	3,3	85	2,6	81	2,4
Flyttet	18	0,7	13	0,8	30	0,9	12	0,4
Andet/sproglige vanskeligheder	25	1	18	1,1	246	7,4	263	7,8
I alt	2.476	100	1.613	100	3.303	100	3.372	100

Anm.: Fordelingen omfatter ikke-respondenter, der er udtrukket og egnede til deltagelse i den pågældende bølge.

Kilde: Egne beregninger på baggrund af variabel 'resultat for kontakt', som leveredes af SFI-Survey i forbindelse med dataindsamlingens afslutning.

Et formål med Ældredatabasen er at danne et forløbsdatasæt, hvorfor det er vigtigt at fastholde respondenter mellem undersøgelsens bølger. Den simple betingede svarprocent ($SSP_{1|i-1}$) udregnes derfor for hver af Ældredatabasens fire bølger, idet denne giver et overblik over, hvor godt respondenter fastholdes fra én bølge til den næste. Tælleren inkluderer i hver bølge individer, som deltog i den pågældende runde og i den foregå-

ende runde, mens nævneren omfatter udtrukne egnede, der besvarede den tidligere runde. Den betingede simple svarprocent kan således også kaldes fastholdelsesprocenten, idet den udtrykker, hvor godt respondenter fastholdes fra én runde til den næste. Det fremgår af tabel 3.4, at svarprocenten i Ældredatabasens anden bølge er hele 89 pct. blandt personer, der også besvarede første bølge, og som fortsat var egnede til deltagelse. I tredje bølge var svarprocenten 88 pct. blandt personer, der også besvarede anden bølge, og i fjerde bølge var svarprocenten 82 blandt personer, der også besvarede tredje bølge. Personer, der først én gang har sagt ja til at deltage, har altså stor sandsynlighed for også at deltage i den næste bølge. Et resultat heraf er, at Ældredatabasen med tiden vil bestå af et mere og mere selekteret udsnit af respondenter, der er villige til at deltage.

TABEL 3.4

Simpel betinget svarprocent ($SSP_{i|i-1}$) for Ældredatabasens fire bølger. Procent.

Betinget	$SSP_{2 1}$	$SSP_{3 2}$	$SSP_{4 3}$
	88,5	88,3	82,1

Anm.: Nævneren omfatter personer, der er udtrukket til deltagelse i den pågældende bølge, og som desuden er egnede til deltagelse, dvs. i live på kontakttidspunktet og ikke udvandret til et andet land siden Ældredatabasens seneste interviewbølge. Nævneren er desuden betinget til respondenter, der besvarede den tidligere bølge.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og dødsdato fra DODE-registret samt migrationsdata fra VNDS-registret.

TVÆRSNITTETS SVARPROCENTER (TSP)

I dette og i det efterfølgende afsnit beregner vi henholdsvis tværsnittets og panelets svarprocenter for Ældredatabasens fire bølger. Disse typer af svarprocenter baserer sig på Cheshire m.fl. (2011)'s metode (som beskrevet i kapitel 1). Formålet med at udregne disse svarprocenter er at kunne sammenligne Ældredatabasens kvalitet med lignende internationale databaser på trods af forskelle i databasernes design. Som sammenligningsgrundlag anvender vi svarprocenter fra det engelske English Longitudinal Survey on Ageing (ELSA) og fra det amerikanske Health and Retirement Study (HRS) (Cheshire m.fl., 2011). For bedst muligt at kunne sammenligne svarprocenterne med Cheshire m.fl. (2011) opgøres disse udelukkende for personer, der er en del af Ældredatabasens oprindelige sample, bestående af seks fødselskohorter født 1920, 1925, 1930, 1935, 1940 og 1945. Vi begrænser, ligesom tidligere, analysen til personer, som er egnede på kontakttidspunktet for den pågældende bølge (dvs. ikke død eller udvandret på kontakttidspunktet), men følger derudover Cheshire m.fl. (2011)'s metodik, idet vi inkluderer alle personer, der var udtrukket til Ældredatabasens første bølge, uanset om de blev udtrukket og kontaktet igen til næste bølge. På dette punkt adskiller tværsnittets og panelets svarprocenter sig således fra den simple svarprocent.

Det er dog vigtigt at bemærke, at Ældredata-basen ikke er direkte sammenlignelig med HRS og ELSA. En væsentlig forskel er, at Ældredata-basen er designet med 5 år mellem hver bølge, mens der blot er 2 år mellem hver bølge i den engelske og amerikanske database. Derfor må man alt andet lige forvente lavere svarprocent i Ældredata-basen over tid, idet der i Ældredata-basen simpelthen er længere tid til at falde fra mellem bølgerne, hvor respondenterne bliver ældre.

Tabel 3.5 viser først og fremmest tværsnittets ubetingede svarprocent for hver af Ældredata-basens fire bølger (TSP_i). De ubetingede tvær-snitts-svarprocenter angiver svarprocenten blandt egnede i hver enkelt bølge for det oprindelige udtræk. Den ubetingede svarprocent opgøres for alle personer, der er en del af ældredata-basens oprindelige udtræk (seks fødselskohorter), som fortsat er 'egnede' i den undersøgte bølge, uanset om de var udtrukket hertil eller ej. Første bølges svarprocent er på 70 pct. ligesom den simple svarprocent, men da ikke-responderer *ikke* følges op i efterfølgende bølger, er tværsnittets svarprocent nødvendigvis faldende bølge for bølge, med 63 pct. i anden bølge, 55 pct. i tredje bølge og 47 pct. i fjerde bølge. Det svarer til, at omkring halvdelen af den oprindelige stikprøve, som fortsat var egnede i fjerde bølge, deltog i undersøgelsen i 2012 (15 år efter at de første gang deltog i Ældredata-basen).

TABEL 3.5

Tværsnittets ubetingede svarprocenter (TSP_i) for Ældredata-basens fire bølger. Procent.

Ubetinget	TSP ₁	TSP ₂	TSP ₃	TSP ₄
	70,3	62,5	55,3	47,0

Anm.: Tværsnittets svarprocenter udregnes udelukkende for de seks oprindelige fødselsår-gange, som er samlet til Ældredata-basens første runde. I hver bølges svarprocent indgår i nævneren alle egnede, som blev udtrukket til deltagelse i Ældredata-basens første bølge – også personer, der *ikke* blev samlet til deltagelse i den pågældende bølge.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredata-basen og dødsdato fra DODE-registret samt migrationsdata fra VNDS-registret.

Det må siges at være et udmærket resultat, når man tager respondenternes alder i betragtning. Sammenligner vi dette resultat med lignende internationale databaser om ældre, må de ubetingede tvær-snitts-svarprocenter siges at ligge på niveau. I det engelske Longitudinal Survey on Ageing (ELSA) ligger de ubetingede svarprocenter således noget lavere end i Ældredata-basen, og i det amerikanske Health and Retirement Study (HRS) ligger de noget højere (Cheshire m.fl., 2011).⁶ Det er dog klart, at Ældredata-basens panel ville have gavn af, at respondenter, der nægter deltagelse i én runde, blev fulgt op i senere runder. For at imødekomme

6. De ubetingede tvær-snitts-svarprocenter er hhv. 70,2, 46,5, 38,2 og 36,1 i de fire første bølger af ELSA, som gennemføres hvert andet år, og 78,0, 72,5, 72,0 og 70,3 i de fire første bølger af HRS, som ligeledes gennemføres hvert andet år (Cheshire m.fl., 2011).

denne udfordring anvender Ældredatabasens design i stedet et opfyldnings-sample, hvor bortfaldne fra én runde erstattes med et nyt sample i den efterfølgende runde.

Tabel 3.6 viser hernæst tværsnittets betingede svarprocenter, som baserer sig på respondenter, der deltog i den foregående bølge ($TSP_{i|i-1}$). Tælleren inkluderer i hver bølge individer, som deltog i den pågældende runde og i den foregående runde, og nævneren inkluderer personer, der deltog i den foregående bølge, og som fortsat var egnede i den pågældende bølge. De betingede tværsnits-svarprocenter angiver således svarprocenten blandt egnede i hver enkelt bølge for personer, der også har deltaget i den foregående bølge – blandt Ældredatabasens oprindelige udtræk. Den siger således, ligesom den betingede simple svarprocent, noget om fastholdelse mellem bølger. Den er på 88 pct. i anden bølge, 84 pct. i tredje bølge og 83 pct. i fjerde bølge (betinget på deltagelse i den foregående bølge). Den betingede svarprocent synes at være næsten stabil over tid. Det skyldes, at det sample, der deltog i den forrige bølge, er selektivt og har større sandsynlighed for at vedblive med at deltage. Ældredatabasens betingede svarprocenter ligger på et pænt niveau sammenlignet med ELSA og næsten på niveau med HRS⁷.

TABEL 3.6

Tværsnittets betingede svarprocenter ($TSP_{i|i-1}$) for Ældredatabasens fire bølger. Procent.

Betinget	$TSP_{2 1}$	$TSP_{3 2}$	$TSP_{4 3}$
	87,5	83,9	82,6

Anm.: Tværsnittets svarprocenter udregnes udelukkende for de seks oprindelige fødselsårsgange, som er samlet til Ældredatabasens første runde. I hver bølges svarprocent indgår i nævneren alle egnede, som blev udtrukket til deltagelse i Ældredatabasens første bølge – også personer, der *ikke* blev samlet til deltagelse i den pågældende bølge.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og dødsdato fra DODE-registret samt migrationsdata fra VNDS-registret.

PANELETS SVARPROCENTER (PSP)

Panelets ubetingede svarprocent ($PSP_{i,..k}$) angiver svarprocenten for en serie af flere bølger i Ældredatabasen for personer, der er egnede i alle disse bølger. Det fremgår af tabel 3.7, at svarprocenten i første og anden bølge er 62 pct. blandt egnede i disse to bølger. Svarprocenten falder for hver bølge, der tilføjes. Svarprocenten markeret med fed repræsenterer den ”komplette” svarprocent og indikerer, at 47 pct. af alle personer, der blev udtrukket til første bølge, og som var egnede i alle efterfølgende bølger, deltog i alle Ældredatabasens fire bølger. Som beskrevet tidligere,

7. De betingede tværsnits-svarprocenter er hhv. 64,7, 81,5 og 85,6 i anden, tredje og fjerde bølge af ELSA (betinget på deltagelse i den foregående bølge) og 92,6, 94,1 og 94,5 i HRS (Cheshire m.fl., 2011). Bemærk dog, at der er forskelle i definitionen af egnede, som vanskeliggør direkte sammenligning.

er Ældredatabasen på dette punkt *ikke* direkte sammenlignelig med ELSA og HRS, men panelsvarprocenterne i Ældredatabasen ligger på niveau med disse, når man tager i betragtning, at der er flere år imellem Ældredatabasens bølger.⁸

TABEL 3.7

Panelets ubetingede svarprocent ($PSP_{j,\dots,k}$) for Ældredatabasens fire bølger. Procent.

Ubetinget	$PSP_{1,2}$	$PSP_{1,2,3}$	$PSP_{1,2,3,4}$
	62,40	54,1	46,5

Anm.: Panelets svarprocenter udregnes udelukkende for de seks oprindelige fødselsårgange, som er samlet til Ældredatabasens første runde. I hver bølges svarprocent indgår i nævneren alle egnede, også personer, der ikke blev samlet til deltagelse.

Tabel 3.8 præsenterer panelets betingede svarprocenter ($SP_{i,\dots,k|1}$). Disse svarprocenter angiver svarprocenten for en serie af flere bølger i Ældredatabasen for personer, der er egnede og desuden har deltaget i undersøgelsens første bølge. Det fremgår, at svarprocenterne ligger på hhv. 88 pct., 75 pct. og 63 pct. Dette er noget lavere end i både ELSA og HRS⁹, hvilket kan tilskrives, at der er flere år mellem runderne i Ældredatabasen. Per definition er panelets betingede svarprocent for bølge 1 og 2 magen til tværsnittets betingede svarprocent for bølge 2.

TABEL 3.8

Panelets betingede svarprocent for Ældredatabasens fire bølger. Procent.

Betinget	$SP_{1,2 1}$	$SP_{1,2,3 1}$	$SP_{1,2,3,4 1}$
	87,5	74,5	63,3

Anm.: Panelets svarprocenter udregnes udelukkende for de seks oprindelige fødselsårgange, som er samlet til Ældredatabasens første runde. I hver bølges svarprocent indgår i nævneren alle egnede, også personer, der ikke blev samlet til deltagelse.

8. I ELSA og HRS er de ubetingede longitudinale svarprocenter hhv. 45,8, 38,6 og 34,6 pct. for ELSA og 70,4, 67,7 og 64,3 pct. for HRS (for $SP_{1,2}$, $SP_{1,2,3}$, $SP_{1,2,3,4}$).

9. ELSA og HRS er de betingede longitudinale svarprocenter hhv. 81,5, 70,8 pct. for ELSA og 92,6, 87,6 og 83,3 pct. for HRS (for $SP_{1,2|1}$, $SP_{1,2,3|1}$, $SP_{1,2,3,4|1}$).

NON-RESPONS

NON-RESPONS I ÆLDREDATABASEN

I de følgende afsnit præsenteres resultaterne af analysen af non-respons i Ældredatabasen. Først sammenlignes respondenter og ikke-respondenter ud fra en række registerbaserede karakteristika, for at undersøge, hvorvidt der er systematiske forskelle på de to grupper. Dernæst følger en multivariat analyse, der undersøger, hvorvidt socioøkonomisk status eller helbred selvstændigt prædikerer non-respons i databasen.

SAMMENLIGNING AF RESPONDENTER OG IKKE-RESPONDENTER

Et formål med analysen af respons i Ældredatabasen er at undersøge, om ikke-respondenter adskiller sig systematisk fra respondenter. I det følgende karakteriseres disse to grupper ud fra udvalgte registerbaserede indikatorer (den fulde karakteristik kan findes i bilagstabel B2.1/bilagstabel B2.1). Analyserne omfatter, for hver bølge, personer, som er udtrukket og desuden egnede (ikke død eller udvandret).

SOCIODEMOGRAFI

Figur 4.1 viser respondenter og ikke-respondenter fordelt på fødselsårsgang, særligt for hver af Ældredatabasens fire bølger. Det fremgår, at der er signifikant forskel på aldersfordelingen mellem respondenter og ikke-respondenter, således at den yngste kohorte og de ældste kohorter typisk er overrepræsenteret blandt ikke-respondenter, sammenlignet med

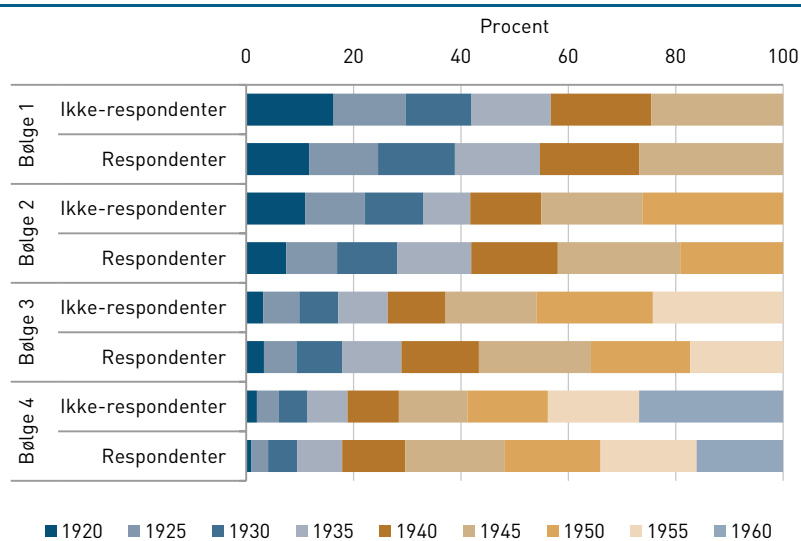
respondenter. Eksempelvis ser vi i fjerde bølge, at de 52-årige og de 87- og 92-årige udgør en større andel af ikke-respondenter end af respondenter.

Sammenligner vi i stedet aldersfordelingen blandt respondenter i fjerde bølge med aldersfordelingen i målpopulationen i 2012 (se bilagstabel B3.1 for målpopulationens fordeling), finder vi ligeledes, at 52-årige er underrepræsenteret i Ældredatabasen, og det samme er 87-årige og 92-årige, alt imens 62- og 67-årige er overrepræsenteret i Ældredatabasen ift. i målpopulationen. Der er således, som det er tilfældet i de fleste surveyundersøgelser blandt voksne og ældre, en tendens til, at personer i den erhvervsaktive alder, samt personer i de ældste aldersgrupper, i højere grad undlader at besvare end de mellemliggende aldersgrupper.

Hvad angår køn afspejler fordelingen blandt respondenter i Ældredatabasen kønsfordelingen i målpopulationen (køn er ikke afbildet her, men findes i bilagstabel B3.3).

FIGUR 4.1

Respondenter og ikke-respondenter fordelt på fødselsårgang, særskilt for hver bølge. Procent.



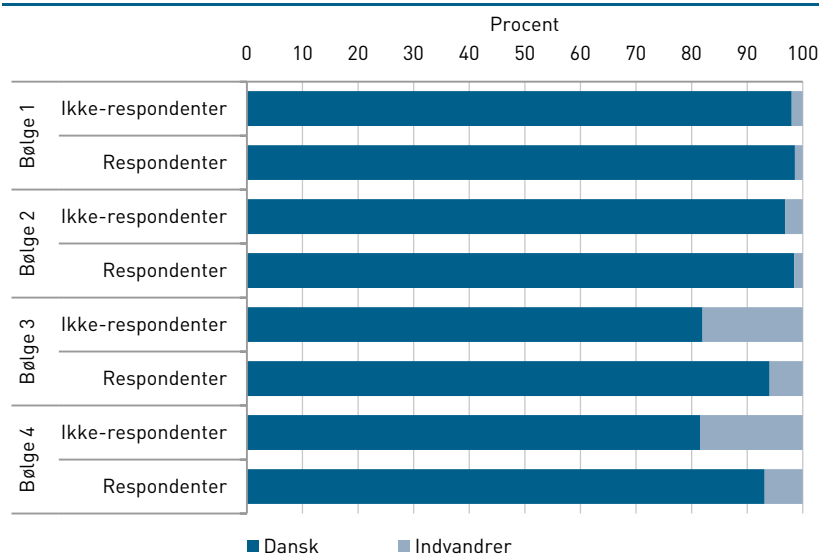
Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på de forskellige fødselsårgange på et 0,001-procents-niveau i hver bølge. Chi²-test.
 Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasens spørgsmål: "Hvilket år er de født?".

