

Espen Dahl

Nytter det å følge opp sykmeldte?

Espen Dahl

Nytter det å følge opp sykmeldte?

© Forskningsstiftelsen Fafo 1996
ISSN 0804-5135

Innhold

Forord	4
Innledning	5
Evalueringsforskning	6
Prosjektdesign	7
Data og metode	9
Resultater	9
Oppsummering og diskusjon	15
Konklusjon	17

Forord

Formålet med prosjektet «Søkelys på sykmeldte» var å undersøke om det er mulig, med forholdsvis enkle og rimelige tiltak fra trygdekontolets side, å korte ned lengden på sykefraværene og dermed redusere sykepengeutbetalingene.

Prosjektet kom i stand på initiativ fra Fylkestrygdekontolet i Buskerud. I nært samarbeid med Fylkestrygdekontolet har fire trygdekontolet i Buskerud deltatt i prosjektet. Av grunner vi kommer til, rapporteres resultatene bare fra tre av dem her. Prosjektet har vært finansiert av Rikstrygdeverket, mens Forskningsstiftelsen Fafo har vært ansvarlig for den vitenskapelige delen av forsøket.

Takk til Bjarte Johannesen som har registrert dataene for kontrollgruppa, og til prosjektkonsulentene ved hvert av de fire deltakende trygdekontoletene. Likeledes fortjener referansegruppa en honnør for sitt engasjement. Den har vært ledet av Ingeborg Sølverud, først i samarbeid med Bjarte Johannesen og deretter med Erik Vaag, begge fra Fylkestrygdekontolet i Buskerud. For øvrig har referansegruppa hatt følgende medlemmer: kontorsjef Sverre Ulsberg, Rikstrygdeverket, rådgivende lege Ole B. Herland, Kongsberg trygdekontolet, Knut Erik Lie, Fylkesarbeidskontolet i Buskerud, og Bente Berge, Kai Becher, Berthe Ramstad og Anne Lise Marstein fra de fire trygdekontoletene som var med i forsøket. Endelig skal Fafo-kollega Axel West Pedersen ha takk for konstruktive kommentarer til analyseopplegg og tolkninger av resultatene.

Oslo, mai 1996

Espen Dahl

Innledning

Velferdsstaten er under omstilling. Utformingen av ytelser og tjenester skal tjene arbeidslinja. Denne rapporten bringer erfaringene fra et forsøk på å få flere sykmeldte raskere friskmeldt og i arbeid.

Hovedspørsmålet som behyses i rapporten er : kan trygdefunksjonærene, ved å snakke med de sykmeldte, bidra til å redusere antall sykefraværsdager og derved redusere sykepenge-utbetalingene?

I 1992 observerte Fylkestrygdekontoret i Buskerud at fylket lå vesentlig over landsgjennomsnittet når det gjaldt utbetaling av sykepenger per innbygger, antall erstattede sykepengedager, og antall langtidssykmeldte. På denne bakgrunn reiste Fylkestrygdekontoret spørsmål om det var noe trygdekontorene kunne gjøre annerledes overfor sykmeldte i oppfølgingsarbeidet. Intensjonen med en sykmelding er å frita en person fra arbeidslivets plikter og krav slik at hun kan hente seg inn og bli frisk. Denne idéen er i økende grad blitt utfordret. I de seinere åra har debatten i stor grad dreid seg om de uheldige og utilsiktede konsekvensene av en sykmelding, særlig de langvarige tilfellene. Fylkestrygdekontoret i Buskerud mente at det var, og er en rekke problemer knyttet til langvarige sykefravær, blant annet følgende:

- Mange sykmeldte går for lenge passivt sykmeldt med påfølgende undergraving av motivasjonen for å gå tilbake til arbeid. De sykmeldte kan bli isolerte, og de kan utvikle en følelse av å være utstøtt. Dette vil igjen føre til at det er vanskeligere å vende tilbake til arbeidet.
- Noen fyller ikke sykdomsvilkåret i folketrygden
- Enkelte bør føres over på attføring på et tidligere tidspunkt
- Mange får ikke den behandlingen de trenger av de rette instansene
- I oppfølgingsrutinen ligger at oppfølging av sykmeldte først skal ta til etter 8 uker i forbindelse med mottakelse av Sykmelding II blanketten. I mange tilfelle er det for seint.
- Sett fra den sykmeldtes side er forholdet til hjelpeapparatet ofte preget av fjernhet og manglende kommunikasjon og informasjon. Dette kan gjelde i forhold til behandlende leger, annen medisinsk ekspertise, arbeidsgivere, og sosial- og trygdekontor.

En effekt av arbeidslinja var at trygdeetaten fikk endringer i regelverket som ga bedre muligheter for trygdekontoret til å gripe mer aktivt inn i sykmeldingsprosessen, og ga trygdefunksjonærene mer vidtgående fullmakter, særlig i forhold til legene. Disse endringene inviterte til en mer offensiv holdning fra trygdekontorets side.

Fylkestrygdekontoret hadde en forestilling om at flere sykmeldte gikk sykmeldt fordi de ikke ble tatt godt nok vare på. Sviktende oppfølging fra trygdekontorets side kunne føre til unødvendig lange sykefravær. Etter fylkestrygdekontorets oppfatning var det alvorlige svakheter ved de gjeldende oppfølgingsrutinene, hvis de overhodet ble fulgt. Den første var at kontakten mellom den sykmeldte og trygdekontoret ble etablert alt for seint. Den andre var at kommunikasjonen foregikk på en upersonlig måte, som regel per brev.

Et forsøksprosjekt burde derfor ha et design som eliminerte disse to svakhetene. Fylkestrygdekontoret var tidlig innstilt på at forsøket burde være kontrollert, og at tiltaket skulle være personlige samtaler med en trygdefunksjonær. Både teori og praktiske erfaringer tilsa at personlig kontakt mellom sykmeldt og trygdekontor kunne føre til at flere ble raskere friskmeldt. Det ble besluttet å starte oppfølgingen etter fire uker.