Figur 4.2 viser respondenter og ikke-respondenter fordelt på herkomst og viser tydeligt, at indvandrere i alle bølger er overrepræsenteret blandt ikke-respondenter. Figuren viser også, at der i bølge 1 og 2 er relativt få indvandrere med i undersøgelsen, hvilket kan være et resultat af, at ikke-vestlige indvandrere er blevet oversamlet fra 2007 og frem, da de var svært underrepræsenterede tidligere.

Sammenligner vi herkomst blandt Ældredatabasens respondenter med fordelingen i målpopulationen i de 4 år (se bilagstabel B3.2 for målpopulationens fordeling), finder vi, at indvandrere er stærkt underrepræsenteret i Ældredatabasens to første bølger, men let overrepræsenteret i de to seneste bølger, hvor personer med ikke-vestlig herkomst da også er samlet med dobbelt sandsynlighed.

FIGUR 4.2

Respondenter og ikke-respondenter fordelt på herkomst, særskilt for hver bølge. Procent.



Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på herkomst på et 0,001-procents-niveau i bølge 1, 2 og 3. Chi²-test.

1. Efterkommere af indvandrere er kodet som "dansk".

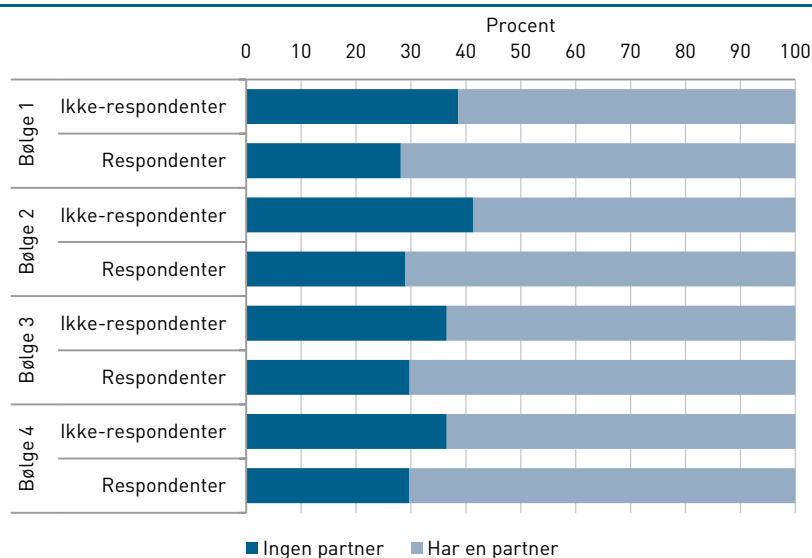
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasens samt IE_TYPE fra GRUND-registret.

Figur 4.3 viser respondenter og ikke-respondenter fordelt på civilstand. Det fremgår, at enlige (personer, der lever uden en partner) i alle bølger er overrepræsenterede blandt ikke-respondenter. I Ældredatabasens første bølge er det således 39 pct. af ikke-respondenter, som er enlige, mens de enlige udgør blot 28 pct. af respondenter. Den samme tendens gør sig gældende i anden, tredje og fjerde bølge.

Sammenligner vi i stedet Ældredatabasens respondenter med målpopulationen (se målpopulationens fordeling i bilagstabel B3.4/bilagstabel B3.4), finder vi, at enlige er underrepræsenteret i Ældredatabasens første bølge, men at fordelingen svarer til fordelingen i populationen i de øvrige tre bølger.

FIGUR 4.3

Respondenter og ikke-respondenter fordelt på civilstand (e-falle), særskilt for hver bølge. Procent.



Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på civilstand på et 0,001-procents-niveau i hver bølge. Pr-test.

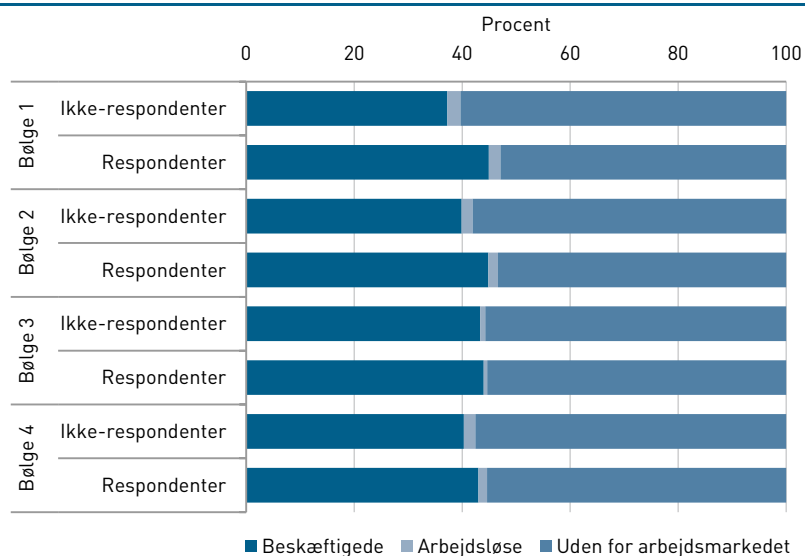
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og e-falle fra GRUND-registret.

SOCIOØKONOMISK STATUS

Figur 4.4 viser respondenter og ikke-respondenter fordelt på arbejdsmarkedstilknytning. Det fremgår, at mere end halvdelen af Ældredatabasens udtræk står uden for arbejdsmarkedet. Dette afspejler naturligt, at der er tale om den ældre befolkningsgruppe. Det fremgår ydermere, at personer, som er beskæftigede, er overrepræsenteret blandt respondenter. Eksempelvis er 45 pct. af respondenter i beskæftigelse, mens de beskæftigede udgør 37 pct. af ikke-respondenter. Tilknytning til arbejdsmarkedet er således associeret med større deltagelse i undersøgelsen.

FIGUR 4.4

Respondenter og ikke-respondenter fordelt på tilknytning til arbejdsmarkedet, særskilt for hver bølge. Procent.



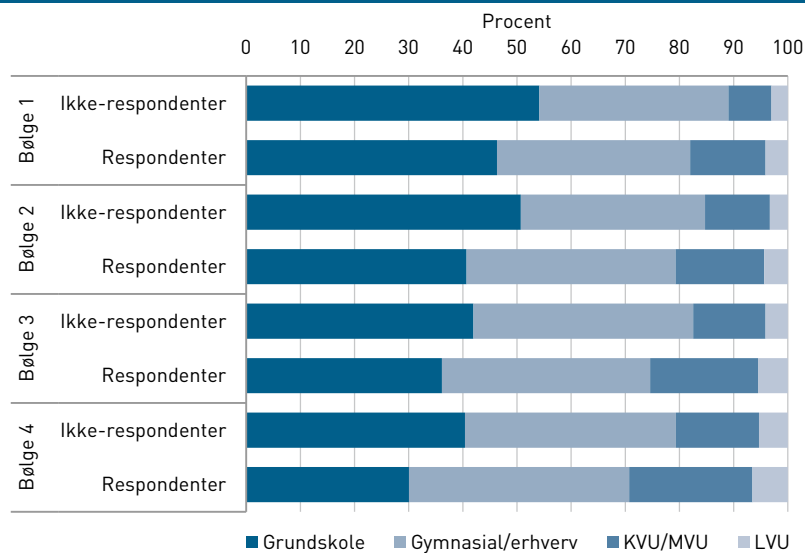
Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på tilknytning til arbejdsmarkedet på et 0,01-procents-niveau i bølge 1, 2 og 4. Chi²-test.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og socstil fra GRUND-registret.

Figur 4.5 viser respondenter og ikke-respondenter fordelt på uddannelsesniveau. Det fremgår overordnet, at størstedelen af Ældredatabasens udtræk har grundskolen som højest fuldførte uddannelse. Dette afspejler, at der er tale om den ældre befolkningsgruppe. Det fremgår, at personer med en kort uddannelsesbaggrund er overrepræsenteret blandt ikke-respondenter i alle fire bølger. I første bølge har således 54 pct. af ikke-respondenterne grundskolen som højest fuldførte uddannelse, mens denne gruppe udgør 46 pct. af respondenter. På samme måde er de højtuddannede overrepræsenteret blandt respondenter. I den fjerde bølge er det således 30 pct. af respondenterne, som har en kort, mellemlang eller lang videregående uddannelse, mens disse grupper udgør 20 pct. af ikke-respondenter.

FIGUR 4.5

Respondenter og ikke-respondenter fordelt på uddannelsesniveau, særskilt for hver bølge. Procent.



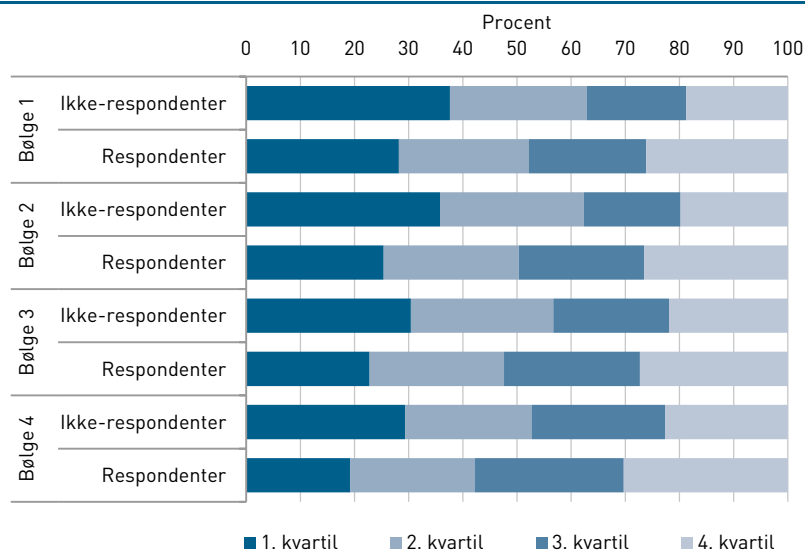
Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på uddannelsesniveau på et 0,001-procents-niveau i hver bølge. Chi²-test.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og hfaudd fra GRUND-registret.

Den socioøkonomiske skævhed i respons synliggøres også, når vi anskuer indkomstniveau og formue. Figur 4.6 viser respondenter og ikke-respondenter fordelt på indkomst-kvartiler, hvor første kvartil repræsenterer personer, der har en indkomst i den laveste fjerdedel af den samlede indkomstfordeling osv. Det fremgår, at personer i den laveste indkomstkvartil er overrepræsenteret blandt ikke-respondenter, mens personer i det højeste indkomstkvartil er overrepræsenteret blandt respondenter. Eksempelvis befinder 37 pct. af ikke-respondenter sig inden for det laveste indkomstkvartil i den første bølge, mens denne gruppe udgør 28 pct. af respondenter. Det samme gør sig gældende for anden, tredje og fjerde bølge. Formue er ikke afbildet her, men en opdeling af respondenter og ikke-respondenter på formuekvartiler viser samme billede med personer inden for de laveste formuekvartiler overrepræsenteret blandt ikke-respondenter.

FIGUR 4.6

Respondenter og ikke-respondenter fordelt på kvartiler for ækvivaleret disponibel indkomst, særskilt for hver bølge. Procent.



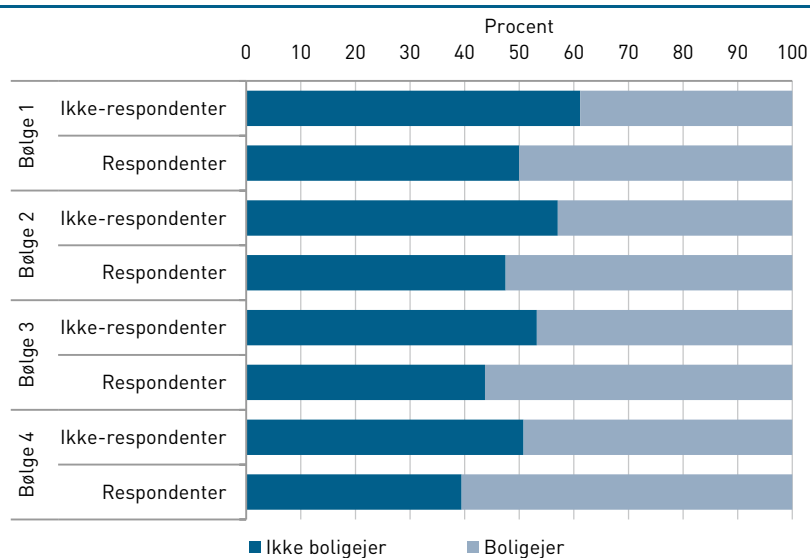
Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på indkomst på et 0,001-procents-niveau i hver bølge. Chi²-test.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og dispon_ny fra GRUND-registret.

En alternativ socioøkonomisk indikator er boligforhold, herunder ejerforhold¹⁰. Figur 4.7 viser respondenter og ikke-respondenter fordelt på ejerforhold. Det fremgår, at boligejere er overrepræsenteret blandt respondenter, mens restgruppen bestående af lejere, andelshavere og andre er overrepræsenteret blandt ikke-respondenter. Blandt respondenter er det i første bølge således 50 pct., som ejer deres bolig, mens denne gruppe udgør 39 pct. af ikke-respondenter. Det samme er gældende i fjerde bølge, hvor 61 pct. blandt respondenterne ejer deres egen bolig, mens det kun er 49 pct. blandt ikke-respondenterne.

FIGUR 4.7

Respondenter og ikke-respondenter fordelt på, om de har egen bolig eller ej, særskilt for hver bølge. Procent.



Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på, om de har egen bolig på et 0,001-procents-niveau i bølge 1, 2 og 3. I bølge 4 er forskellen signifikant på et 0,01-signifikans-niveau. Pr-test.
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og ovskej02_ny og ovskej07 fra GRUNDregistret.

HELBRED

De danske sundhedsregistre rummer en række detaljerede oplysninger, som er registreret i forbindelse med danskernes forbrug af sundhedsydelser, og som med rimelighed kan tjene som indikatorer for individers helbredssituation.

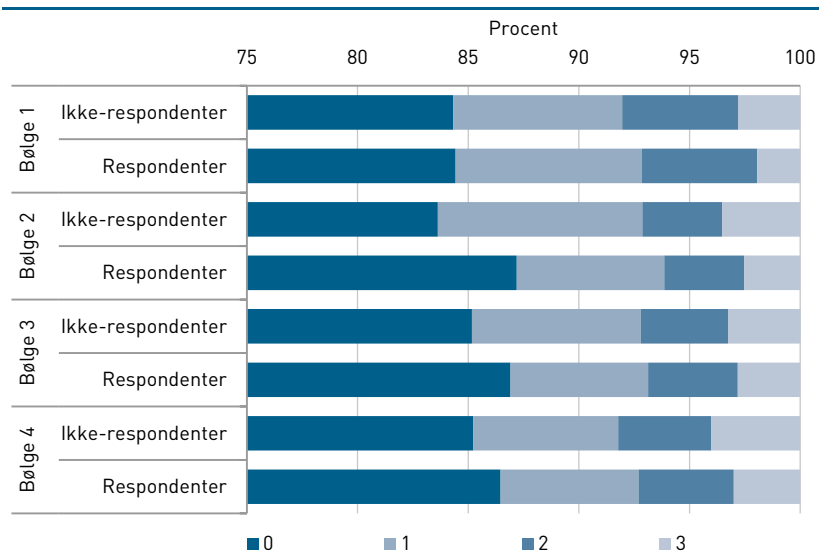
For at undersøge, om der er forskel på respondenter og ikke-respondenters sygdomshistorik, har vi undersøgt diagnoser (aktions- og

10. De registeroplysninger, vi i denne undersøgelse har haft adgang til, rummer ikke direkte informationer om, hvorvidt en person er ejer, lejer, andelshaver eller andet. I stedet har vi på baggrund af variablene ovskej02_ny og ovskej07 dannet en indikator for, hvorvidt personer har et ejerforhold for en bolig. Restgruppen er personer, som ikke ejer en bolig, herunder lejere, andelshavere mv.

bidiagnoser) stillet i forbindelse med hospitalskontakter de seneste 5 år. Disse er kodet ud fra Charlsons Comorbidity Indeks, hvor en række alvorlige diagnoser tildeles henholdsvis et eller flere point (se bilagstabel B1.1 for kodning baseret på Johansen & Fynbo, 2011). Indekset er grupperet således, at 0 indikerer ingen af de 19 diagnoser, mens værdien 1 til 3 indikerer alvorligheden af morbiditet eller komorbiditet. Det fremgår af figur 4.8, at størstedelen af Ældredatabasens population ikke er diagnosticeret med nogen af Charlson diagnoserne inden for de seneste 5 år. Det fremgår desuden, at der er signifikant forskel på fordelingen mellem respondenter og ikke-respondenter i bølge 2, 3 og 4. Andelen, der scorer ét point eller mere på Charlson indekset, er signifikant større blandt ikke-respondenter end blandt respondenter. Forskellene er dog ikke store. Størst er forskellene i anden bølge, hvor 16 pct. af ikke-respondenter scorer ét point eller mere på indekset, alt imens denne gruppe udgør 13 pct. af respondenter.

FIGUR 4.8

Respondenter og ikke-respondenters fordeling på Charlson indekset, særskilt for hver bølge. Procent.



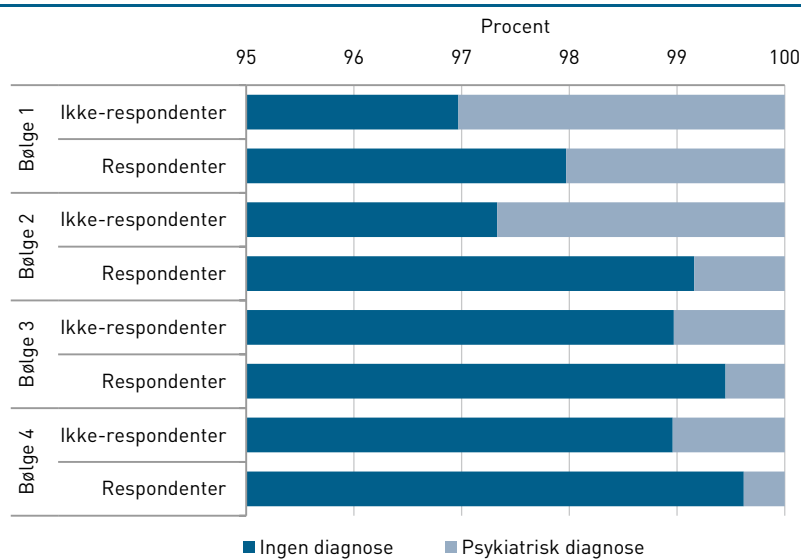
Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på Charlson indekset på et 0,001-procents-niveau i bølge 2 og 3. I bølge 1 og 4 er forskellen signifikant på et 0,01-signifikans-niveau. Chi²-test. Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og addiag fra Landspatient-registret.

Den sidste indikator for helbred, som vi har adgang til gennem sundhedsregistrene, handler om brug af psykiatriske tilbud, herunder ambulant behandling og indlæggelser. Vi koder for en indlæggelse eller ambulant behandling på et psykiatrisk hospital eller lignende de seneste 5 år og begrænser ydermere til diagnoser, der afspejler udvalgte psykiske lidelser (se bilag-

stabel B1.2 for kodning). Det fremgår af figur 4.9, at kun en ganske lille andel af Ældredatabasens udtræk har fået stillet en psykisk diagnose inden for de seneste 5 år. Det fremgår ydermere, at der er signifikant forskel på fordelingen mellem ikke-respondenter og respondenter. Eksempelvis har 3 pct. af ikke-respondenter en psykisk diagnose i Ældredatabasens første bølge, mens denne gruppe udgør 2 pct. af respondenterne.

FIGUR 4.9

Respondenter og ikke-respondenter fordelt på om de har en psykisk diagnose, særskilt for hver bølge. Procent.



Anm.: Der er signifikante forskelle mellem respondenter og ikke-respondenters fordeling på Charlson indekset på et 0,001-procents-niveau i bølge 2 og 4. I bølge 1 og 3 er forskellen signifikant på et 0,01-signifikans-niveau. χ^2 -test.
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og adlæg fra Landspatient-registret.

HVAD PRÆDIKTERER NON-RESPONS?

I det følgende undersøger vi, hvorvidt socioøkonomisk status og helbred er systematisk korreleret med non-respons, når der samtidig tages højde for øvrige karakteristika, og gennemgår udvalgte resultater af modellerne. De estimerede modeller har generelt en lav forklaringskraft, hvilket indikerer, at der er få observerbare karakteristika, som prædikerer non-respons. Resultater fra de estimerede modeller findes i bilagstabel B4.1

MODEL FOR NON-RESPONS

For at anskueliggøre, hvorvidt socioøkonomisk status og helbred har en systematisk betydning for non-respons, når der samtidig kontrolleres for øvrige karakteristika, estimerer vi en logistisk model for hver enkelt af

Ældredatabasens fire bølger, hvis outcome-variabel måler henholdsvis non-respons (1) og respons (0) blandt egnede i den pågældende bølge (dvs. udtrukket til deltagelse, i live og befinder sig i landet). Tabel 4.1 viser datagrundlaget for hver analyse af hver bølges non-respons. I dette afsnit viser vi udvalgte estimater fra Ældredatabasens fjerde bølge (2012). Resultater angående non-respons i de øvrige bølger findes ligeledes i bilagstabel B4.1.

Analysen omfatter 13.337 respondenter fra første, anden og tredje bølge, som stadig var egnede i fjerde bølge samt den nye kohorte af 52-årige (1960), der tilføjes ved hver runde. Heraf har 14 personer missing på én eller flere af modellens indikatorer, disse er ekskluderet fra analysen, som ender med at bero på 13.323 personer.

Der er foretaget en modelsøgning, som giver anledning til følgende forklarende variable: fødselsårgang, køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. Vi ville gerne kunne kontrollere for uddannelsesniveau, men relativt mange personer i Ældredatabasen (>1.000) har manglende uddannelsesinformationer, og det har ikke været muligt for os at finde oplysninger om disse personernes uddannelsesniveau. Derfor har vi valgt at udelade uddannelse fra modellen, da det ellers ville betyde, at mere end 1.000 personer udelades fra analysen. Dette ville særligt være problematisk, da nærmere granskning afslørede, at de manglende uddannelsesoplysninger i høj grad er korreleret med både etnisk herkomst og høj alder. Ved at inkludere uddannelse i modellen ville vi derfor ekskludere den ældre del af befolkningen samt en stor del af mennesker født uden for Danmark. I ovenstående kapitel er det muligt at se den deskriptive fordeling på uddannelse for respondenter og ikke-respondenter, også for de uoplyste på uddannelsesniveau.

TABEL 4.1

Fordeling på outcome-variabel for logistisk regressionsmodel over non-respons. Antal og procent.

Outcome	Bølge 1		Bølge 2		Bølge 3		Bølge 4	
	Antal	Pct.	Antal	Pct.	Antal	Pct.	Antal	Pct.
0 = Respons	5.864	70,3	8.207	83,6	9.633	74,5	9.965	74,7
1 = Non-respons	2.476	29,7	1.613	16,4	3.303	25,5	3.372	25,3
Total	8.340	100	9.820	100	12.936	100	13.337	100
Heraf uden missing	8.300		9.808		12.911		13.323	

Kilde: Ældredatabasens bruttostikprøve.

Følgende afsnit vil beskrive de signifikante baggrundsfaktorer, der prædikterer non-respons i Ældredatabasens fjerde bølge i 2012. For at lette

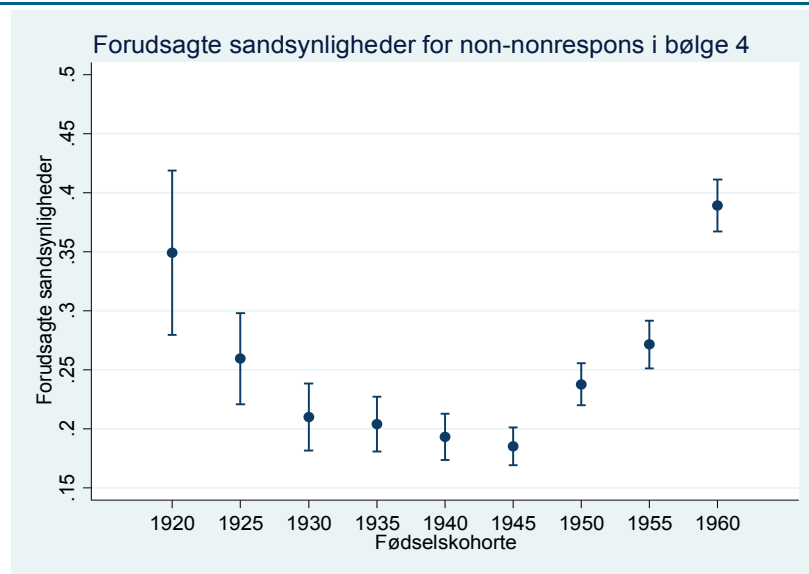
analysens fortolkning rapporteres resultaterne ved hjælp af estimerede sandsynligheder.

SOCIODEMOGRAFI

Det fremgår af figur 4.10, at der er en U-formet sammenhæng mellem alder (her målt ved fødselskohorte) og sandsynligheden for non-respons. Den højeste sandsynlighed for non-respons findes blandt de ældste og yngste kohorter, mens de mellemliggende kohorter gennemsnitligt har en lavere sandsynlighed for non-respons. Eksempelvis er den gennemsnitlige estimerede sandsynlighed for non-respons ca. 15 procentpoint større blandt kohorten født 1920 (92 år i 2012) end blandt kohorten født 1940 (72 år i 2012). På samme vis er sandsynligheden for non-respons ca. 19 procentpoint højere blandt kohorten født 1960 (52 år i 2012) end blandt kohorten født 1940 (72 år i 2012). Denne U-formede sammenhæng, som ses, selv når der kontrolleres for socioøkonomisk status og helbred, indikerer, at personer i den erhvervsaktive alder og de ældste ældre, er mindre tilbøjelige til at besvare Ældredatabasens spørgeskema end personer, der netop har nået tilbagetrækningsalderen.

FIGUR 4.10

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af fødselskohorte. Estimat for sandsynlighed og konfidensinterval.



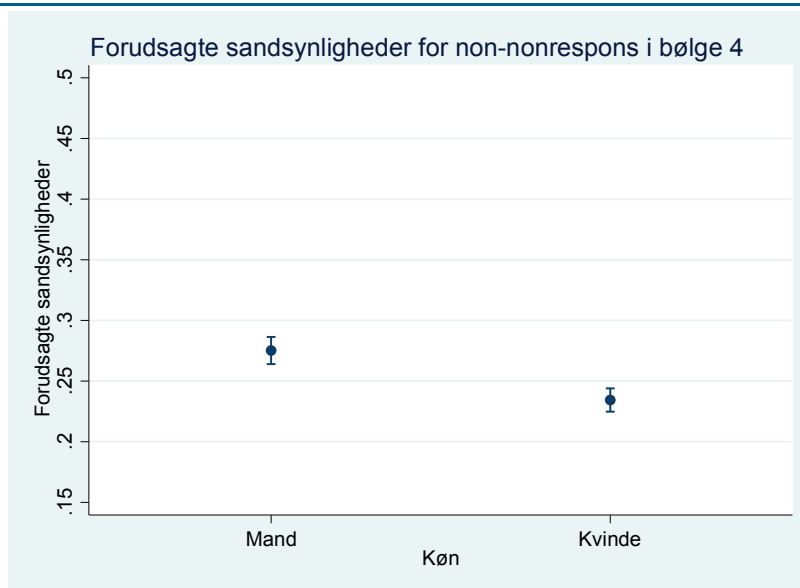
Anm.: Multivariat model med kontrol for: køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasens spørgsmål: "Hvilket år er De født?".

Det fremgår af figur 4.11, at sandsynligheden for non-respons blot er marginalt større blandt mænd end blandt kvinder. Den gennemsnitlige estimerede sandsynlighed for non-respons er ca. 5 procentpoint højere blandt mænd end blandt kvinder. Selv når der kontrolleres for socioøkonomisk status, alder og helbredsindikatorer, er mænd således mindre tilbøjelige til at deltage i Ældredata-basen end kvinder. Dette er i overensstemmelse med fund fra de fleste spørgeskemaundersøgelser.

FIGUR 4.11

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af køn. Estimat for sandsynlighed og konfidensinterval.



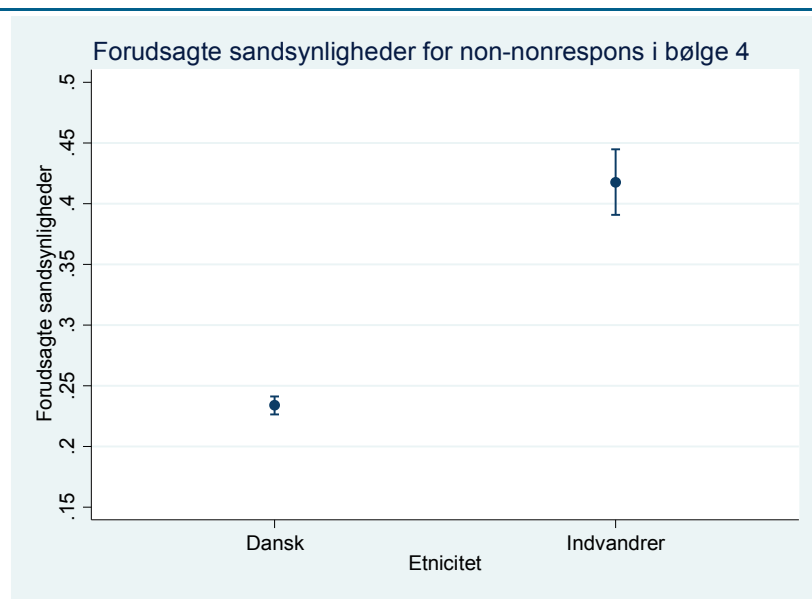
Anm.: Multivariat model med kontrol for: fødselsårgang, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredata-basen og køn fra GRUND-registret.

Figur 4.12 viser de gennemsnitlige estimerede sandsynligheder for non-respons i Ældredatabasens fjerde bølge, for danskere og indvandrere. Bemærk, at indvandrere med ikke-vestlig baggrund er samlet med dobbelt sandsynlighed i Ældredatabasens tredje og fjerde bølge med det formål at øge repræsentativiteten i denne gruppe. Det fremgår, at indvandrere, som er udtrukket, er mindre tilbøjelige til at deltage i undersøgelsen end danskere. Blandt indvandrere er den estimerede sandsynlighed for non-respons 19 procentpoint højere end blandt danskere, der er dog en vis usikkerhed på dette estimat. Forskellen skyldes formentlig sprogvanskeligheder, idet Ældredatabasens spørgeskema ikke er oversat til andre sprog end dansk. Den deskriptive analyse viser dog, at personer med anden etnisk herkomst er fint repræsenteret i Ældredatabasens tredje og fjerde bølge, hvor personer med ikke-vestlig baggrund er oversamlet med dobbelt sandsynlighed.

FIGUR 4.12

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af etnicitet. Estimat for sandsynlighed og konfidensinterval.



Anm.: Multivariat model med kontrol for: kohorte, køn, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.

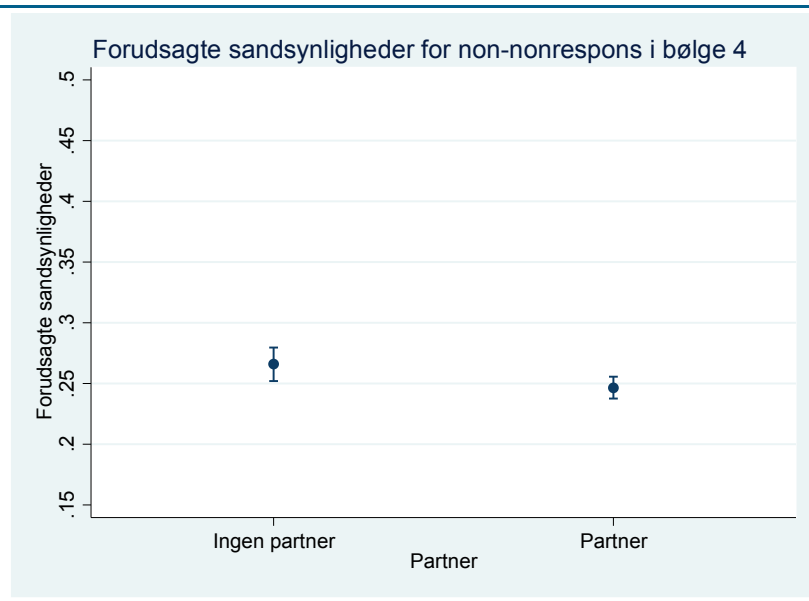
1. Efterkommere er kodet som dansk.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og ietype fra GRUND-registret.

Figur 4.13 viser de gennemsnitlige estimerede sandsynligheder for non-respons for ældre med og uden en partner. Figuren viser, at der er en lille statistisk signifikant forskel sådan, at ældre uden en partner har en lidt højere sandsynlighed for non-respons (2 procentpoint) end ældre med en partner. Altså er ældre med en partner, selv når vi kontrollerer for socio-økonomisk status, alder og helbredsindikatorer, en smule mere tilbøjelige til at besvare Ældredatabasens spørgeskema end ældre uden en partner.

FIGUR 4.13

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af partner. Estimat for sandsynlighed og konfidensinterval.



Anm.: Multivariat model med kontrol for: kohorte, køn, etnisk herkomst, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.

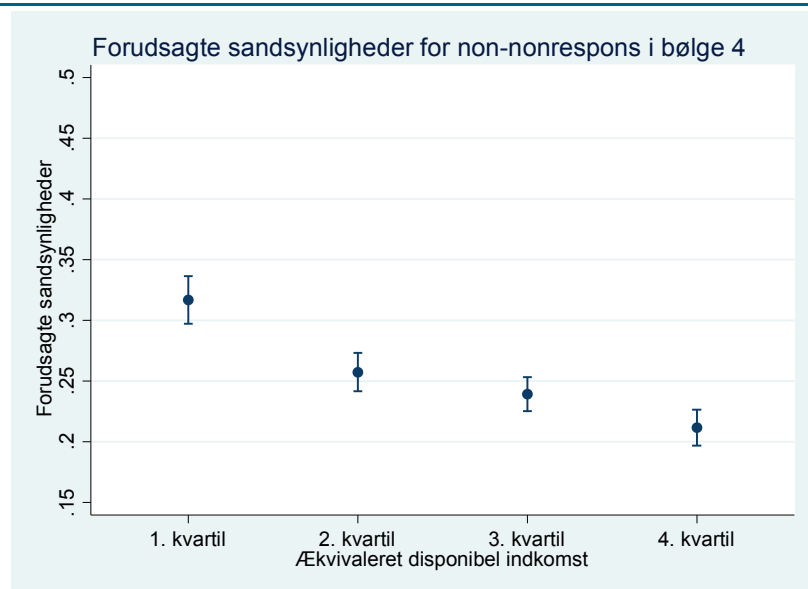
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og efalle fra GRUND-registret.