En slik oppfølgingsmodell er ikke like anvendelig overfor alle slags lidelser. Den intervensjonsmodellen Fylkestrygdekontoret hadde i tankene ville antakelig egne seg best overfor sykmeldte med lidelser som ikke har en klart og entydig klinisk medisinsk definisjon, der varigheten kan variere mye fra tilfelle til tilfelle, og der uførhetsgraden kan være vanskelig å fastsette nøyaktig. Kanskje vil hjelpeapparatets oppfølgings- og behandlingssinnsats overfor personer med denne typen diffuse lidelser være noe tilfeldig, lite målrettet og ikke alltid like effektiv. I slike tilfeller kan også en langvarig sykmelding føre til passivitet og isolasjon, noe som i sin tur kan være destruktivt både for evnen og motivasjonen til å vende tilbake til jobben.

Evalueringsforskning

Oppfølging av langtidssykmeldte har vært et satsingsområde for trygdekontorene i en årrekke. En serie ulike tiltak er blitt satt i verk, og mange er blitt evaluert. Særlig har gruppe for trygdemedisin ved Universitetet i Oslo drevet et aktivt og bredt evalueringsarbeid (Tellnes og medarbeidere 1992). Et hovedfunn fra disse evalueringene var at oppfølgingstiltakene virker, men at de ikke alltid er like effektive.

De metodiske tilnærmingene som har vært brukt har imidlertid noen svakheter. Ett problem er at eksperimentelle metoder ikke har vært brukt. Det er dermed usikkert om det er tiltakene eller andre forhold som endrer seg samtidig, som har generert de observerte forskjellene. Selv om eksperimentelle designs selvsagt også er belemret med svakheter ved utprøving av sosialpolitiske tiltak, er disse alt i alt likevel mindre enn for de fleste andre metoder som er anvendt for å belyse om sosialpolitiske tiltak virker etter hensikten.

I Norge finnes noen få sosialpolitiske forsøksprosjekter med eksperimentelle opplegg. Det er enkelte arbeidsmarkedstiltak og sosialhjelpstiltak som har vært evaluert med eksperimentelle metoder (Pedersen 1993, Røed og Try 1993, Raaum og Torp 1993, Raaum, Torp og Goldstein 1993, Try 1992, 1993a,b). I disse studiene avdekkes som regel en positiv virkning av tiltakene i forventet retning, men den kan være svak og usikker, og omfatter ikke alltid alle undergrupper.

De norske erfaringene er i overensstemmelse med resultatene fra evalueringene av den såkalte «workfare»-politikken i USA. Fra begynnelsen av 1980-tallet har workfaretiltak avløst de tidligere welfaretiltakene i en rekke amerikanske stater. Tiltakene har til dels vært ulikt organisert og hatt ulike målsettinger. De fleste har imidlertid delt en målsetting om at arbeidsløse sosialhjelpsmottakere skal ut av fattigdommen og inn i jobb. Målgruppa har i hovedsak vært enslige mødre.

En lang rekke av disse workfaretiltakene har vært evaluert med eksperimentelle metoder. En omfattende gjennomgåelse av en serie evalueringer, alle med eksperimentelle design, konkluderer at det nytter med jobbfremmende tiltak over sosialhjelpsmottakere, men at «forventningene bør være beskjedne» (Gueron og Pauly 1991). De noe nedslående resultatene har ført til at myndighetene i USA i øyeblikket er skeptiske til hele workfare-konseptet, og flere rikspolitikere mener at resultatene av workfarepolitikken har vært magre i forhold til hva de har kostet det offentlige (Wiseman 1995).

Overføringsverdien av sosialpolitiske tiltak av den typen vi nå har nevnt til vårt område, må ikke overdrives. Å være sykmeldt er noe annet enn å være arbeidsledig eller sosialhjelpsmottaker. Vi kjenner kun til to undersøkelser av sykefravær som har et tilnærmet eksperimentelt design med en tiltaksgruppe og en kontrollgruppe. I den ene ble langtidssykmeldte innkalt til trygdekontorets rådgivende lege for konsultasjon og

samtale (Melsom i NOU 1986:22). I konsultasjonen ble det lagt vekt på informasjon og motivasjon i forhold til den sykmeldte. Begge grupper ble fulgt over ett år. Etter fire måneder var differansen mellom utbetalte sykepenges per sykmeldt i hver gruppe 2334 kroner i favør av tiltaksgruppa. Etter ett år var differansen økt til 4557 kroner per person. Det går ikke fram av rapporten fra prosjektet om inndelingen i tiltaksgruppe og kontrollgruppe var tilfeldig. I den andre undersøkelsen ble såkalte gjengangeryskmeldte innkalt til samtale (Melsom og Noreik 1988). Tiltaksgruppa og kontrollgruppa var tilfeldig valgt ut. Av de innkalte 78 personene i kontrollgruppa møtte halvparten, det vil si 39 personer. Antakelig er ikke disse representative for alle innkalte. Over ni måneder påviste forfatterne en tendens til at de innkalte som møtte hadde færre sykefraværsdager og lavere sykepengeutbetalinger i gjennomsnitt enn kontrollgruppa. Differansen i sykefraværsdager og sykepengeutbetaling mellom disse to gruppene var henholdsvis 3540 kroner og 1,2 dager.

Disse resultatene antyder at tiltak i form av samtale med sykmeldte gir en viss begrenset effekt på sykefraværenes varighet og på sykepengeutbetalingene.