SOCIOØKONOMISK STATUS

Det fremgår ligeledes af figur 4.14, at personer i de laveste indkomstkvarterer har større sandsynlighed for non-respons end personer i de højeste indkomstkvarterer. De ældre, der har en indkomst i den laveste fjerdedel af den samlede indkomstfordeling, har 11 procentpoint højere sandsynlighed for non-respons sammenlignet med ældre, der har en indkomst i den højeste fjerdedel af den samlede indkomstfordeling. Det tyder derfor på, at ældre med en mindre indkomst er mindre tilbøjelige til at besvare ældreskemaet.

FIGUR 4.14

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af disponibel indkomst. Estimat for sandsynlighed og konfidensinterval.



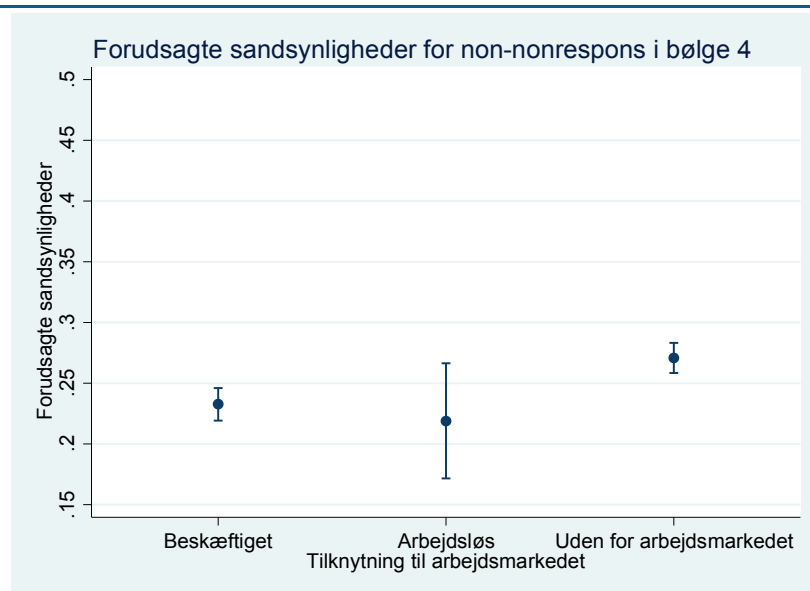
Anm.: Multivariat model med kontrol for: kohorte, køn, etnisk herkomst, partner, flytning, boligforhold, arbejdsmarkedstil-knytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.
N = 13.323.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og dispon_ny fra GRUND-registret.

Figur 4.15 viser den estimerede sandsynlighed for non-respons alt efter arbejdsmarkedstilknytning. Det fremgår, at personer, der står uden for arbejdsmarkedet (som eksempelvis er pensionerede, modtager førtidspension eller aldrig har haft et arbejde), generelt har en større sandsynlighed for non-respons end personer, der er i beskæftigelse. Der er ikke signifikant forskel på beskæftigede og arbejdsløses sandsynlighed for non-respons, men personer, der står uden for arbejdsmarkedet, har 4 procentpoint større sandsynlighed for non-respons end personer, der er i beskæftigelse.

FIGUR 4.15

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af tilknytning til arbejdsmarkedet. Estimat for sandsynlighed og konfidensinterval.



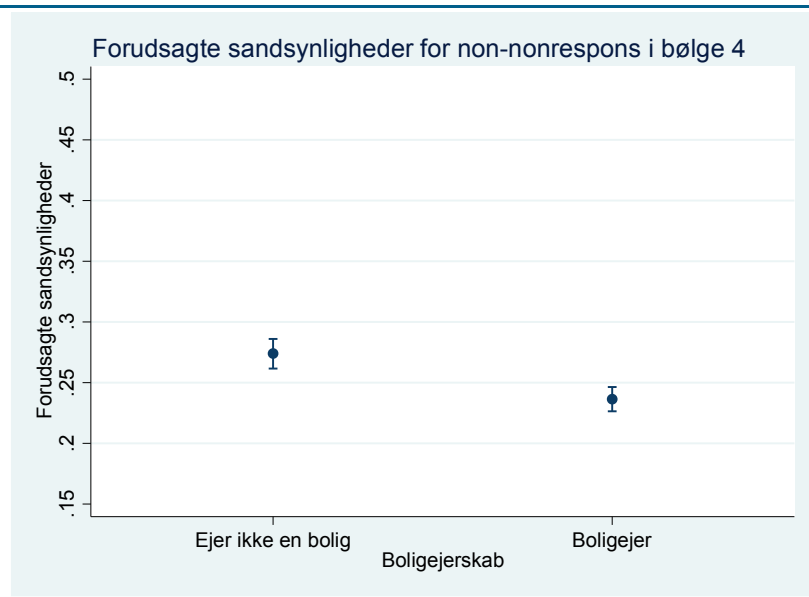
Anm.: Multivariat model med kontrol for: kohorte, køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.

Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og socstil fra GRUND-registret.

Figur 4.16 viser sandsynligheden for non-respons fordelt på boligforhold. At være lejer, andelshaver mv. frem for boligejer er forbundet med en større sandsynlighed for non-respons. Sandsynlighed for non-respons er omkring 3 procentpoint højere blandt lejere og andelshavere end blandt boligejere.

FIGUR 4.16

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af boligejerskab. Estimat for sandsynlighed og konfidensinterval.



Anm.: Multivariat model med kontrol for: kohorte, køn, etnisk herkomst, partner, flytning, disponibel indkomst, arbejdsmarkedstilknøytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.

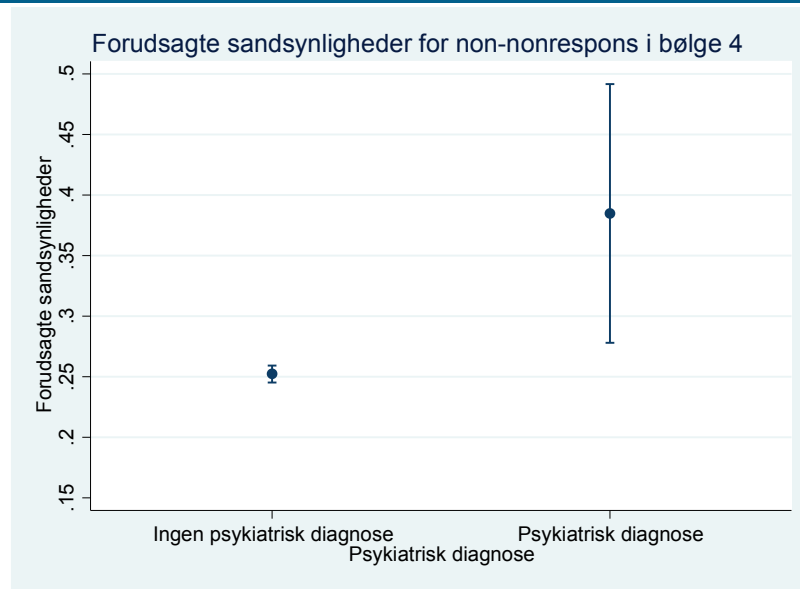
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og ovskej02_ny og ovskej07 fra GRUND-registret.

HELBRED

Hvor de socioøkonomiske variable indikerer, at der er en systematisk sammenhæng mellem lav socioøkonomisk status og sandsynlighed for non-respons, så genfinder vi ikke denne sammenhæng, når vi betragter helbredsindikatorer. Figur 4.17 viser den estimerede sandsynlighed for non-respons blandt ældre med og uden en psykiatrisk diagnose. Figuren viser, at der er relativ stor forskel i sandsynligheden for non-respons, alt efter om de ældre har psykiatrisk diagnose eller ej. Uden en diagnose har ældre, kontrolleret for demografi, økonomi og alder, en sandsynlighed for non-respons på 25 pct. Har de derimod en diagnose, er deres sandsynlighed 38 pct. – altså en stigning i sandsynlighed for non-respons på 13 procentpoint.

FIGUR 4.17

Gennemsnitlig forudsagt sandsynlighed for non-respons i bølge 4 på baggrund af score på psykiatrisk diagnose. Estimat for sandsynlighed og konfidensinterval.



Anm.: Multivariat model med kontrol for: kohorte, køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknøytning, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. N = 13.323.
Kilde: Egne beregninger på baggrund af bruttostikprøven for Ældredatabasen og addiag fra Landspatient-registret.

REPRÆSENTATIVITET

Vi har i det foregående sammenlignet respondenter og ikke-respondenter på udvalgte karakteristika, og vi har estimeret sandsynligheden for non-respons. I det følgende undersøger vi, hvilken betydning non-respons har for repræsentativiteten i hver af Ældredatabasens fire bølger. Dette gør vi ved at sammenligne fordelingen blandt respondenter med fordelingen i målpopulationen. Målpopulationen består som bekendt af alle danske ældre fra ni udvalgte kohorter. Data om målpopulationen stammer fra Statistikbanken (Danmarks Statistik, 2016), hvor følgende indikatorer er tilgængelige: alder, køn, herkomst, partner og arbejdsmarkedstilknøytning. Vi tester for hver af disse indikatorer, om der er forskel på fordelingen blandt respondenter og fordelingen blandt målpopulationen. Dette tester vi ved hjælp af en χ^2 -test, og vi afrapporterer med stjerner signifikansniveauet for p-værdien.

MÅLPOPULATION OG RESPONDENTER

I tabel 4.2 viser vi aldersfordelingen blandt respondenter i Ældredatabasens fire bølger og blandt målpopulationen, som består af ældre i den danske befolkning. Det fremgår af tabellen, at der er signifikante afvigelser mellem respondenteres aldersfordeling og aldersfordelingen i målpopulationen i alle fire bølger. Imidlertid viser tabellen samtidig, at der er tale om ganske små forskelle.

TABEL 4.2

Aldersfordeling blandt målpopulation og respondenter samt test af forskelle (chi²-test). Særskilt for bølger. Procent.

År	1997		2002		2007		2012	
	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.
52	26,4	26,8	21,4	19,2	19,1	17,3	18,1	16,1
57	18,9	18,5	22,5	22,8	19,0	18,5	17,3	17,9
62	16,3	15,8	15,8	16,1	19,6	20,9	17,0	17,8
67	14,4	14,3	13,1	13,8	13,4	14,5	17,1	18,5
72	13,7	12,8	11,1	11,2	10,6	11,1	11,3	11,8
77	10,3	11,7	9,7	9,5	8,3	8,5	8,3	8,3
82	-	-	6,4	7,5	6,5	6,0	5,8	5,5
87	-	-	-	-	3,5	3,3	3,7	3,1
92	-	-	-	-	-	-	1,4	1,0
Tot.	100	100	100	100	100	100	100	100
N/n	300.750	5.864	341.970	8.207	374.141	9.633	402.078	9.965
Sign. chi ² -test		***		***		***		***

I tabel 4.3 viser vi, hvordan henholdsvis respondenter i Ældredatabasen og målpopulationen fordeler sig efter køn. Det fremgår, at kønsfordelingen blandt respondenter ikke adskiller sig signifikant fra fordelingen i målpopulationen i Ældredatabasens tre første bølger. I fjerde bølge er kvinder signifikant overrepræsenteret i Ældredatabasen i sammenligning med målpopulationen, om end der størrelsesmæssigt er tale om en mindre forskel.

TABEL 4.3

Fordeling efter køn blandt målpopulation og respondenter, samt test af forskelle (chi²-test). Særskilt for bølger. Procent.

År	1997		2002		2007		2012	
	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.
Mænd	47,9	48,2	47,6	47,5	47,5	46,6	47,6	46,5
Kvinder	52,1	51,8	52,4	52,5	52,5	53,4	52,4	53,6
Total	100	100	100	100	100	100	100	100
N/n	300.750	5.864	341.970	8.207	374.141	9.633	402.078	9.965
Sign. chi ² -test								**

I tabel 4.4 viser vi, hvordan henholdsvis respondenter i Ældredata-basen og målpopulationen fordeler sig efter herkomst. Det fremgår af tabellen, at der er signifikante afvigelser mellem respondenter og målpopulationen i alle fire bølger. I Ældredata-basens to første bølger er indvandrere underrepræsenteret blandt respondenter i forhold til målpopulationen, mens etnisk danske ældre er overrepræsenteret blandt respondenter. I de to seneste bølger af Ældredata-basen er dette mønster imidlertid ændret. Dette skyldes indførelsen af en ny praksis i indsamlingsmetoden, hvor ældre med ikke-vestlig oprindelse er udtrukket med dobbelt sandsynlighed. Vi ser, at denne ændrede praksis har haft den ønskede virkning, idet ældre indvandrere er godt repræsenteret – faktisk endda overrepræsenteret – blandt respondenter i 2007 og 2012 sammenlignet med målpopulationen.

TABEL 4.4

Fordeling efter herkomst blandt målpopulation og respondenter, samt test af forskelle (chi²-test). Særskilt for bølger. Procent.

År	1997		2002		2007		2012	
	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.
Indvand.	3,9	1,4	4,4	1,6	5,0	6,0	5,8	6,9
Dansk	96,1	98,6	95,6	98,4	95,0	94,0	94,2	93,1
Total	100	100	100	100	100	100	100	100
N/n	300.750	5.864	341.970	8.207	374.141	9.633	402.078	9.965
Sign. chi ² -test		***		***		***		***

Anm.: Dansk omfatter her også efterkommere af indvandrere.

I tabel 4.5 viser vi, hvordan henholdsvis respondenter i Ældredata-basen og målpopulationen fordeler sig efter familietype. I den første og seneste bølge af Ældredata-basen adskiller fordelingen blandt respondenter sig signifikant fra fordelingen i målpopulationen, mens der i de to midterste type er opnået repræsentativitet. I første bølge er enlige således overrepræsenteret blandt respondenter, mens enlige i sidste bølge er underrepræsenteret blandt respondenter. I de to midterste bølger af Ældredata-basen er der opnået repræsentativitet af både enlige ældre og ældre med en partner.

TABEL 4.5

Fordeling efter familietype blandt målpopulation og respondenter samt test af forskelle (chi²-test). Særskilt for bølger. Procent.

År	1997		2002		2007		2012	
	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.	Målpop.	Resp.
Partner	73,7	71,9	71,7	71,0	69,6	70,2	68,6	70,3
Enlig	26,3	28,1	28,3	29,0	30,4	29,8	31,4	29,7
Total	100	100	100	100	100	100	100	100
N/n	300.631	5.864	341.898	8.207	374.140	9.633	402.078	9.965
Sign. chi ² -test		***						***

I tabel 4.6 viser vi, hvordan henholdsvis respondenter i Ældredata-basen og målpopulationen fordeler sig efter arbejdsmarkedstilknytning. Data om arbejdsmarkedstilknytning, som er udtrukket fra Statistikbanken, omfatter alle 50-årige og ældre, hvorfor data ikke perfekt repræsenterer målpopulationen for Ældredata-basen, der jo omfatter udvalgte aldersgrupper (52, 57, 62 år etc.). Tilmed er data om målpopulationens arbejdsmarkedstilknytning kun tilgængeligt for året 2012. Alligevel kan en sammenligning give et praj om Ældredata-basens repræsentativitet, hvad angår arbejdsmarkedstilknytning blandt respondenter i Ældredata-basens fjerde bølge. Det fremgår, at der er signifikant forskel på fordelingen blandt respondenter i Ældredata-basen og i målpopulationen, idet beskæftigede er overrepræsenteret blandt respondenter, mens personer, der står uden for arbejdsmarkedet, er underrepræsenteret blandt respondenter. Det er dog ikke sikkert, om denne forskel kan skyldes, at målpopulationen omfatter flere aldersgrupper end de, som er udtrukket til Ældredata-basen.

TABEL 4.6

Fordeling efter arbejdsmarkedstilknytning blandt målpopulation og respondenter i 2012 samt test af forskelle (χ^2 -test). Procent.

År	2012	
	Målpopulation	Respondenter
Arbejdsmarkeds-tilknytning		
Beskæftigede	40,6	43,0
Arbejdsløse	1,7	1,6
Uden for arbejdsstyrken	57,7	55,4
Total	100	100
N/n	2.067.137	9.952
Sign. χ^2 -test		***

Anm.: Tilknytning til arbejdsmarkedet er for målpopulationen opgjort blandt 50-årige og ældre, hvilket betyder, at fordelingen ikke er direkte sammenlignelig med survey-populationen, der blot inkluderer følgende aldersgrupper 52-, 57-, 62-, 67-, 72-, 77-, 82-, 87- og 92-årige.

Alt i alt viser sammenligningen af respondenter i Ældredata-basen med målpopulationen, at der er signifikante forskelle i fordelingen for de fleste af de undersøgte indikatorer. Imidlertid er der tale om størrelsesmæssigt ganske små forskelle, hvorfor vi ikke ser grund til at være bekymrede om Ældredata-basens repræsentativitet. Vi finder dog alligevel grund til at anbefale, at der for fremtiden udarbejdes vægte til Ældredata-basen i forbindelse med dataindsamlingen.

BORTFALD

BORTFALD MELLEM ÆLDRE DATABASENS BØLGER

I det foregående kapitel undersøgte vi non-respons i Ældredatabasen på tværs af hver enkelt af Ældredatabasens fire bølger. I det efterfølgende kapitel undersøger vi omfanget og karakteret af bortfald *mellem* Ældredatabasens bølger. Fastholdelse i Ældredatabasen er vigtig for panelets repræsentativitet, men årsagerne til bortfald mellem runder kan være mange. I forløbsdatabaser med ældre befolkningsgrupper er det særligt vigtigt at skelne mellem aktivt og passivt bortfald, hvor passivt bortfald betegner bortfald pga. død, mens aktivt bortfald betegner bortfald af andre (evt. ukendte) årsager.

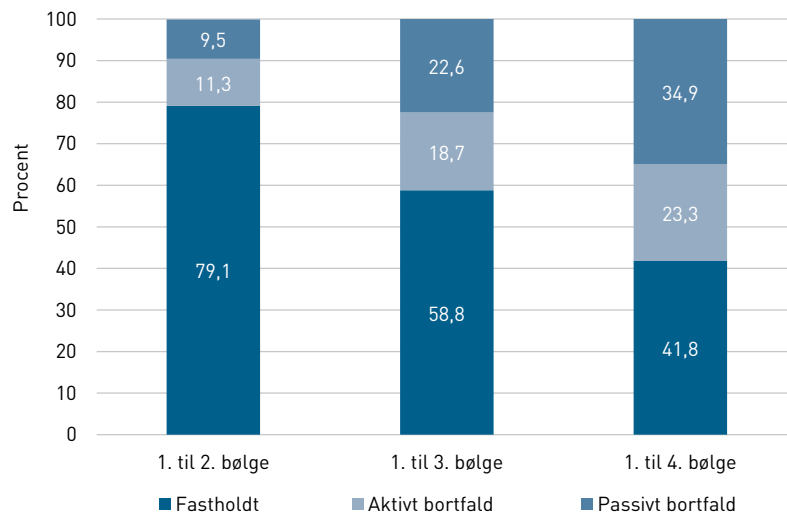
Figur 5.1 viser aktivt og passivt bortfald i Ældredatabasen blandt respondenter, som besvarede Ældredatabasens første bølge.¹¹ Det fremgår, at en betydelig andel af panelets bortfald over tid kan tilskrives dødelighed (dvs. passivt bortfald). Det fremgår ligeledes, at det passive bortfald stiger bølge for bølge, i takt med at deltagerne i den oprindelige stikprøve ældes. Ud af alle respondenter, som deltog i Ældredatabasens første bølge, var 10 pct. gået bort før undersøgelsens anden bølge, 23 pct. var gået bort før undersøgelsens tredje bølge, og 35 pct. var gået bort før undersøgelsens fjerde bølge.¹²

11. Og som ydermere ikke migrerede i perioden (28 personer sorteret fra).

12. En mere detaljeret opdeling på fødselsår viser desuden, at det passive bortfald er størst blandt de ældste kohorter. Således er hele 81 pct. af den ældste kohorte (født 1920) faldet fra pga. død i fjerde runde (2012), mens dette gælder for 13 pct. af den yngste kohorte (født 1945).

FIGUR 5.1

Passivt og aktivt bortfald i Ældredatabasen blandt første bølges respondenter, fordelt på anden, tredje og fjerde bølge. Procent.



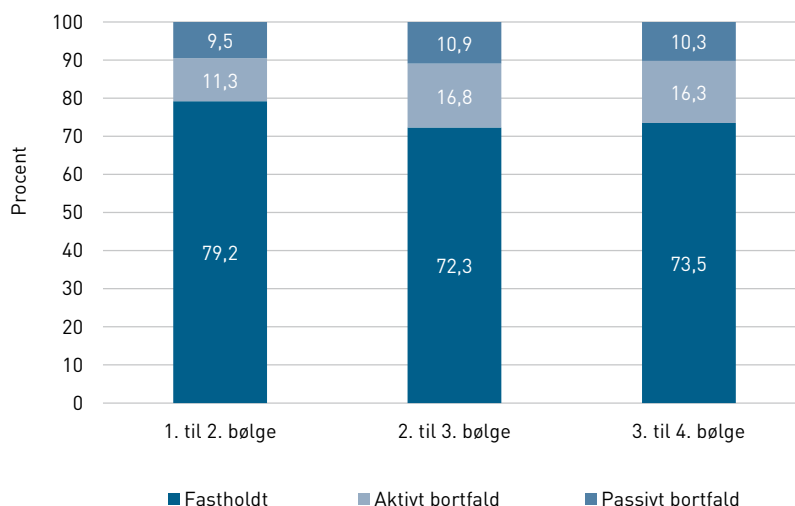
Anm.: Bortfaldsprocenterne er udregnet for respondenter i Ældredatabasens første bølge. Personer, som er udvandet i perioden 1997 til 2012, indgår ikke i beregningsgrundlaget. N = 5.838.

Kilde: Ældredatabasens bruttostikprøve koblet til dødsdato fra DODE-registret og migrationsdata fra VNDS-registret.

Mens figur 5.1 illustrerer bortfaldet over fire bølger for første bølges respondenter, viser figur 5.2 bortfaldet mellem hver enkelt bølge. Det fremgår, at både det aktive og passive bortfald er nogenlunde stabilt over tid. Således er det passive bortfald fra én bølge til den næste på 10-11 pct. blandt den tidlige bølges respondenter, mens det aktive bortfald ligger mellem 11 og 17 pct.

FIGUR 5.2

Passivt og aktivt bortfald mellem bølger i Ældredatabasen fordelt på anden, tredje og fjerde bølge. Procent.



Anm.: Bortfaldsprocenterne er for hver bølge udregnet for forrige bølges respondenter. Personer, som er udvandret i perioden siden forrige bølge, er ikke i beregningsgrundlaget. N = 5.853, N = 8.193, N = 9.126.

Kilde: Ældredatabasens bruttostikprøve koblet til dødsdato fra DODE-registret og migrationsdata fra VNDS-registret.

Særligt aktivt bortfald er problematisk, hvis det ikke er tilfældigt ("missing completely at random", MCAR), da dette kan forskyde databasens repræsentativitet således, at blot et selektivt udsnit af ældre er repræsenteret i undersøgelsen. For at imødekomme denne udfordring er Ældredatabasens indsamling designet med tilføjelse af et opfyldnings-udtræk i hver ny bølge, således at ikke-respondenter og bortfaldne erstattes med interviews med nye respondenter fra samme fødselsårgang i hver af de efterfølgende bølger.

HVAD PRÆDIKERER AKTIVT BORTFALD?

I det følgende findes resultaterne af en bortfaldsanalyse, hvor vi undersøger, hvorvidt bortfald er systematisk korreleret med lav socioøkonomisk status eller ringe helbred. Passivt bortfald reflekterer i sagens natur den naturlige selektivitet, der sker i aldrende befolkninger, og er for så vidt ikke problematisk for databasens repræsentativitet. Vi fokuserer derfor udelukkende på aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge og begrænser således analysen til respondenter, der besvarede Ældredatabasens tredje bølge, og som stadig var i live ved Ældredatabasens fjerde bølge.

MODEL FOR AKTIVT BORTFALD

For at anskueliggøre, hvorvidt socioøkonomisk status og helbred har en systematisk betydning for non-respons, når der samtidig kontrolleres for øvrige karakteristika, estimerer vi en logistisk model for aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge, hvis outcome-variabel måler henholdsvis aktivt bortfald (1) og respons (0) i fjerde bølge blandt personer, der deltog i tredje bølge. Passivt bortfald er sorteret fra, og analysen omfatter således udelukkende personer, der var egnede til deltagelse i Ældredatabasens fjerde bølge. Analysen omfatter 8.634 respondenter fra tredje bølge, som stadig var egnede i fjerde bølge, hvoraf 18 pct. faldt fra i fjerde bølge (tabel 5.1).

Der er foretaget en modelsøgning, som giver anledning til følgende forklarende variable: fødselsårgang, køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose og Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse. På disse indikatorer har 34 personer missing på én eller flere af modellens forklarende variable og er således ekskluderet fra analysen (tabel 5.1). Som mål for socioøkonomisk status inkluderer vi i modellen husstandsækvivaleret indkomst¹³.

TABEL 5.1

Fordeling på outcome-variabel for logistisk regressionsmodel over aktivt bortfald. Antal og procent.

Outcome	Antal	Pct.
0 = fastholdt	7.067	81,9
1 = aktivt bortfald	1.567	18,1
Total	8.634	100
Heraf, uden missing	8.600	-

Anm.: Analysen af aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge omfatter respondenter fra tredje bølge, der er egnede i fjerde bølge, dvs. at passivt bortfald er sorteret fra.

Kilde: Ældredatabasens bruttostikprøve.

Denne variabel omfatter imidlertid ikke alle former for pensionsindkomster, ligesom den ikke afspejler formue såsom friværds og aktier. Som et robusthedstjek har vi derfor ligeledes estimeret tre alternative modeller, hvor indkomst erstattes med hhv. uddannelsesniveau, formue og samlede pensionsindkomster. Disse modeller afrapporteres ikke i nærværende notat, men hovedresultaterne er de samme som i den valgte model¹⁴.

13. Vi vælger indkomst frem for uddannelsesniveau, idet 473 personer har missing på uddannelse i registrene. Missing på uddannelse er samtidig stærkt korreleret med alder og etnisk herkomst. Derfor udelader vi uddannelse fra den multivariate model. Som en alternativ indikator for socioøkonomisk status inddrager vi husstandsækvivaleret disponibel indkomst.

14. En markant forskel optræder dog i den alternative model, hvor uddannelse inkluderes i stedet for indkomst. Her bliver effekten af at tilhøre de ældste kohorter mindre, ligesom betydningen af etnisk herkomst er insignifikant i denne model. Dette skyldes imidlertid, at uddannelse i registrene er missing for 473 personer (svarende til 5,5 pct. af analysepopulationen), og samtidig stærkt kor-

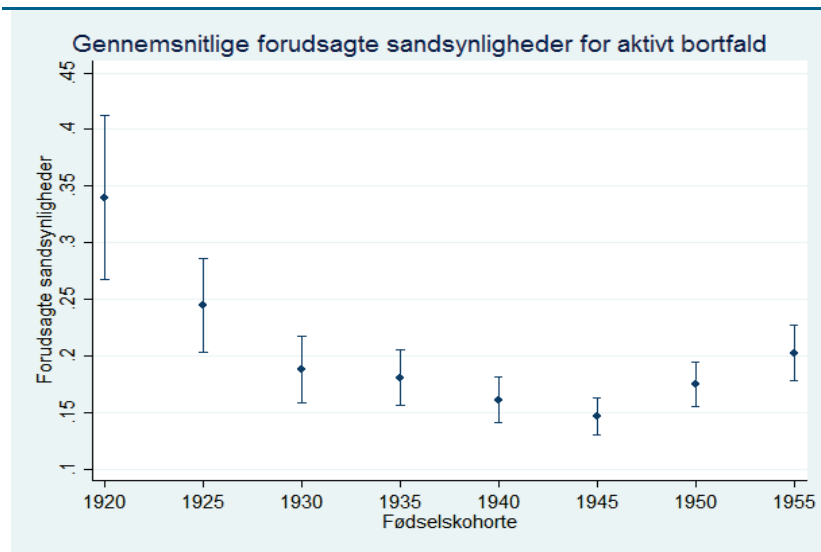
I de efterfølgende afsnit gennemgår vi udvalgte resultater fra modellen (visualiseret ved gennemsnitlige estimerede sandsynligheder). Ligesom i modellerne for non-respons har de estimerede bortfaldsmodeller generelt en lav forklaringskraft, hvilket indikerer, at der er få observerbare karakteristika, som prædikerer bortfald. Resultater fra de estimerede modeller findes desuden i bilagstabel B4.1 bilagstabel B4.1.

SOCIODEMOGRAFI

Det fremgår af figur 5.3, at der er en negativ sammenhæng mellem alder (her målt ved fødselskohorte) og sandsynligheden for bortfald. Den højeste sandsynlighed for bortfald findes blandt de ældste kohorter, mens de yngre kohorter gennemsnitligt generelt har en lavere sandsynlighed for bortfald. Eksempelvis er den gennemsnitlige estimerede sandsynlighed for bortfald ca. 14 procentpoint højere blandt kohorten født i 1920 (92 år i 2012) end blandt kohorten født i 1955 (57 år i 2012).

FIGUR 5.3

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt på fødselskohorte.



Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Aanalysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.

Kilde: Ældredatabasens bruttostikprøve. Variablen for kohorte.

Denne sammenhæng, som ses, selv når der kontrolleres for socioøkonomisk status og helbred, indikerer, at ældre er mere tilbøjelige til at falde

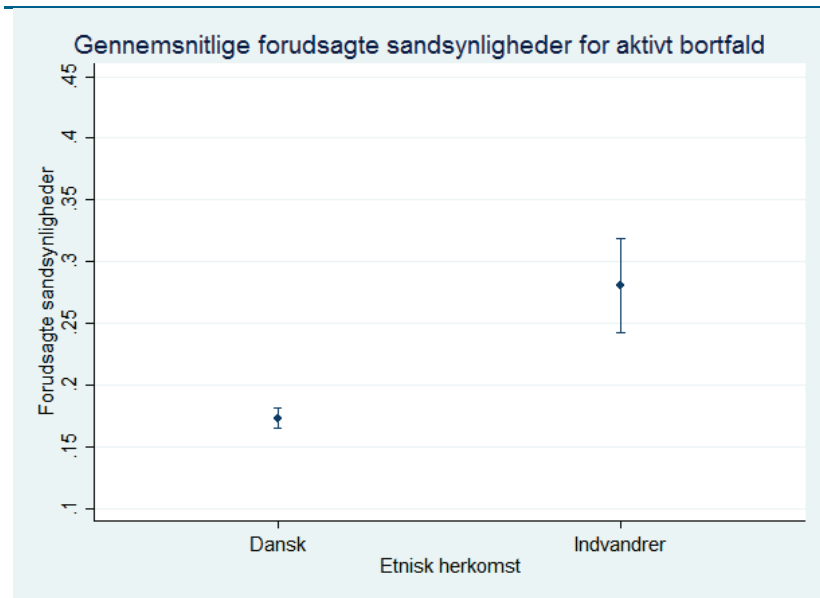
relateret med henholdsvis alder og etnisk herkomst. Modellen viser dog tydeligt en uddannelsesmæssig gradient i sandsynligheden for bortfald mellem bølger.

fra imellem Ældredatabasens bølger. Hvis vi antager, at de registerbase-rede helbredsindikatorer, vi har inkluderet i modellen, ikke afspejler de ældres fulde helbredsprofil, kan det større frafald blandt de ældste ældre skyldes, at denne gruppe har mindre psykisk overskud til at svare på spørgeskemaer og blive ved med at deltage i undersøgelser end yngre, eksempelvis grundet fysiske skavanker eller sociale omstændigheder.

Estimaterne for køn og partner er ikke signifikante i modellen, men etnisk herkomst derimod har betydning for sandsynligheden for bortfald. Det fremgår af figur 5.4, at sandsynligheden for bortfald er 11 procentpoint højere blandt indvandrere end blandt etnisk danske respon-denter.

FIGUR 5.4

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt på etnisk herkomst.



Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Analysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: alder, køn, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknøytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.

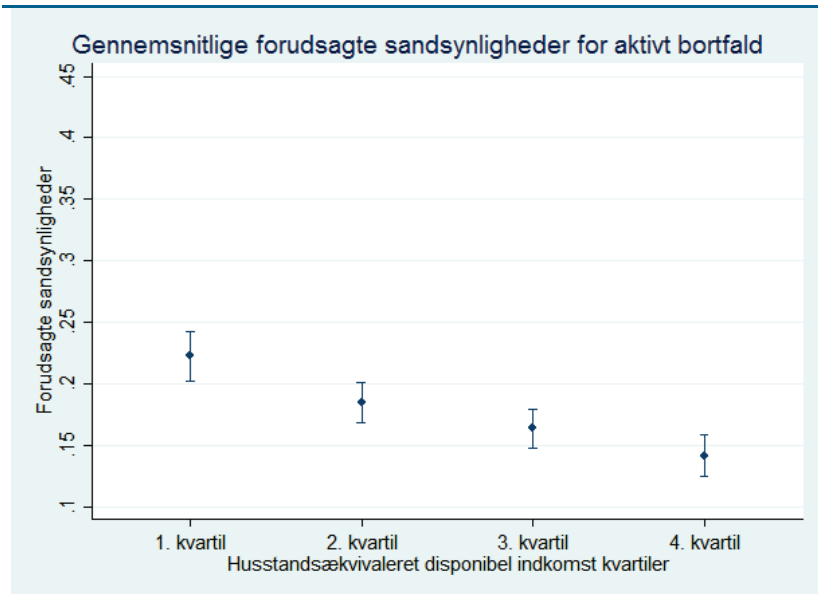
Kilde: Ældredatabasens bruttostikprøve og variabelen for etnicitet i etype fra GRUND-registret.

SOCIOØKONOMISK STATUS

Figur 5.5 viser den estimerede sandsynlighed for bortfald mellem tredje og fjerde bølge blandt personer fra husstande i første, anden, tredje og fjerde indkomstkvartil. Det fremgår, at personer fra de laveste indkomstgrupper har højere sandsynlighed for bortfald end personer fra de højeste indkomstgrupper. Blandt den fattigste fjerdedel af analysepopulationen er sandsynligheden for bortfald 8 procentpoint højere end blandt den rigeste fjerdedel. Denne gradient findes, uanset at der kontrolleres for fødselsårgang, køn, helbredsindikatorer osv.

FIGUR 5.5

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt på kvartiler for disponibel indkomst.