Prosjektdesign

Under forberedelsene til prosjektet ble det avholdt flere møter der representanter fra flere lokale trygdekontorer i Buskerud deltok. Til slutt vedtok fire kontorer å delta i prosjektet. Disse lå i kommuner av forskjellig størrelse og næringstype. Etter at forsøkskommunene var klare, ble det nedsatt en arbeidsgruppe der prosjektmedarbeiderne fra disse fire kontorene deltok i utformingen av det endelige prosjektopplegget.

Det ble tidlig bestemt at tiltaket skulle være personlige samtaler som prosjektmedarbeideren skulle ta med de sykmeldte. Av økonomiske grunner skulle de sykmeldte få tre samtaler i gjennomsnitt, etter fire, sju og elleve uker, dersom de ikke ble friskmeldt i mellomtida. Hvis det var penger til overs kunne antallet samtaler utvides.

Prosjektet ble lagt så nært opp til et klassisk eksperiment som praktisk mulig. Dette innebærer en tilfeldig tilordning av aktuelle personer til en tiltaksgruppe og en kontrollgruppe, og at det «stimulus» personene i tiltaksgruppa utsettes for er så lik som mulig (Pocock 1983). Idealet for det kontrollerte eksperimentet er utprøving av medikamenter. I slike forsøk er det ikke særlig vanskelig å holde stimulus konstant, det vil si servere samme pille. I praksis er dette umulig i den typen sosiale eksperimenter vi har å gjøre med her.

Den konkrete og eksakte utformingen av prosjektet ble grundig diskutert i prosjektgruppa for å oppnå en mest mulig samstemt forståelse av designet, innkallingsprosedyrer og samtalenes innhold. Det ble derfor utarbeidet et samtalskjema som skulle følges i samtalene (se vedlegg). I samtaleprogrammet ble det óg innarbeidet en progresjon fra første til tredje samtale for å forhindre at samtalene skulle bli en trettende og irriterende repetisjonsøvelse for de sykmeldte.

For å komme i betraktning måtte de sykmeldte tilfredsstillende fire kriterier. 1) Av grunner vi har diskutert, falt valget av diagnoser på muskel-/skjelettsykdommer og nervøse lidelser 2) øvre aldersgrense ble satt til 65 år, 3) kun ansatte ble inkludert og 4) bare 100 prosent sykmeldte var aktuelle.

Hvert trygdekontor, fire i tallet, skulle inkludere 70 fireukers sykmeldte i tiltaksgruppa. Det totale antall tiltaksdeltakere ble altså 280 personer. Den like fordelingen av sykmeldte på hvert trygdekontor måtte fravikes noe fordi det minste kontoret ikke klarte å fylle opp sin kvote. Derfor innlemmet et større kontor flere enn 70 personer.

280 personer, 70 fra hvert av de deltagende kontorene, ble også innlemmet i kontrollgruppa. Overfor personer i denne gruppa ble det ikke iverksatt spesielle tiltak utover de ordinære oppfølgings tiltak trygdekontorene har rutiner for. Som regel er, eller var, oppfølgingsinnsatsen fra de fleste trygdekontorenes side ikke særlig offensiv og omfattende. Likevel betyr dette at en del mennesker i kontrollgruppa har fått hjelp og behandling hos andre instanser. Dette verken ville eller kunne vi forhindre.

I etterhånd har det dessverre vist seg at ett trygdekontor har fulgt en annen prosedyre for innlemming av sykmeldte i tiltaksgruppa enn de tre øvrige kontorene. Dette kontoret har ekskludert sykmeldte som etter å ha fått innkallingsbrevet ringte og sa at de regnet med å bli friskmeldt i løpet av kort tid. 19 personer, det vil si 24 prosent av alle i tiltaksgruppa, ble ekskludert på dette grunnlaget. Denne kategorien sykmeldte er innlemmet i tiltaksgruppa ved de andre kontorene. Vi mente det var viktig å fange opp denne kategorien fordi innkallingsbrevet i seg selv har karakter av tiltak og kunne antas å ha en «helbredende» virkning. Det var dessuten absolutt nødvendig av hensyn til sammenlignbarheten med kontrollgruppa å ha med alle i tiltaksgruppa med et fravær på over fire uker, selv om enkelte forventet friskmelding umiddelbart etter dette. Grunnen var at vi selvsagt ikke hadde noen mulighet til å sortere kontrollpersoner etter slike subjektive kriterier. Dette betyr at 19 personer ved dette kontoret som med høy sannsynlighet ville hatt svært korte fravær er erstattet med personer med lavere sannsynlighet for korte fravær. Tiltaksgruppa ved dette kontoret avviker dermed systematisk fra tiltaksgruppa ved de tre andre kontorene¹ og dessuten fra kontrollgruppa i eget kontor. Vi har derfor funnet det riktig å holde dette ene kontoret utenfor analysen. *Analysen bygger dermed på sykmeldte ved tre trygdekontorer i Buskerud fylke og omfatter 202 personer i tiltaksgruppa og 205 personer i kontrollgruppa.*

Samtidig som de aktuelle kandidatene til tiltaksgruppa ble inkludert, foregikk arkivering av kandidater til kontrollgruppa i egne mapper ved de enkelte kontorene. I og med at den siste som ble sykmeldt skulle få mulighet til å gå den tilmålte sykmeldingsperioden på ett år helt ut, strakk prosjektet seg over bortimot to år. Mot slutten av prosjektperioden ble 280 personer fra disse mappene tilfeldig valgt ut, og deres vitale data registrert.