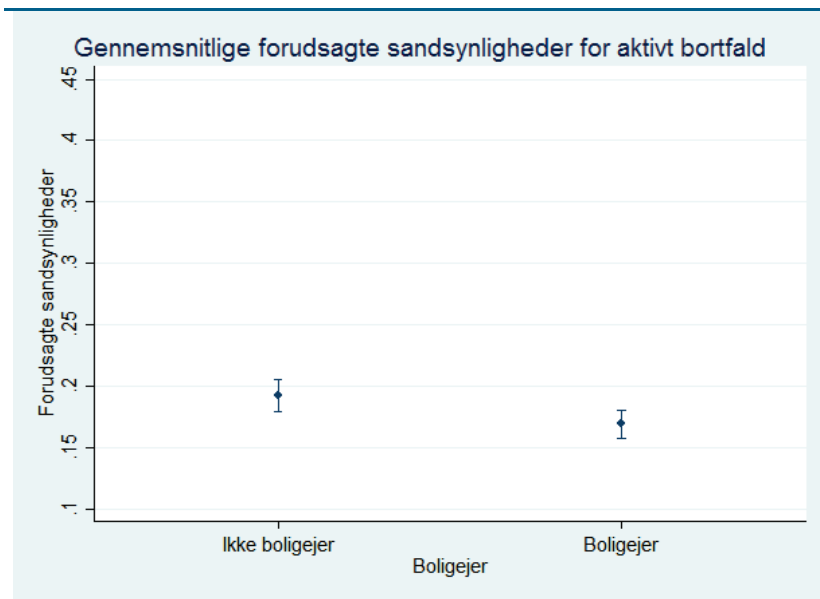
Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Analysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: fødselskø, køn, etnisk herkomst, partner, flytning, boligforhold, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.

Kilde: Ældredatabasens bruttostikprøve og hfauss fra GRUND-registret.

Figur 5.6 viser sandsynligheden for bortfald fordelt på boligforhold. At være lejer, andelshaver mv. frem for boligejer er forbundet med en højere sandsynlighed for bortfald. Sandsynligheden for bortfald er således omkring 2 procentpoint højere for lejere end for boligejere. Da der er kontrolleret for flytning, skyldes denne sammenhæng *ikke*, at lejere flytter oftere end boligejere.

FIGUR 5.6

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt på boligejere og lejere.



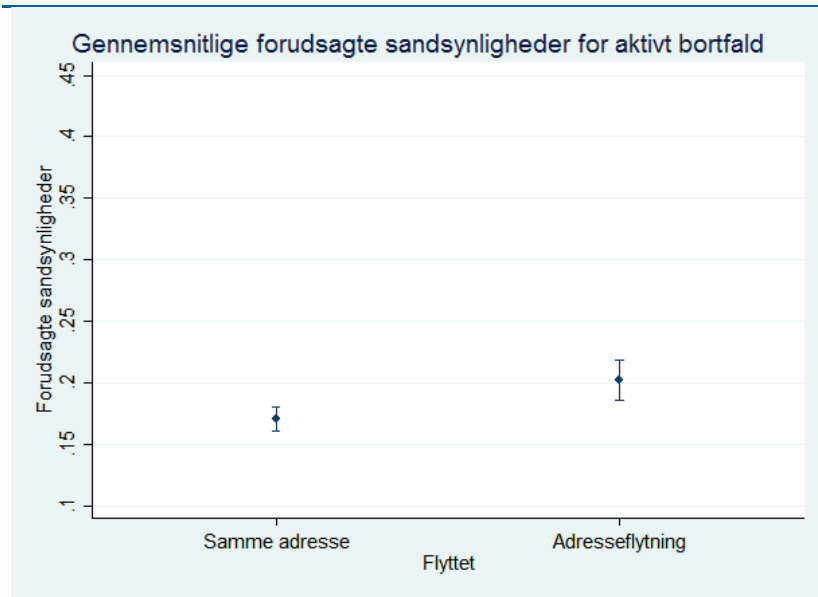
Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Analysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: fødselskorte, køn, etnisk herkomst, partner, flytning, disponibel indkomst, arbejdsmarkedstilknytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.

Kilde: Ældredatabasens bruttostikprøve samt ovskej07 fra GRUND-registret.

Figur 5.7 viser netop sandsynligheden for bortfald for personer, der har skiftet adresse siden seneste bølge og altså dermed antages at være flyttet. Flytning synes at prædiktere bortfald. Således har personer, der er flyttet siden seneste bølge, omkring 3 procentpoint højere sandsynlighed for bortfald, end personer, der ikke er flyttet.

FIGUR 5.7

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt efter, om respondenter er flyttet siden tredje bølge.



Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Analysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: fødselskohorte, køn, etnisk herkomst, partner, disponibel indkomst, boligforhold, arbejdsmarkedstilknøytning, psykiatrisk diagnose, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.

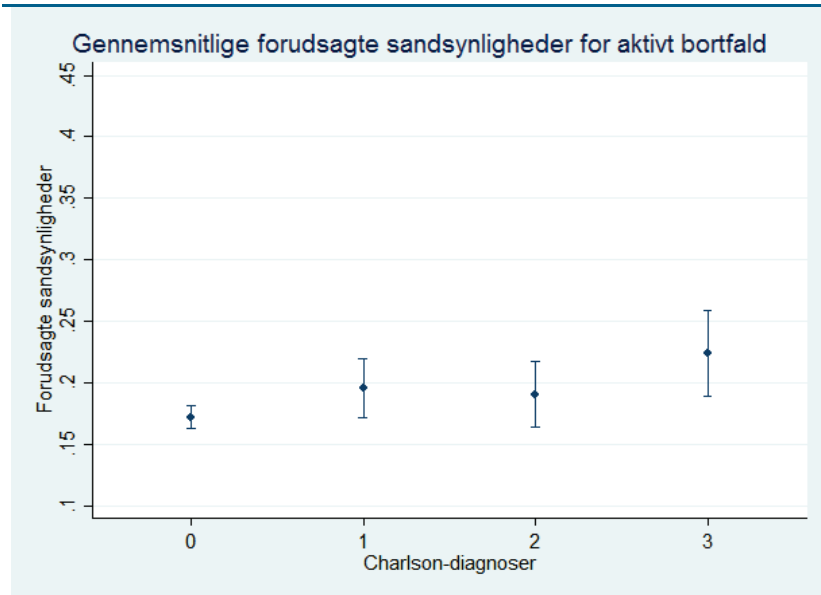
Kilde: Ældredatabasens bruttostikprøve samt bobikom fra GRUND-registret.

HELBRED

Bortfald kan også tænkes at hænge sammen med et forværret helbred siden seneste bølge. Dette undersøger vi nærmere her ved at se på hospitalsindlæggelser i perioden mellem tredje og fjerde bølge. Figur 5.8 viser den estimerede sandsynlighed for bortfald blandt følgende grupper: personer uden en hospitalsindlæggelse siden den seneste bølge (Charlson indeks = 0), personer med en hospitalsindlæggelse, gradueret efter diagnosernes alvorlighed (charlson = 1 til 3). Der er udelukkende en signifikant forskel på sandsynligheden for bortfald for den alvorligste score på Charlson indekset. Hos personer uden en hospitalsindlæggelse er sandsynligheden for bortfald ca. 5 procentpoint lavere end blandt personer med en høj score på Charlson's comorbidity indeks. Det er dog et fåtal, der har en alvorlig score på Charlson indekset (færre end 6 pct.).

FIGUR 5.8

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt efter Charlson-diagnoser.



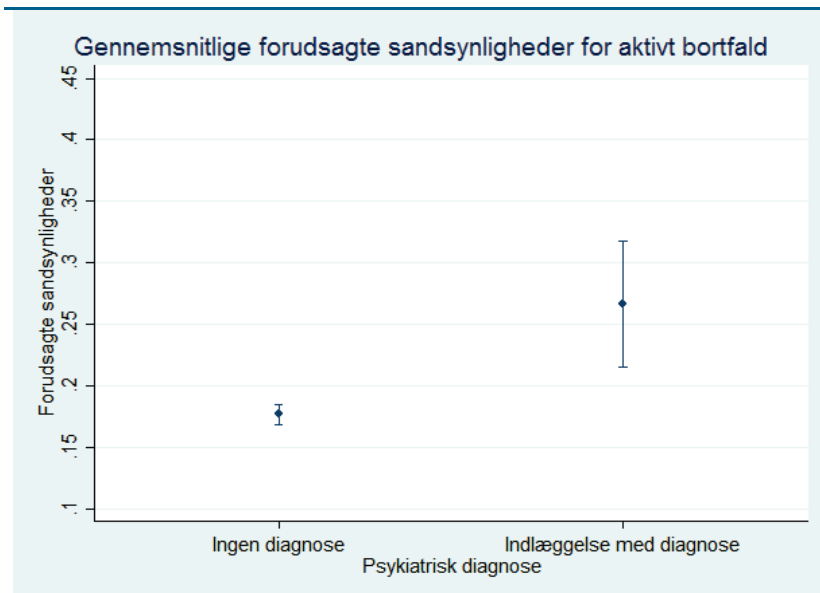
Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Analysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: fødselskohorte, køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknøytning, psykiatrisk diagnose.

Kilde: Ældredatabasens bruttostikprøve samt adiag fra Landspatient-registret.

Figur 5.9 viser den estimerede sandsynlighed for bortfald blandt følgende grupper: personer uden en hospitalskontakt med en psykiatrisk diagnose siden den seneste bølge (0), personer med en hospitalsindlæggelse og en psykiatrisk diagnose (1). Selvom det er ganske få respondenter (færre end 3 pct.), der har oplevet en indlæggelse med en psykiatrisk diagnose mellem tredje og fjerde bølge, er der en signifikant forskel på sandsynligheden for bortfald mellem grupperne. Hos personer uden en psykiatrisk diagnose er sandsynligheden for bortfald ca. 9 procentpoint lavere end blandt personer med en psykiatrisk diagnose.

FIGUR 5.9

Forudsagt sandsynlighed for bortfald mellem Ældredatabasens tredje og fjerde bølge, fordelt efter psykiatrisk diagnose.



Anm.: Postestimation efter logit-model. Modellens outcome er aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge. Analysen omfatter udelukkende 'egnede', dvs. eksklusivt 'passivt bortfald'. Der er kontrolleret for: fødselskorte, køn, etnisk herkomst, partner, flytning, disponibel indkomst, boligforhold, arbejdsmarkedstilknytning, Charlson indeks baseret på diagnoser stillet ved evt. hospitalsindlæggelse.

Kilde: Ældredatabasens bruttostikprøve samt adilag fra Landspatient-registret.

PANELETS REPRÆSENTATIVITET

Vi har i det foregående undersøgt omfanget og karakteren af bortfald mellem Ældredatabasens bølger. I det følgende undersøger vi, hvilken betydning det akkumulerede bortfald har for repræsentativiteten i Ældredatabasens 'panel'. Panelet består af de respondenter, der har besvaret alle fire bølger (n = 2.406). Panelet er interessant, fordi det udgør den gruppe af respondenter, som vi gennem Ældredatabasen har data for

over en 15-årig periode, dvs. i 1997, 2002, 2007 og 2012. Vi undersøger panelets repræsentativitet ved at sammenligne fordelingen blandt respondenter i panelet med fordelingen i målpopulationen. Målpopulationen består som bekendt af alle danske ældre fra de udvalgte kohorter. Data for målpopulationen stammer fra Statistikbanken (Danmarks Statistik, 2016). Følgende indikatorer er tilgængelige i Statistikbanken for målpopulationen: alder, køn, herkomst, partner, arbejdsmarkedstilknytning. Vi tester for hver af disse indikatorer, om der er forskel på fordelingen blandt respondenter og målpopulationen ved hjælp af en chi²-test, og vi afrapporterer med stjerne signifikansniveauet for p-værdien. Opgørelsen er baseret på den seneste bølge, hvor panelrespondenterne deltog, det vil sige i 2012.

MÅLPOPULATION OG PANEL

I tabel 5.2 viser vi aldersfordelingen blandt respondenter i Ældredatabasens panel i 2012 og blandt målpopulationen, som består af ældre i den danske befolkning i samme år (fra de samme seks kohorter). Det fremgår, at aldersfordelingen i panelet ikke er signifikant forskellig fra aldersfordelingen i målpopulationen, hvilket betyder, at vi har aldersmæssig repræsentativitet i Ældredatabasens panel.

TABEL 5.2

Aldersfordeling blandt målpopulation og panelrespondenter samt test af forskelle (chi²-test) i 2012. Procent.

År Alder	2012	
	Målpopulation	Panelrespondenter
67	35,9	36,7
72	23,7	24,4
77	17,5	17,8
82	12,2	11,3
87	7,8	7,2
92	3,0	2,7
Total	100	100
N/n	191.706	2.406
Signifikans for chi ² -test		i.

I tabel 5.3 viser vi, hvordan henholdsvis panelrespondenter i Ældredata-basen og målpopulationen fordeler sig efter køn. Det fremgår, at kønsfordelingen i panelet ikke er signifikant forskellig fra kønsfordelingen i målpopulationen, hvilket betyder, at vi har kønsmæssig repræsentativitet i Ældredata-basens panel.

TABEL 5.3

Fordeling efter køn blandt målpopulation og panelrespondenter, samt test af forskelle (χ^2 -test) i 2012. Procent.

År	2012	
	Målpopulation	Panelrespondenter
Køn		
Mænd	45,2	45,1
Kvinder	54,8	54,9
Total	100	100
N/n	191.706	2.406
Signifikans for χ^2 -test		i.

I tabel 5.4 viser vi, hvordan henholdsvis målpopulationen og panelrespondenter i Ældredata-basen fordeler sig efter herkomst. Det fremgår, at der er signifikant forskel på fordelingen i Ældredata-basens panel og i målpopulationen. Således er indvandrere underrepræsenteret i Ældredata-basens panel, mens etnisk danske ældre er overrepræsenteret. Dette skyldes, at panelet tager udgangspunkt i respondenter fra Ældredata-basens første bølge, hvor ældre med ikke-vestlig oprindelse endnu ikke blev udtrukket med dobbelt sandsynlighed.

TABEL 5.4

Fordeling efter herkomst blandt målpopulation og panelrespondenter samt test af forskelle (χ^2 -test) i 2012. Procent.

År	2012	
	Målpopulation	Panelrespondenter
Herkomst		
Indvandrer	4,2	1,2
Dansk eller efterkommer	95,8	98,8
Total	100	100
N/n	191.706	2.406
Signifikans for χ^2 -test		***

I tabel 5.5 viser vi, hvordan henholdsvis målpopulationen og panelrespondenter i Ældredata-basen fordeler sig efter familietype. Det fremgår, at fordelingen af ældre med partner og ældre uden partner i panelet ikke er signifikant forskellig fra fordelingen i målpopulationen, hvilket betyder, at der er repræsentativitet i Ældredata-basens panel.

TABEL 5.5

Fordeling efter familietype blandt målpopulation og panelrespondenter, samt test af forskelle (χ^2 -test) i 2012. Procent.

År	2012	
	Målpopulation	Panelrespondenter
Familietype		
Partner eller anden familietype	60,0	61,4
Enlig	40,0	38,6
Total	100	100
N	191.706	2.406
Signifikans for χ^2 -test		i.

I tabel 5.6 viser vi, hvordan henholdsvis respondenter i Ældredatabasens panel og målpopulationen fordeler sig efter arbejdsmarkedstilknytning. Data om arbejdsmarkedstilknytning, som er udtrukket fra Statistikbanken, omfatter alle 67-årige og ældre, hvorfor data ikke perfekt repræsenterer målpopulationen for Ældredatabasens panel, der jo omfatter udvalgte aldersgrupper (67, 72, 77, 82, 87 og 92 år i 2012). Alligevel kan en sammenligning give et præj om Ældredatabasens repræsentativitet, hvad angår arbejdsmarkedstilknytning blandt respondenter i Ældredatabasens fjerde bølge. Det fremgår, at der er signifikant forskel på fordelingen blandt respondenter i Ældredatabasens panel og i målpopulationen, idet beskæftigede er overrepræsenteret blandt panelets respondenter, mens personer, der står uden for arbejdsmarkedet, er underrepræsenteret blandt respondenter. Det er dog ikke sikkert, om denne forskel kan skyldes, at målpopulationen omfatter flere aldersgrupper end de, som indgår i Ældredatabasens panel.

TABEL 5.6

Fordeling efter arbejdsmarkedstilknytning blandt målpopulation og panelrespondenter samt test af forskelle (χ^2 -test) i 2012. Procent.

År	2012	
	Målpopulation	Panelrespondenter
Arbejdsmarkedstilknytning		
Beskæftigede	8,7	11,3
Arbejdsløse	0,0	0,0
Uden for arbejdsstyrken	91,3	88,7
Total	100	100
N	847.566	2.402
Signifikans for χ^2 -test		***

Anm.: Tilknytning til arbejdsmarkedet er for målpopulationen opgjort blandt 67-årige og ældre, hvilket betyder, at fordelingen ikke er direkte sammenlignelig med panelpopulationen, der blot inkluderer følgende aldersgrupper 67-, 72-, 77-, 82-, 87- og 92-årige.

OPSUMMERING OG KONKLUSION

RESPONS OG BORTFALD I ÆLDREDATABASEN

Ud fra de undersøgte parametre – svarprocent, non-respons og bortfald – vurderer vi, at Ældredatabasen har en fornuftig kvalitet som forløbsundersøgelse. Langt de fleste adspurgte respondenter vil gerne deltage, og størstedelen vil deltage gentagne gange. Ældredatabasens respondenter er et rimeligt repræsentativt snit af befolkningen i målgruppen. Der er mindre skævheder i forhold til alder, socioøkonomisk position og helbred, og som ved alle survey-undersøgelser er de allersvageste grupper svære at få fat i. Der er dog noget, man fremadrettet kunne gøre for at forbedre databasens repræsentativitet.

Det, at man i Ældredatabasens design har valgt at undlade at genkontakte nægttere fra tidligere bølger, har konsekvenser i forhold til fastholdelse af respondenter. En ændring af dataindsamlingens design, hvor man i højere grad forsøger at overtale nægttere fra tidligere runder til at deltage, vil imidlertid være en omkostningstung strategi. I stedet er Ældredatabasen designet med et opfyldnings-sample, som sikrer, at bortfaldne fra én runde erstattes af nye respondenter i den næste runde.

Ligeledes giver de systematiske forskelle på respondenter og ikke-respondenter, særligt hvad angår socioøkonomisk position, anledning til at anbefale, at der udarbejdes vægte til hver enkelt af Ældredatabasens bølger. Vægtning vil være vigtigt i deskriptive opgørelser over ældres vilkår, men mindre vigtigt i panelanalyser, som har fokus på kausalitet.

SVARPROCENTER I ÆLDREDATABASEN

Vi evaluerede Ældredatabasens svarprocent baseret på tre forskellige beregningsmetoder, som kan anvendes i panelbaserede spørgeskemaundersø-

søgelser. Vi beregnede hhv. den simple svarprocent, tværsnittets svarprocent og panelets svarprocent.

Anvender vi den simple svarprocent, som fortæller noget om dataindsamlingens succes, må Ældredatabasen sige at have en god kvalitet. 7 ud af 10 adspurgte respondenter vælger at deltage i Ældredatabasens første bølge, og omkring 8 ud af 10 respondenter fastholdes fra én bølge til den næste. Dette er et flot resultat, sammenlignet med eksempelvis den europæiske Survey on Health Ageing and Retirement in Europe (SHARE).

Anvender vi i stedet de beregninger af tværsnittets og panelets svarprocenter, som gør det muligt at sammenligne med Cheshire (2011)'s beregninger for andre internationale databaser, får vi derimod noget lavere svarprocenter i Ældredatabasen end i eksempelvis det amerikanske HRS. Dette skyldes, at disse beregninger af svarprocenter omfatter *alle* udtrukne, der var en del af den oprindelige stikprøve i den første bølge, uanset om de også blev kontaktet i efterfølgende bølger. Denne type svarprocenter fortæller således noget om Ældredatabasens repræsentativitet i forhold til målpopulationen, uanset dataindsamlingens design. Ældredatabasens kvalitet kan på dette område siges at lide under, at 'nægttere' fra én bølge *ikke* forsøges overtalt til at deltage i de efterfølgende bølger. Alligevel ligger Ældredatabasens svarprocenter på et acceptabelt niveau sammenlignet med eksempelvis det engelske Longitudinal Study on Ageing (ELSA).

NON-RESPONS

Der er signifikante forskelle på respondenternes og ikke-respondenternes fordeling på henholdsvis sociodemografiske og socioøkonomiske indikatorer. Eksempelvis er der forskel på aldersfordelingen mellem respondenter og ikke-respondenter, hvor de yngste og ældste fødselskohorter generelt er overrepræsenteret blandt ikke-respondenter. Ligeledes er personer, der er trådt ud af arbejdsmarkedet, overrepræsenteret blandt ikke-respondenter, og det samme gælder for personer med en kort uddannelsesbaggrund og personer i de laveste indkomstkvarterer. Også lejere er overrepræsenteret blandt ikke-respondenter i modsætning til boligejere. Ligeledes adskiller de to grupper sig på udvalgte helbredsindikatorer, om end der her er tale om små forskelle.

At der er socioøkonomiske forskelle på respondenter og ikke-respondenter giver anledning til bekymring – hvis non-respons og bortfald er systematisk korreleret med socioøkonomisk status, er der sandsynlighed for selektionsbias. For at undersøge, hvorvidt de deskriptive forskelle genfindes i et multivariat design, gennemførte vi en række modeller, som undersøger, hvorvidt observerbare indikatorer systematisk prædikterer non-respons.

Resultaterne af den multivariate analyse viser, at modellens observerbare karakteristika kun forklarer en ganske lille del af variansen i

respons/non-respons. De primære årsager til non-respons kan vi således ikke observere. Følgende faktorer er dog systematisk forbundet med en højere sandsynlighed for non-respons: Mand (vs. kvinde); Ældste og yngste kohorter (vs. mellemste); Indvandrere (vs. dansk eller efterkommer); Laveste indkomstkvarter (vs. mellemste); Forladt arbejdsmarkedet (vs. i beskæftigelse); Lejere (vs. boligejere); Psykiatrisk diagnose (vs. ingen diagnose). Der er dog tale om begrænsede effektstørrelser, hvorfor vi ikke er voldsomt bekymrede for selektion i forhold til disse forhold. En konklusion er dog, at det ofte vil være nødvendigt at kontrollere for socioøkonomisk status og sociodemografi, når der foretages analyser på baggrund af Ældredatabasens enkelte bølger. Ligeledes giver resultaterne anledning til at overveje at indarbejde vægte i databasen.

BORTFALD

En opgørelse over omfanget af bortfald viste, at 35 pct. af Ældredatabasens oprindelige respondenter fra den første bølge var afgået ved døden før kontakttidspunktet for den fjerde bølge ('passivt bortfald'). I samme 15-årige periode var 23 pct. af de oprindelige respondenter faldet fra pga. andre årsager ('aktivt bortfald').

Ældredatabasen er designet til at tage hånd om begge disse typer af bortfald ved at tilføre et opfyldnings-sample i hver bølge som erstatning for bortfaldne. Bortfald fra én bølge til den næste er imidlertid stadig en udfordring, såfremt man ønsker at undersøge udviklingen over tid. Hvor passivt bortfald afspejler det naturlige bortfald i en aldrende befolkningsgruppe, giver det aktive bortfald anledning til bekymring om databasens repræsentativitet over tid. Vi undersøgte årsagerne til aktivt bortfald mellem Ældredatabasens tredje og fjerde bølge i et multivariat design, hvor vi ekskluderede passivt bortfald.

Vi fandt, ligesom i analysen af non-respons, at modellen havde en ganske lav forklaringskraft, og vi har således ikke held med at forklare meget af variansen i bortfald med observerbare karakteristika. Resultaterne viser dog, at følgende observerbare karakteristika er systematisk korelateret med en højere sandsynlighed for aktivt bortfald mellem bølger: Ældste (vs. yngste); Indvandrere (vs. dansk); Laveste indkomstkvarter (vs. højeste); Lejere (vs. boligejere); Alvorlig komorbiditet (vs. ingen indlæggelser); Psykiatrisk diagnose (vs. ingen indlæggelser); samt at have flyttet adresse siden seneste bølge. Effektstørrelserne er dog små, hvorfor der ikke er grund til alvorlig bekymring omkring databasens repræsentativitet over tid. Alligevel kan et systematisk bortfald relateret til lav socioøkonomisk status og ringe helbred være med til at skjule en negativ udvikling over tid. En konklusion på bortfaldsanalysen er derfor, at det i panelanalyser, som har fokus på kausalitet, er nødvendigt at tage højde for socioøkonomisk status og helbred. Ligeledes anbefaler vi, at der udarbejdes

vægte til hver af Ældredatabasens bølger. Vægte vil være relevante i deskriptive opgørelser af ældres vilkår.

REPRÆSENTATIVITET

For at undersøge, hvorvidt non-respons og bortfald har konsekvenser for databasens repræsentativitet, sammenlignede vi fordelingen blandt respondenter i Ældredatabasen med fordelingen i målpopulationen på en række indikatorer: alder, køn, herkomst, familietype og arbejdsmarkeds-tilknytning. Sammenligningen viser, at selvom der er signifikante forskelle i fordelingen blandt henholdsvis respondenter og målpopulation for de fleste af disse indikatorer, så er der tale om størrelsesmæssigt ganske små forskelle. Ældredatabasen kan således trygt anvendes til deskriptive formål. Når der i fremtiden udarbejdes vægte til databasen, vil det dog være relevant at undersøge, hvorvidt vægtning medfører mindre ændringer i deskriptive resultater.

BILAG

BILAG 1 KODNING AF REGISTERBASEREDE HELBREDSINDIKATORER

Vi beregner for hver person i bruttostikprøven en tilnærmet Charlsons Comorbidity Indeks score, baseret på aktionsdiagnoser stillet i forbindelse med de seneste 5 års kontakter til sygehusvæsnet. Det samlede indeks beregnes ved at lægge indeks-scoren for hver enkelt diagnose sammen. Bilagstabel B1.1 viser ICD-10 koder og tilhørende Charlson Comorbidity Score (jf. Johansen & Fynbo, 2011). Da imidlertid kun få personer i bruttostikprøven scorer højt på Charlsons Comorbidity indeks, grupperer vi i denne analyse efterfølgende indekset i fire grupper 0, 1, 2 og 3, hvor 0 indikerer en indeks-score på 0, 1 indikerer en indeks-score på 1, 2 indikerer en indeks-score på 2, og 3 indikerer en indeksscore på mere end 2.

BILAGSTABEL B1.1

ICD-10 diagnoser til Charlson Comorbidity Indeks (engelsk).

Charlson Comorbidity Category	ICD-10	Charlson Comorbidity Indeks Score
Myocardial infarction	I21; I22; I23	I
Congestive heart failure	I50; I11.0; I13.0; I13.2	I
Peripheral vascular disease	I70; I71; I72; I73; I74; I77	I
Cerebrovascular disease	I60–I69; G45; G46	I
Dementia	F00–F03; F05.1; G30	I
Chronic pulmonary disease	J40–J47; J60–J67; J68.4; J70.1; J70.3; J84.1; J92.0; J96.1; J98.2; J98.3	I
Connective tissue disease	M05; M06; M08; M09; M30; M31; M32; M33; M34; M35; M36; D86	I
Ulcer disease	K22.1; K25–K28	I
Mild liver disease	B18; K70.0–K70.3; K70.9; K71; K73; K74; K76.0	I
Diabetes type 1;	E10.0; E10.1; E10.9	I
Diabetes type 2	E11.0; E11.1; E11.9	I
Hemiplegia	G81; G82	2
Moderate to severe renal disease	I12; I13; N00–N05; N07; N11; N14; N17–N19; Q61	2
Diabetes with end organ damage type 1 and type 2	E10.2–E10.8 E11.2–E11.8	2
Any tumor	C00–C75	2
Leukemia	C91–C95	2
Lymphoma	C81–C85; C88; C90; C96	2
Moderate to severe liver disease	B15.0; B16.0; B16.2; B19.0; K70.4; K72; K76.6; I85	3
Metastatic solid tumor	C76–C80	6
Aids	B21–B24	6

Kilde: Denne oversigt og kodningen stammer fra Johansen & Fynbo (2011).

Bilagstabel B1.2 viser ICD-10 koder for diagnoser, der kan betegnes som psykiske diagnoser. Vi koder en variabel for en psykiatrisk diagnose, såfremt en person i bruttostikprøven har haft en hospitalskontakt med én af disse diagnoser i løbet af de seneste 5 år.

BILAGSTABEL B1.2

ICD-10 koder for diagnose ved kontakt til psykiatrisk facilitet (engelsk).

Psychiatric diagnosis group	ICD-10
Organic, including symptomatic, mental disorders	F00-F09
Mental and behavioural disorders due to psychoactive substance use	F10-F19
Schizophrenia, schizotypal and delusional disorders	F20-F29
Affective disorders	F30-F39
Neurotic, stress-related and somatoform disorders	F40-F49
Behavioural syndromes associated with physiological disturbance and physical factors	F50-F59
Disorders of adult personality and behavior	F60-F69
Mental retardation	F70-F79
Disorders of psychological development	F80-F89
Behavioural and emotional disorders with onset usually occurring in childhood and adolescence	F90-F98

BILAG 2 TEST AF FORSKELLE

BILAGSTABEL B2.1

Respondenter og ikke-respondenter fordelt på baggrundskarakteristika. Antal og procent.

	1997		2002		2007		2012	
	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter
<i>Sociodemografiske faktorer</i>								
<i>Kohorte</i>		***		***		***		***
1920	403	687	177	612	104	319	67	97
	16,28	11,72	10,97	7,46	3,15	3,31	1,99	0,97
1925	333	753	179	777	223	582	138	311
	13,45	12,84	11,1	9,47	6,75	6,04	4,09	3,12
1930	303	840	177	922	239	818	177	543
	12,24	14,32	10,97	11,23	7,24	8,49	5,25	5,45
1935	365	927	140	1.129	304	1.065	254	830
	14,74	15,81	8,68	13,76	9,2	11,06	7,53	8,33
1940	464	1.084	213	1.324	357	1.392	323	1.175
	18,74	18,49	13,21	16,13	10,81	14,45	9,58	11,79
1945	608	1.573	304	1.870	558	2.010	430	1.845
	24,56	26,82	18,85	22,79	16,89	20,87	12,75	18,51
1950			423	1.573	717	1.782	506	1.777
			26,22	19,17	21,71	18,5	15,01	17,83
1955					801	1.665	573	1.782
					24,25	17,28	16,99	17,88
1960							904	1.605
							26,81	16,11
<i>Køn</i>		**		-		***		**
Mand	1.099	2.824	808	3.897	1.689	4.489	1.671	4.629
	44,39	48,16	50,09	47,48	51,14	46,6	49,56	46,45
Kvinde	1.377	3.040	805	4.310	1.614	5.144	1.701	5.336
	55,61	51,84	49,91	52,52	48,86	53,4	50,44	53,55

Tabellen fortsættes

BILAGSTABEL B2.1 FORTSAT

Respondenter og ikke-respondenter fordelt på baggrundskarakteristika. Antal og procent.

	1997		2002		2007		2012	
	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter
<i>Etnisk herkomst</i>		-		***		***		***
Dansk	2.427	5.781	1.562	8.079	2.707	9.059	2.749	9.280
	98,02	98,58	96,84	98,44	81,96	94,04	81,52	93,13
Indvandre	49	83	51	128	596	574	623	685
	1,98	1,42	3,16	1,56	18,04	5,96	18,48	6,87
<i>Partner</i>		***		***		***		***
Har ikke en partner	956	1.650	666	2.379	1.206	2.867	1.231	2.959
	38,61	28,14	41,29	28,99	36,51	29,76	36,51	29,69
Har en partner	1.520	4.214	947	5.828	2.097	6.766	2.141	7.006
	61,39	71,86	58,71	71,01	63,49	70,24	63,49	70,31
<i>Partner død inden for 5 år</i>		-		*		-		-
Nej	1.495	4.179	953	5.797	2.100	6.744	2.141	6.978
	94,44	95,3	93,25	94,89	95,45	95,29	96,14	95,28
Ja	88	206	69	312	100	333	86	346
	5,56	4,7	6,75	5,11	4,55	4,71	3,86	4,72
<i>Civilstatus</i>		***		***		***		***
Enke/Enkemand	410	874	301	1.275	446	1.406	410	1.331
	16,65	14,97	18,7	15,57	13,9	14,7	12,6	13,53
Skilt	365	637	227	922	454	1.175	518	1.257
	14,83	10,91	14,1	11,26	14,15	12,29	15,92	12,78
Gift	1.437	4.027	895	5.517	1.937	6.318	1.937	6.447
	58,37	68,98	55,59	67,36	60,36	66,06	59,53	65,53
Ugift	250	300	187	476	372	665	389	803
	10,15	5,14	11,61	5,81	11,59	6,95	11,95	8,16

Tabellen fortsættes

BILAGSTABEL B2.1 FORTSAT

Respondenter og ikke-respondenter fordelt på baggrundskarakteristika. Antal og procent.

	1997		2002		2007		2012	
	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter
<i>Socioøkonomiske faktorer</i>								
<i>Ækvivaleret husholdningsindkomst</i>		***		***		***		***
1. kvartil	926	1.646	577	2.079	1.003	2.192	989	1.914
	37,61	28,19	35,79	25,37	30,39	22,79	29,34	19,23
2. kvartil	623	1.402	428	2.045	869	2.385	787	2.282
	25	24	26,55	24,95	26,33	24,8	23,35	22,93
3. kvartil	452	1.263	288	1.896	705	2.414	831	2.742
	18,36	21,63	17,87	23,13	21,36	25,1	24,65	27,55
4. kvartil	461	1.527	319	2.176	723	2.626	764	3.014
	18,72	26,16	19,79	26,55	21,91	27,31	22,66	30,29
<i>Højeste uddannelse</i>		***		***		***		***
Grundskole	1.131	2.419	723	3.086	1.162	3.257	1.141	2.836
	45,68	41,25	44,82	37,6	35,18	33,81	33,84	28,46
Gymnasial/erhverv	729	1.863	485	2.928	1.124	3.465	1.100	3.836
	29,44	31,77	30,07	35,68	34,03	35,97	32,62	38,49
KVU/MVU	166	726	171	1.241	370	1.796	433	2.147
	6,7	12,38	10,6	15,12	11,2	18,64	12,84	21,55
LVU	63	215	47	329	114	496	150	617
	3	4	2,91	4,01	3,45	5,15	4,45	6,19
Uoplyst ¹	387	641	187	623	533	619	548	529
	15,63	10,93	11,59	7,59	16,14	6,43	16,25	5,31
<i>Ejer sin egen bolig</i>		***		***		***		***
Ikke boligejer	1.506	2.918	920	3.894	1.756	4.204	1.712	3.925
	61,17	49,98	57,07	47,51	53,21	43,71	50,79	39,44
Boligejer	956	2.920	692	4.302	1.544	5.413	1.659	6.027
	38,83	50,02	42,93	52,49	46,79	56,29	49,21	60,56

Tabellen fortsættes

BILAGSTABEL B2.1 FORTSAT

Respondenter og ikke-respondenter fordelt på baggrundskarakteristika. Antal og procent.

	1997		2002		2007		2012	
	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter
<i>Tilknytning til arbejdsmarkedet</i>		***		***		-		**
Beskæftigede	917	2.623	643	3.676	1.430	4.230	1.359	4.278
Arbejdsløse	37,25	44,93	39,89	44,85	43,33	43,98	40,31	42,99
Uden for arbejdsmarkedet	60	126	33	146	31	62	72	160
	2,44	2,16	2,05	1,78	0,94	0,64	2,14	1,61
	1.485	3.089	936	4.374	1.839	5.325	1.940	5.514
	60,32	52,91	58,06	53,37	55,73	55,37	57,55	55,41
<i>Folkepension</i>		***		***		***		***
Modtager andet med folkepension	1.931	4.854	1.307	7.155	2.730	8.638	2.808	9.175
Modtager kun folkepension	78,43	83,14	81,18	87,36	85,07	90,32	86,29	93,26
	531	984	303	1.035	479	926	446	663
	21,57	16,86	18,82	12,64	14,93	9,68	13,71	6,74
<i>Flyttet siden sidste interview</i>		***		***		**		***
Ikke flyttet	1.297	2.672	759	3.448	1.287	3.518	2.801	8.722
Flyttet	52,38	45,57	47,06	42,01	39,02	36,52	83,07	87,53
	1.179	3.192	854	4.759	2.011	6.114	571	1.243
	47,62	54,43	52,94	57,99	60,98	63,48	16,93	12,47

Tabellen fortsættes

BILAGSTABEL B2.1 FORTSAT

Respondenter og ikke-respondenter fordelt på baggrundskarakteristika. Antal og procent.