Sykmeldte – både kandidater til tiltaksgruppa og til kontrollgruppa – ble identifisert etter tre ukers sykmelding, og arkivert i egne mapper. Sykmeldte som skulle inngå i tiltaksgruppa ble tilskrevet og bedt om å møte på trygdekontoret innen én uke. I de tilfelle friskmelding ble utstedt mellom tredje og fjerde uke, ble personen strøket. Dette gjaldt selvsagt personer både i tiltaks- og kontrollgruppa. I og med at prosjektet skulle løpe parallelt med den daglige drift ved trygdekontorene, var det nødvendig å fordele arbeidsbelastningen over tid. Dette tilsa et design der mindre puljer av sykmeldte ble innkalt til samtale med jevne mellomrom, og hvor det var avsatt tid også til annen- og tredjegangssamtaler. Det ble derfor utarbeidet en detaljert innkallingsplan (se vedlegg).

Det ble innprentet at utvelgingen av tiltakspersoner såvel som kontrollpersoner skulle være tilfeldig. Tilfeldighetsmekanismen ble ivaretatt ved å spre sykmeldingsblankettene utover et bord med baksida opp. Deretter ble det nødvendige antallet til tiltaksgruppa trukket tilfeldig. Den samme prosedyren ble fulgt for kontrollgruppa. Til begge grupper ble alle som tilfredsstilte de angitte kriteriene valgt ut.

¹ Gjennomsnittlig antall sykefraværsdager er da også høyest ved dette kontoret; 125 dager mot henholdsvis 122, 116 og 110 dager.

Data og metode

Våre viktigste avhengige variable er antall dager med sykefravær og brutto utbetalte sykepenger. Vi skal se på forskjeller mellom tiltaksgruppa og kontrollgruppa, og hvordan forskjellene er for menn og kvinner, ulike aldersgrupper og for ulike diagnoser. I tillegg undersøker vi gradert sykmelding og opphørsstatus. Opphørsstatus har tre verdier: friskmeldt, maksimal sykmeldingstid og «oppført av andre grunner».

Datakildene er edb-arkivene til de trygdekontorene som har deltatt i forsøket. Fra disse arkivene har medarbeiderne på prosjektet overført opplysninger om hver enkelt person i tiltaksgruppa fra sykmeldingsblankettene til egne registrerings skjemaer (se vedlegg). I anonymisert form ble disse sendt til Fafo for overføring til edb. Opplysningene for kontrollgruppa er blitt registrert på tilsvarende skjemaer av Bjarte Johannesen fra Fylkes-trygdekontoret i Buskerud og oversendt Fafo.

For å undersøke om sykepengeutbetalingene og antall fraværsdager er signifikant forskjellig i de to gruppene, har vi benyttet en vanlig statistisk test av forskjeller mellom gjennomsnittstall, og multippel lineær regresjonsanalyse. I analysen av antall sykefraværsdager har vi i tillegg anvendt såkalt forløpsanalyse, det vil si overlevelsesanalyse og Cox-regresjon. Det fortrinn disse statistiske tilnærmingene har sammenlignet med andre teknikker er at de er istand til å håndtere personer som gikk hele sykmeldingsperioden ut, eller som forlot sykepengeordningen av «andre årsaker». I forløpsanalysen vil maksimal sykmeldingstid og «oppført av andre grunner» behandles som «sensurerte» i og med at vi ikke vet hvordan det har gått med disse. Vi gjør oppmerksom på at friskmelding ikke alltid er ensbetydende med «tilbake i arbeid», men som regel er det slik.

Resultater

Sammensetningen av tiltaksgruppa og kontrollgruppa

Poenget med å trekke kandidater til tiltaksgruppa og kontrollgruppa tilfeldig, er å forhindre at det oppstår systematiske forskjeller mellom gruppene. De forskjeller vi vil observere mellom gruppene kan dermed med større sannsynlighet tilskrives «tiltaket», i og med at det er det eneste som i prinsippet skiller de to gruppene fra hverandre. I kvasiekperimentelle design, som ofte innebærer selvseleksjon, er ikke denne sikkerheten innebygget. At utvelgingen er tilfeldig, er derimot ingen garanti for at gruppene blir

Tabell 1 Tiltaksgruppe og kontrollgruppe etter kjønn, alder og diagnose. Prosent, gjennomsnitt og standardavvik (SD).

	Tiltak	Kontroll	Alle
Kjønn			
Menn	49	42	46
Kvinner	51	58	54
Fødselsår			
Gj.snitt	1953	1951	1952
SD	12	11	11
Diagnose			
Muskel/skjel.	87	87	87
Psykisk	13	13	13
N	202	205	407

helt like. Også tilfeldighetene kan generere ulik gruppesammensetning. Tabell 1 viser hvordan de to gruppene er sammensatt med hensyn til kjønn, alder og diagnose.

Sammenlignet med kontrollgruppa er det i tiltaksgruppa en større andel menn, og gjennomsnittsalderen er to år lavere. Diagnosefordelingen er helt lik i de to gruppene. Forskjellene i kjønns- og alderssammensetning er ikke signifikante, det vil si statistisk sikre, etter konvensjonelle statistiske kriterier.

Randomiseringsprosedyren har således gitt grupper som ikke er helt like. De forskjellene vi kan observere mellom gruppene taler for å kontrollere for kjønn og spesielt alder i analysen, siden det er veldokumentert at eldre har lengre sykefravær enn yngre (Hippe og Hagen 1991).

Om samtalene

Det var satt av penger til i gjennomsnitt tre samtaler pr. sykmeldt. Tabell 2 gir nøkkeltall for de samtalene som tiltaksgruppa ble innkalt til.

Tabell 2 Fordeling av antall samtaler (prosent), deres varighet i minutter (gjennomsnitt), og andel telefonsamtaler (prosent).

	Samtaler andel	Varighet i minutter	Telefon andel
1 samtale	44	25	20
2 samtaler	28	18	35
3 samtaler+	28	16	34

Nær 45 prosent av tiltakspersonene fikk kun én samtale. For øvrig fikk like mange to og tre samtaler, 28 prosent.

Som en vel kunne vente, avtar gjennomsnittlig samtaletid med antall samtaler, fra 25 minutter ved første samtale til 16 minutter ved siste.

Én av fem av førstegangssamtalene ble gjort over telefon. Andelen telefonsamtaler ved annen- og tredjegangssamtale ligger på samme nivå, om lag 35 prosent. I intervjuerinstruksjonen var det gjort klart at telefonsamtale kun skulle gjennomføres i «nødstilfelle». At såpass mange av samtalene er gjort over telefon illustrerer antakelig at det ofte er vanskelig å få langvarig sykmeldte til å møte opp personlig på trygdekontoret.

Utfallet av forsøket

Tabell 3 viser resultatene av forsøket målt med fire indikatorer.

Det viktigste trekket ved tabell 3 er at i gjennomsnitt har tiltaksgruppa 19 færre sykefraværsdager, og nær 4360 kroner lavere sykepengeutbetaling enn kontrollgruppa. Imidlertid er det bare forskjellen i sykefraværsdager som nærmer seg statistisk signifikans. At disse resultatene ikke er sikrere i statistisk forstand, skyldes at det er betydelig variasjon innad i begge gruppene, noe de høye standardavvikene tydelig demonstrerer.

Gruppeforskjellene i opphørsstatus og i hvorvidt de sykmeldte hadde gradert sykmelding, er små og heller ikke statistisk signifikante.

Når standardavvikene er såpass store, melder spørsmålet seg om tiltaket har vært spesielt effektivt overfor spesielle undergrupper av sykmeldte. Vi har begrensede muligheter til å studere undergrupper, men materialet tillater oss å se på forskjeller etter kjønn, alder og diagnose. Tabell 4, 5 og 6 viser utfallsvariablene varighet og utbetalte sykepenger for henholdsvis menn og kvinner, unge og eldre, og for muskel/skjelettsykdommer og psykiske lidelser.

Tabell 3 Gradert sykmelding, opphørsstatus, gjennomsnittlig varighet i sykefraværsdager, gjennomsnittlig utbetalt sykepenger per sykmeldt etter tiltaksgruppe, kontrollgruppe og alle.

	Tiltak	Kontroll	Differanse (K-T)	Alle
Varighet i dager*	110	129	19	119
standardavvik	(98)	(105)	(102)	
Kroner utbetalt**	38 939	43 295	4356	41 138
standardavvik	(42 984)	(41 620)		(42 304)
Opphørsstatus:				
Friskmeldte %	88	89	1	88
Maksimal tid %	9	10	1	10
Andre opphørsårsaker %	3	1	-2	2
Gradert sykmelding %	21	27	6	24

* Signifikanssannsynlighet .06

** Signifikanssannsynlighet .30

Tabell 4 Gjennomsnittlig antall sykefraværsdager og gjennomsnittlig utbetalt sykepenger per sykmeldt etter kjønn, tiltaksgruppe (T) og kontrollgruppe (K).

	Menn			Kvinner		
	T	K	diff (K-T)	T	K	diff (K-T)
Varighet i dager*	102	128	26	117	129	12
standardavvik	(93)	(97)		(102)	(110)	
Kroner utbetalt**	41 813	50 256	8 443	36 204	38 163	1959
standardavvik	(44 087)	(42 956)		(41 940)	(40 016)	
N=	98	87		103	118	

* Signifikanssannsynlighet: Menn .06, kvinner: .39

** Signifikanssannsynlighet: Menn .19, kvinner: .72

I tabell 5 er de «unge» født i 1950 eller senere, mens de «eldre» er født i 1949 eller tidligere.

Tabell 5 Gjennomsnittlig antall sykefraværsdager og gjennomsnittlig utbetalt sykepenger per sykmeldt etter alder, tiltaksgruppe (T) og kontrollgruppe (K).

	Unge			Eldre		
	T	K	diff (K-T)	T	K	diff (K-T)
Varighet i dager*	106	116	10	115	147	32
standardavvik	(95)	(93)		(102)	(117)	
Kroner utbetalt**	36 366	38 602	2 236	42521	49 920	7 399
standardavvik	(40 431)	(37 467)		(46 315)	(46 284)	
N=	117	120		84	85	

* Signifikanssannsynlighet: Unge .43, eldre .06

** Signifikanssannsynlighet .Unge .66, eldre .30

Tabell 6 Gjennomsnittlig antall sykefraværsdager og gjennomsnittlig utbetalte sykepenger per sykmeldt etter diagnosegruppe, tiltaksgruppe (T) og kontrollgruppe (K).

	Muskel-/skjelettsykdommer			Psykisk		
	T	K	diff (K-T)	T	K	diff (K-T)
Varighet i dager*	104	129	25	131	133	2
Standardavvik	(89)	(104)	(125)	(108)	(117)	
Kroner utbetalt**	37 611	43 192	5 581	42 489	45 164	2 675
Standardavvik	(41 622)	(41 365)	(48 283)	(44 492)	(46 284)	
N=	273	178		26	26	

* Signifikanssannsynlighet Muskel/skjel.: .02, Psykisk: .95

** Signifikanssannsynlighet Muskel/skjel.: .21, Psykisk: .84

Tabell 4, 5 og 6 viser det samme utfallsmønsteret i alle grupper. Uansett kjønn, alder og diagnosegruppe er antallet sykefraværsdager lavere i tiltaksgruppa enn i kontrollgruppa, og utbetalte sykepenger likeså. Et gjennomgående trekk er videre at tiltaket synes å virke relativt sterkere på antall dager med sykefravær enn på utbetalte sykepenger.

Likevel registrerer vi til dels markerte forskjeller fra undergruppe til undergruppe. Forskjellene er større for menn enn for kvinner, for gamle enn for unge, og for dem som har muskel/skjelett sykdommer enn for dem med psykiske lidelser. Blant menn, eldre, og muskel/skjelettplagete er forskjellene i fraværsdager statistisk sikre, eller svært nær.