	1997		2002		2007		2012	
	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter
<i>Helbredsmæssige faktorer</i>								
<i>Psykiatrisk diagnose</i>		**		***		**		***
Ingen diagnose	2.401	5.745	1.570	8.138	3.269	9.580	3.337	9.927
	97	98	97,33	99,16	98,97	99,45	98,96	99,62
Psykiatrisk diagnose	75	119	43	69	34	53	35	38
	3,03	2,03	2,67	0,84	1,03	0,55	1,04	0,38
<i>CCI indeks</i>								
0	2.088	4.951	1.349	7.156	2.813	8.372	2.874	8.616
	84,33	84,43	83,63	87,19	85,17	86,91	85,23	86,46
1	189	493	149	547	252	599	221	622
	7,63	8,41	9,24	6,67	7,63	6,22	6,55	6,24
2	130	306	58	297	130	390	142	428
	5,25	5,22	3,6	3,62	3,94	4,05	4,21	4,3
3	69	114	57	207	108	272	135	299
	2,79	1,94	3,53	2,52	3,27	2,82	4	3

Tabellen fortsættes

BILAGSTABEL B2.1 FORTSAT

Respondenter og ikke-respondenter fordelt på baggrundskarakteristika. Antal og procent.

	1997		2002		2007		2012	
	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter	Ikke-respondent	Respondenter
<i>Patienttype</i>								
Skadestuepatient		-		-		**		*
Nej	2.360	5.604	1.595	8.082	3.241	9.517	3.303	9.812
	95,32	95,57	98,88	98,48	98,12	98,8	97,95	98,46
Ja	116	260	18	125	62	116	69	153
	4,68	4,43	1,12	1,52	1,88	1,2	2,05	1,54
Indlæggelse		-		-		-		-
Nej	2.092	5.005	1.521	7.808	9.329	3.137	3.328	9.836
	84,49	85,35	94,3	95,14	95	94,97	98,7	98,71
Ja	384	859	92	399	491	166	44	129
	15,51	14,65	5,7	4,86	5	5,03	1,3	1,29
Ambulant		***		-		**		-
Nej	2.297	5.304	1.539	7.789	3.121	8.963	3.001	8.785
	92,77	90,45	95,41	94,91	94,49	93,04	89	88,16
Ja	179	560	74	418	182	670	371	1.180
	7,23	9,55	4,59	5,09	5,51	6,96	11	11,84
<i>Kontakt til sygehuse</i>								
0-10 henvendelser		-		**		-		-
	1.502	3.771	847	4.651	1.497	4.694	1.410	4.585
	65,96	67,4	58,33	61,03	52,31	52,32	47,8	49,37
11-20 henvendelser	436	1.067	306	1.671	662	2.185	745	2.426
	19,15	19,07	21,07	21,93	23,13	24,35	25,25	26,12
Flere end 20 henvendelser	339	757	299	1.299	703	2.093	795	2.276
	14,89	13,53	20,59	17,05	24,56	23,33	26,95	24,51

1. Manglende observationer i uddannelse er stærkt korreleret med alderdom og etnisk herkomst. Der er således et relativt højt antal manglende observationer på uddannelsesvariablen, hvorfor vi har valgt at tydeliggøre andelen af manglende observationer ved at inkludere en kategori i uddannelse, der hedder 'uplyst'.

Kilde: Bruttonsample fra Ældredatabasen og udvalgte variable fra Danmarks Statistik.

BILAG 3 UDVALGTE SOCIOØKONOMISKE VARIABLES FORDELINGER I MÅLPOPULATIONEN

BILAGSTABEL B3.1

Målpopulationens fordeling efter aldersgrupper, i årene 1997, 2002, 2007 og 2012. Andel (procent).

Aldersgruppe	Årstal			
	1997	2002	2007	2012
52 år	26,4	21,4	19,1	18,1
57 år	18,9	22,5	19,0	17,3
62 år	16,3	15,8	19,6	17,0
67 år	14,4	13,1	13,4	17,1
72 år	13,7	11,1	10,6	11,3
77 år	10,3	9,7	8,3	8,3
82 år	-	6,4	6,5	5,8
87 år	-	-	3,5	3,7
92 år	-	-	-	1,4
Total	100,0	100,0	100,0	100,0
Målpopulation, N	300.750	341.970	374.141	402.078

Kilde: Statistikbanken, Folketal 1. januar 2012 (DST, 2016).

BILAGSTABEL B3.2

Målpopulationens fordeling efter herkomst i årene 1997, 2002, 2007 og 2012. Andel (procent).

Herkomst	Årstal			
	1997	2002	2007	2012
Indvandrere	3,9	4,4	5,0	5,8
Dansk eller efterkommer	96,1	95,6	95,0	94,2
Total	100,0	100,0	100,0	100,0
Målpopulation, N	300.750	341.970	374.141	402.078

Kilde: Statistikbanken, Folketal 1. januar 2012 (DST, 2016).

BILAGSTABEL B3.3

Målpopulationens fordeling efter køn i årene 1997, 2002, 2007 og 2012. Andel (procent).

Køn	Årstal			
	1997	2002	2007	2012
Mænd	47,9	47,6	47,5	47,6
Kvinder	52,1	52,4	52,5	52,4
Total	100,0	100,0	100,0	100,0
Målpopulation, N	300.750	341.970	374.141	402.078

Kilde: Statistikbanken, Folketal 1. januar 2012 (DST, 2016).

BILAGSTABEL B3.4

Målpopulationens fordeling efter husstand i årene 1997, 2002, 2007 og 2012. Andel (procent).

Civilstand	Årstal			
	1997	2002	2007	2012
Partner eller anden familietype	73,7	71,7	69,6	68,6
Enlig	26,3	28,3	30,4	31,4
Total	100,0	100,0	100,0	100,0
Målpopulation	300.631	341.898	374.140	402.078

Kilde: Statistikbanken, Voksne, husstandstype, 1. januar 2012 (DST, 2016).

BILAG 4 GENNEMSNITLIGE MARGINALE EFFEKTER

BILAGSTABEL B4.1

De gennemsnitlige marginale effekter på sandsynligheden for non-respons på baggrund af sociodemografiske, socioøkonomiske og helbreds-mæssige karakteristika. Særskilt for hver bølge. På baggrund af logistisk regression.

	Bølge 1	Bølge 2	Bølge 3	Bølge 4
<i>Fødselsår</i>				
1920	-0,0269 [-0,07;0,01]	0,0185 [-0,01;0,05]	0,0139 [-0,03;0,05]	0,156 *** [0,08;0,23]
1925	-0,0663 *** [-0,11;-0,03]	0,00116 [-0,03;0,03]	0,0440 ** [0,01;0,08]	0,0663 ** [0,02;0,11]
1930	-0,0826 *** [-0,12;-0,05]	-0,0142 [-0,04;0,01]	0,000891 [-0,03;0,03]	0,0167 [-0,02;0,05]
1935	-0,0422 * [-0,08;-0,01]	-0,0452 *** [-0,07;-0,02]	0,00544 [-0,02;0,03]	0,0107 [-0,02;0,04]
1940	Reference	Reference	Reference	Reference
1945	-0,0117 [-0,04;0,02]	0,0215 [-0,00;0,05]	0,0352 ** [0,01;0,06]	-0,00804 [-0,03;0,02]
1950		0,102 *** [0,07;0,13]	0,118 *** [0,09;0,15]	0,0445 ** [0,02;0,07]
1955			0,145 *** [0,12;0,17]	0,0784 *** [0,05;0,11]
1960				0,196 *** [0,16;0,23]
<i>Køn</i>				
Kvinde	Reference	Reference	Reference	Reference
Mand	0,0128 [-0,01;0,03]	0,0370 *** [0,02;0,05]	0,0524 *** [0,04;0,07]	0,0410 *** [0,03;0,06]
<i>Etnisk herkomst</i>				
Dansk	Reference	Reference	Reference	Reference
Indvandrer	0,0597 [-0,02;0,14]	0,0853 ** [0,02;0,15]	0,218 *** [0,19;0,25]	0,184 *** [0,16;0,21]

Tabellen fortsættes

BILAGSTABEL B4.1 FORTSAT

De gennemsnitlige marginale effekter på sandsynligheden for non-respons på baggrund af sociodemografiske, socioøkonomiske og helbreds-mæssige karakteristika. Særskilt for hver bølge. På baggrund af logistisk regression.

	Bølge 1	Bølge 2	Bølge 3	Bølge 4
<i>Partner</i>				
Ingen partner	Reference	Reference	Reference	Reference
Partner	-0,0649 *** [-0,09;-0,04]	-0,0508 *** [-0,07;-0,03]	-0,0428 *** [-0,06;-0,02]	-0,0194 * [-0,04;-0,00]
<i>Flytning inden for 5 år</i>				
Ikke flyttet	Reference	Reference	Reference	Reference
Flyttet	-0,0480 *** [-0,07;-0,03]	-0,0195 * [-0,03;-0,00]	-0,0181 * [-0,03;-0,00]	0,0147 [-0,01;0,04]
<i>Ækvivaleret disponibel indkomst</i>				
1. kvartil	0,0317 * [0,00;0,06]	0,0318 ** [0,01;0,05]	0,0367 ** [0,01;0,06]	0,0595 *** [0,04;0,08]
2. kvartil	Reference	Reference	Reference	Reference
3. kvartil	-0,0268 [-0,06;0,00]	-0,0369 *** [-0,06;-0,02]	-0,0381 *** [-0,06;-0,02]	-0,018 [-0,04;0,00]
4. kvartil	-0,0499 ** [-0,08;-0,02]	-0,0364 ** [-0,06;-0,01]	-0,0470 *** [-0,07;-0,02]	-0,0458 *** [-0,07;-0,02]
<i>Boligforhold</i>				
Ejer ikke egen bolig	Reference	Reference	Reference	Reference
Ejer egen bolig	-0,0710 *** [-0,09;-0,05]	-0,0361 *** [-0,05;-0,02]	-0,0356 *** [-0,05;-0,02]	-0,0375 *** [-0,05;-0,02]
<i>Tilknytning til arbejdsmarkedet</i>				
Beskæftiget	Reference	Reference	Reference	Reference
Arbejdsløs	0,0307 [-0,04;0,10]	-0,00497 [-0,05;0,04]	-0,00567 [-0,08;0,07]	-0,0137 [-0,06;0,03]
Uden for arbejdsmarkedet	0,0243 [-0,00;0,05]	0,0181 [-0,00;0,04]	0,0207 [-0,00;0,04]	0,0381 *** [0,02;0,06]

Tabellen fortsættes

BILAGSTABEL B4.1 FORTSAT

De gennemsnitlige marginale effekter på sandsynligheden for non-respons på baggrund af sociodemografiske, socioøkonomiske og helbreds-mæssige karakteristika. Særskilt for hver bølge. På baggrund af logistisk regression.

	Bølge 1	Bølge 2	Bølge 3	Bølge 4
<i>Psykiatriske diagnoser</i>				
Ingen diagnose	Reference	Reference	Reference	Reference
Diagnose	0,0281 [-0,04;0,09]	0,120 ** [0,04;0,20]	0,0597 [-0,03;0,15]	0,133 * [0,03;0,24]
<i>CCI indeks</i>				
0	Reference	Reference	Reference	Reference
1	-0,0427 * [-0,08;-0,01]	0,0315 * [0,00;0,06]	0,0341 * [0,00;0,06]	-0,00213 [-0,03;0,03]
2	-0,0261 [-0,07;0,02]	-0,0145 [-0,05;0,02]	-0,00195 [-0,04;0,04]	-0,00244 [-0,04;0,03]
3	0,0202 [-0,05;0,09]	0,0368 [-0,01;0,08]	0,0191 [-0,02;0,06]	0,0361 [-0,01;0,08]
N	8.300	9.808	12.911	
Pseudo R2	0,024	0,038	0,047	0,0541

Anm.: Variablene er målt separat ift. hver bølge. For bølge 1 er variablene således målt i 1997, for bølge 2 i 2002, for bølge 3 i 2007 og i bølge 4 i 2012.

Kilde: Ældredatabasens bruttostikprøve og udvalgte variable fra Danmarks Statistiks registre.

BILAGSTABEL B4.2

De gennemsnitlige marginale effekter på sandsynligheden for aktivt bortfald mellem tredje og fjerde bølge på baggrund af sociodemografiske, -økonomiske og helbredsmæssige karakteristika. På baggrund af logistisk regression.

	Aktivt bortfald
<i>Fødselsår</i>	
1920	0.178 *** [0.104; 0.253]
1925	0.0833 *** [0.0382; 0.128]
1930	0.0266 [-0.00834; 0.0615]
1935	0.0196 [-0.0109; 0.0501]
1940	Reference
1945	-0.0146 [-0.0397; 0.0105]
1950	0.0135 [-0.0151; 0.0421]
1955	0.0415 * [0.00799; 0.0749]
<i>Køn</i>	
Kvinde	Reference
Mand	-0.0153 [-0.0323; 0.00163]
<i>Etnisk herkomst</i>	
Dansk	Reference
Indvandrere	0.108 *** [0.0687; 0.147]
<i>Partner</i>	
Har ikke en partner	Reference
Har en partner	0.00469 [-0.0143; 0.0237]
<i>Flytning</i>	
Er ikke flyttet	Reference
Er flyttet	0.0303 ** [0.0106; 0.0500]
<i>Disponibel husstandsækvivaleret indkomst, kvartiler</i>	
1.kvartil	Reference
2.kvartil	-0.0371 ** [-0.0617; -0.0126]
3. kvartil	-0.0580 *** [-0.0849; -0.0310]
4. kvartil	-0.0794 *** [-0.109; -0.0500]
<i>Egen bolig</i>	
Har ikke egen bolig	Reference
Har egen bolig	-0.0249 ** [-0.0438; -0.00604]
<i>Tilknytning til arbejdsmarkedet</i>	
Beskæftigede	Reference
Arbejdsløse	-0.000851 [-0.0721; 0.0704]
Uden for arbejdsmarkedet	0.0158 [-0.00826; 0.0399]

Tabellen fortsættes

BILAGSTABEL B4.2 FORTSAT

De gennemsnitlige marginale effekter på sandsynligheden for aktivt bortfald mellem tredje og fjerde bølge på baggrund af sociodemografiske, -økonomiske og helbredsmæssige karakteristika. På baggrund af logistisk regression.

<i>Psykiske diagnoser</i>	
Har ikke en psykisk diagnose	Reference
Har en psykisk diagnose	0.0885 *** [0.0368; 0.140]
<i>CCI indeks</i>	
0	Reference
1	0.0238 [-0.00192; 0.0496]
2	0.0186 [-0.0100; 0.0472]
3	0.0516 ** [0.0151; 0.0881]
N	8600

Anm.: Outcome er aktivt bortfald, hvor 1 repræsenterer bortfald mellem 3. og fjerde bølge, mens 0 repræsenterer fastholdelse mellem tredje og fjerde bølge. Analysen omfatter kun egnede, derved er passivt bortfald sorteret fra.

Kilde: Danmarks Statistik

LITTERATUR

- Arbejderbevægelsens Erhvervsråd (2011): *Teknisk baggrundspapir om indkomst-definitioner og datagrundlag*.
- Cheshire, H., M. Ofstedal, S. Scholes & M. Schröder (2011): "A Comparison of Response Rates in the English Longitudinal Study of Ageing and the Health and Retirement Study". *Longitudinal and Life Course Studies*, 2(2), s. 127-144.
- Cao, H. & D. Hill (2005): *Activ versus Passive Panel Attrition. The Health and Retirement Study*. Michigan: Institute for Social Research, University of Michigan.
- De Luca, G. & F. Peracchi (2005): "Survey Participation in the first Wave of SHARE". *The Survey of Health, Ageing and Retirement in Europe—Methodology, MEA, Mannheim*, s. 88-104.
- Hsiao, C. (2014): *Analysis of Panel Data* (No. 54). UK: Cambridge University Press.
- Johansen, M.B. & C. Fynbo (2011): "Comparison of Charlson Comorbidity Indeks with SAPS and APACHE Scores for Prediction of Mortality Following Intensive Care". *Clinical Epidemiology*, 3, s. 203-211.
- Lauritzen, H.H. (2012): *Ældres ressourcer og behov i perioden 1997-2012. Nye- ste viden på baggrund af Ældredatabasen*. København: SFI – Det Nationale Forskningscenter for Velfærd, 14:10.
- Lynn, P. (2005): *Outcome Categories and Definitions of Response Rates for Panel Surveys and other Surveys Involving Multiple Data Collection Events from the same Units*. Unpublished manuscript. Colchester: University of Essex.
- (<http://www.nonresponse.org/uploadi/editor/1129102463Outcome%20categories%20same%20units.pdf>)

- Sonnega, A., J.D. Faul, M.B. Ofstedal, K.M. Langa, J.W. Phillips & D.R. Weir (2014): "Cohort Profile: The Health and Retirement Study (HRS)". *International Journal of Epidemiology*, 43(2), s. 576-585.
- Stephens, A., E. Breeze, J. Banks & J. Nazroo (2012): "Cohort Profile: The English Longitudinal Study of Ageing". *International Journal of Epidemiology*, dys168.
- Winkels, J.W. & S.D. Withers (2000): "Panel Attrition", in Rose, D. (ed.). *Researching Social and Economic Change: The Uses of Household Panel Studies* (Vol. 12). Psychology Press.