Det er litt forbausende at de to utfallsvariablene fraværsdager og utbetalingsbeløp ikke stemmer bedre overens. 19 dagers forskjell i sykefraværet ledsages av en forskjell på bare 4356 kroner i utbetalte sykepenger. Gjennomsnittlig utbetalte sykepenger per dag per person i undersøkelsen er 344 kroner. 19 dagers redusert sykefravær svarer dermed til 6536 kroner. Forskjellen i teoretisk innspart beløp og faktisk innspart beløp er således 2180 kroner. Hva skyldes denne diskrepansen?

En mulig forklaring på denne uoverensstemmelsen er at sykepengeutbetalingene ikke bare avhenger av sykefraværets lengde, men også av de sykmeldtes «sykepengegrunnlag». At sykepengegrunnlaget er forskjellig i de to gruppene, vises i tabell 7. Her er «sykepengegrunnlaget» per person beregnet med utgangspunkt i totalt utbetalt kronebeløp og totalt antall dager med sykefravær innen hver gruppe.

Tabell 7 Sykepengegrunnlag i kroner per dag per person i henholdsvis tiltaksgruppa og kontrollgruppa etter fraværens varighet.

	Totalt	Under 150 dager	151 dager og mer
Tiltak	354	330	376
Kontroll	336	328	341
Differanse(T-K)	18	2	35
N=	407	302	105

Tabell 7 viser at for alle fraværsdager er sykepengegrunnlaget per dag per person i tiltaksgruppa 18 kroner høyere enn i kontrollgruppa. For fravær under 150 dager er forskjellen forsvinnende liten, bare to kroner. For fravær som varer 150 dager eller mer er forskjellen mellom de to gruppene betydelig, hele 35 kroner per person per dag. En grunn til at innsparingene i sykepenger er mindre enn reduksjonen i antall fraværsdager, er derfor at personer i tiltaksgruppa med lange fravær får mer utbetalt pr. dag enn sykmeldte i kontrollgruppa. Det foregår altså en sterkere seleksjon etter inntekt blant personer i tiltaksgruppa enn i kontrollgruppa. Det kan med andre ord se ut til at det er

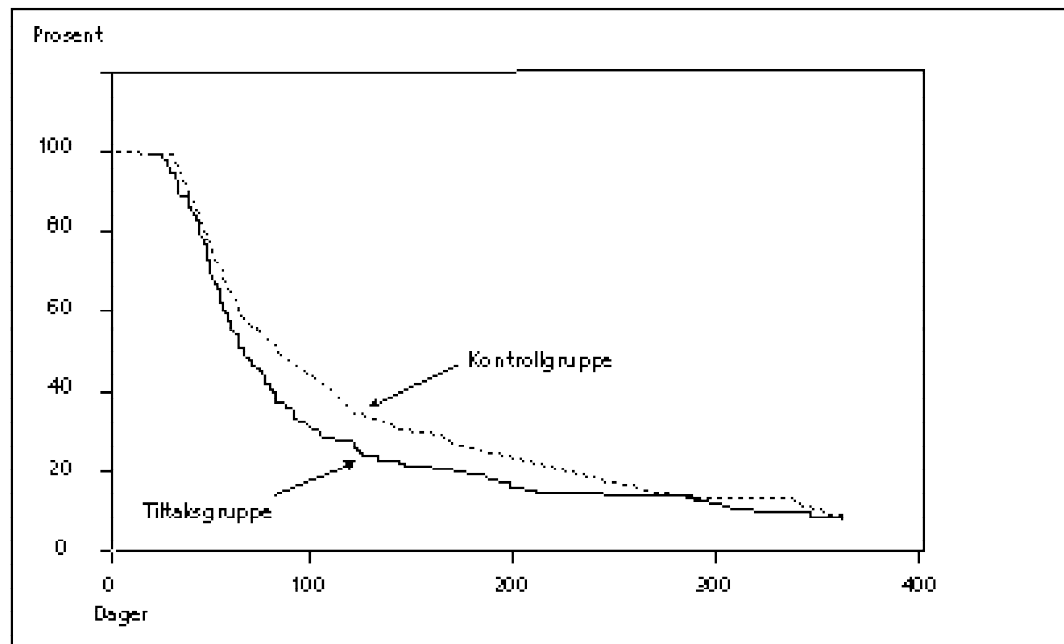
lettere å redusere lengden av sykefraværene til sykmeldte med lavere inntekter. Hva dette i sin tur skyldes, kan ikke slås fast uten supplerende data, men vi kan foreslå to tolkninger som nødvendigvis er litt spekulative.² En medisinsk tolkning er at grupper med lavere inntekter kan ha mindre alvorlige sykdommer og høyere grad arbeidsforhet noe som gjør at intervensjonen, det vil si samtalene, virker mer effektivt. En sosiologisk tolkning er at sykmeldte med lavere inntekter er mer sensitive og lydhøre for signalene fra trygdefunksjonærene fordi de oppfattes som autoriteter.

Vi nevnte foran at tiltaksgruppa og kontrollgruppa ikke er likt sammensatt med hensyn på kjønn og alder. Dette bør det kontrolleres for statistisk. Vi har derfor utført diverse multivariate regresjonsanalyser for å undersøke om ulik gruppesammensetning påvirker resultatene (tall ikke vist). Resultatet av disse analysene understøtter hovedkonklusjonen så langt.

Forløpsanalyse

Det er flere problemer med å anvende ordinær regresjonsanalyse på såkalte forløpsdata som varighet av sykefravær er. Det alvorligste problemet er kanskje at for en del sykmeldte har vi ufullstendige opplysninger. Vi vet for eksempel intet om hvordan det har gått med dem som har gått sykmeldingsperioden ut. Vi kjenner ikke til om de senere er blitt friskmeldt eller fått uførepensjon. Disse individene er «sensurerte» og må behandles deretter. Ordinær regresjonsanalyse kan ikke håndtere sensurerte data, men det kan forløpsanalytiske teknikker. Uten å gå i detaljer om de statistiske finurlighetene ved denne klassen av metoder, skal vi vise resultatene fra to litt ulike typer forløpsanalyser. Den første kalles overlevelsesanalyse, den andre cox-regresjon. Figur 1 viser overlevelseskurvene for henholdsvis tiltaksgruppa og kontrollgruppa slik de blir beregnet i en overlevelsesanalyse.

Figur 1 Overlevelseskurver for tiltaksgruppa og kontrollgruppa.



Figur 1 viser en ganske klar tendens til at personer i tiltaksgruppa temmelig raskt oppviser høyere friskmeldingsrater enn personer i kontrollgruppa. Denne tendensen ser

² Foran påviste vi at andelen med gradert sykmelding var noe høyere i kontrollgruppa enn i tiltaksgruppa. En multipel regresjonsanalyse avkrefter imidlertid at dette forklarer den moderate forskjellen i utbetalte sykepenge mellom de to gruppene.

imidlertid ut til å avta etter rundt sju måneder. Deretter nærmer kurvene seg hverandre for så å falle nesten sammen fra om lag ni måneder. Som også tabell 3 viste, er det ikke nevneverdig forskjell på de to gruppene når det gjelder tilbøyelighet til å gå sykepengeperioden ut; forskjellen kommer først og fremst til uttrykk i reduserte fraværsdager blant dem som friskmeldes.

Forskjellen i varighet mellom de to gruppene gjenspeiles også i medianverdiene, det vi si den verdien som deler fordelingen i to like deler. For tiltaksgruppa er medianverdien 68 dager, mens den er 85 dager for kontrollgruppa. Forskjellen mellom medianverdiene er altså 17 dager og avviker således lite fra gjennomsnittsforskjellen på 19 dager.

Signifikanstesten som denne analysen gir viser at forskjellen i overlevelse mellom de to gruppene såvidt er statistisk sikker etter konvensjonelle kriterier (sig=0.03).

Vi har vist at tiltaksgruppa og kontrollgruppa ikke er likt sosialt sammensatt. Det har vi ikke tatt hensyn til i denne overlevelsesanalysen. Analysen nedenfor i tabell 8, viser effekten av å tilhøre henholdsvis tiltaksgruppa og kontrollgruppa på varigheten av sykefraværet, når vi tar hensyn til at de to gruppene er ulikt sammensatt hva gjelder kjønn, alder (fødselsår) og diagnose.

Tabell 8 Resultater fra forløpsanalyse, Cox-regresjon. Kontrollerte sammenhenger.

-2 Log Likelihood	3693,816		
	Kji-kvadrat	Frihetsgrader	Signifikanssannsynlighet
Modellforbedring	10,370	4	,0346
Variabel	Ekspensiert koeffisient	Frihetsgrader	Signifikanssannsynlighet
Gruppe	1,2295	1	,0528
Fødselsår	1,0103	1	,0233
Kjønn	1,0581	1	,6023
Diagnose	1,0941	1	,2938

I tabell 8 er effekten av hver enkelt variabel kontrollert for effektene av de øvrige variablene i modellen. Det er to trekk som er viktige å legge merke til i tabell 8.

Det viktigste er at den kontrollerte effekten av å tilhøre tiltaksgruppa eller kontrollgruppa på varigheten av sykefraværet er nesten statistisk sikker (sig=.053)³. Resultatet som er vist i figur 1 bekreftes altså i hovedsak, og endres lite om vi tar hensyn til at de to gruppene er noe ulikt sosialt sammensatt.

Det andre som er verdt å merke seg er at for øvrig er det kun alder som øver en signifikant innflytelse på varigheten av sykefraværet. Varigheten øker med økende alder. Verken kjønn eller diagnose innvirker på sykefraværets varighet, «alt annet likt».

³ Cox-regresjon bygger på den forutsetning at forholdet mellom de såkalte hazardratene for to tilfeldige individer er proporsjonalt, og ikke varierer som en funksjon av tiden. Som overlevelseskurvene for de to gruppene viser (figur 1), er denne forutsetningen brutt i vårt materiale. Strengt tatt passer derfor ikke Cox-regresjon så godt i dette tilfellet. Dette kan være en grunn til at koeffisienten for gruppetilhørighet ikke er statistisk signifikant.

Oppsummering og diskusjon

Analysen viser at forsøket med å «snakke sykmeldte tilbake til jobben» har en effekt i forventet retning. Sykefraværet er 17–19 dager kortere i tiltaksgruppa enn i kontrollgruppa og sykepengeutbetalingene er lavere, om lag 4350 kroner. Det samme mønsteret går igjen for menn og kvinner, unge og eldre, og for sykmeldte med muskel/skjelettlidelser og psykiske lidelser. Forskjellene mellom tiltaksgruppa og kontrollgruppa er imidlertid størst for menn, for eldre, og for personer med muskel/skjelettlidelser.

Forskjellene i sykefraværsdager mellom tiltaksgruppa og kontrollgruppa er i noe varierende grad sikre etter konvensjonelle statistiske kriterier, det vil si med en signifikanssannsynlighet på fem prosent. Ikke i noe tilfelle er forskjellene i utbetalte sykepenger statistisk sikker.

Denne uoverensstemmelsen mellom de to utfallsvariablene skyldes blant annet at personer i tiltaksgruppa med lange fravær, det vil si fravær på 151 dager og mer, har betydelig mer utbetalt per dag enn langtidsfraværende i kontrollgruppa. Det er ikke klart hva som ligger bak denne seleksjonen, men forklaringen kan være at grupper med lavere inntekter har mindre alvorlige sykdommer og lettere arbeidsuførhet, eller at grupper med lavere inntekter er mer lydhøre for signalene fra trygdefunksjonærene fordi disse oppfattes som myndighetspersoner.

Samtaletiltaket synes å virke på varigheten av sykefraværet, men ikke på tilbøyeligheten til å gå sykepengeperioden på ett år helt ut. Andelen som går ut hele ettårsperioden er omtrent ti prosent i både tiltaksgruppa og kontrollgruppa. Dersom ett års sykefravær er ett tegn på gryende utstøting fra arbeidslivet, betyr det at samtaletiltaket ikke er særlig egnet til å bremse utstøtingsprosessen. Effekten av tiltaket synes å være begrenset til å korte ned fraværslengden for dem som blir friskmeldt i løpet av ett år. Ved nærmere ettertanke er ikke dette så overraskende. De krefter som virker på utstøtingen fra arbeidslivet er antakelig så sterke at de ikke motvirkes nevneverdig av denne typen tiltak alene.

17 eller 19 dagers reduksjon av sykefraværet, avhengig av statistisk mål (henholdsvis median og gjennomsnitt), blant fire ukers sykmeldte er trygdepolitisk betydningsfullt, selv om resultatet er i grenselandet for statistisk sikkerhet. En grunn til at en såpass klar forskjell ikke er sikrere i statistisk forstand, er at variasjonen innad i begge gruppene er meget betydelig, og at det er forholdsvis få individer i hver gruppe.

Denne heterogeniteten kan ha mange ulike kilder. Et særlig viktig forhold er antakelig at sykdommene innen en og samme diagnosegruppe har svært forskjellig alvorlighetsgrad og uførhetsskapende konsekvenser. Tilsvarende forsøk med mer homogene grupper av sykmeldte ville muligens gitt resultater som var statistisk sikre.

Et spørsmål er hvor stor vekt vi skal legge på de statistiske konvensjonene. Vi kan godta større statistisk usikkerhet enn det som er vanlig, og i større grad betone den trygdepolitiske relevansen av funnene som i dette tilfelle er betydelig. En påminnelse av hva signifikansnivået betyr kan her være på sin plass. Et resultat med en signifikanssannsynlighet lik 0.05 betyr at i fem av 100 tilfeldige undersøkelser vil dette resultatet oppstå på grunn av sjansen alene. Det vanlige signifikansnivået på 0.05 er satt helt vilkårlig, og har ingen matematisk eller substansiell begrunnelse. Forskjellen på en signifikanstest som gir 0.06 og 0.04 er helt minimal, og det er selvsagt ikke slik at den siste verdien beviser at tiltaket har virket mens den første beviser det motsatte (se for eksempel Pocock 1983:205). Det er hele tiden snakk om grader av sannsynlighet langs

et kontinuum.⁴ Også sett på bakgrunn av tidligere tilsvarende forsøk, som ble referert innledningsvis, er det rimelig å anta at forsøk av denne typen gir en viss reell effekt i ønsket retning, selv om vi trenger flere og helst større undersøkelser for å *fastslå dette med større grad av sikkerhet*.

Det er vanskelig å vurdere hvilke konkrete aspekter ved samtalen som har hatt den tilsiktede virkningen. Noen vil antakelig ha opplevd samtalen som en oppmuntring og som en spore til å gjøre noe på egen hånd, slik intensjonen med forsøket var. Andre vil antakelig ha følt seg kontrollert og mistenkeliggjort. Vi kan heller ikke se bort fra at enkelte ble «skremt» til å friskmelde seg. Prosjektmedarbeiderne som sto for samtalen rapporterer imidlertid at de aller fleste sykmeldte reagerte positivt på å bli innkalt. I forlengelsen av dette må vi ikke glemme den såkalte «Hawthorne-effekten». Dette uttrykket betegner at ethvert forsøksiltak har en gunstig effekt, ganske uavhengig av tiltakets essens. Bare det å få oppmerksomhet kan gi gunstige virkninger. Dersom det er en Hawthorne-effekt vi avdekker i dette forsøket, betyr det at det konkrete innholdet i akkurat dette samtaleiltaket er uvesentlig. Det vesentlige er at trygdekontoret eller andre instanser gjør noe.

Et viktig spørsmål som vi ikke har hatt anledning til å undersøke i denne omgang er hvordan det går med de fire ukers sykmeldte på lengre sikt. Vi har begrenset oss til kun å kartlegge korttidsvirkningen, det vil si ettårsvirkningen, av samtaleiltaket. Et viktig spørsmål er selvsagt om de friskmeldte blir raskt sykmeldt igjen, eller om de ikke skiller seg vesentlig fra «normalbefolkningen» i så måte? Trygdekontorene sitter på opplysninger som gjør det mulig å etterspore langtidsvirkningene av denne typen intervensjon. Slike data bør nyttiggjøres.

I et kontrollert forsøk er det avgjørende at personer plukkes tilfeldig til både tiltaks- og kontrollgruppa. På papiret skisserte vi et opplegg som skulle sørge for at så skjedde. Det ble imidlertid ikke etablert kontrollrutiner for å overvåke at prosedyrene ble etterlevd i praksis. Det kan for eksempel tenkes at personer som skulle vært i tiltaksgruppa aldri ble inkludert fordi prosjektmedarbeiderne fant det «urimelig» eller hensiktsløst å kalle dem inn til samtale. Dersom disse kandidatene til tiltaksgruppa har høy sannsynlighet for langvarig fravær, oppstår en systematisk skjevhet som fører til at resultatene for tiltaksgruppa overvurderes. Denne problemstillingen har vært diskutert med prosjektmedarbeiderne i referansegruppa. De avviser at de har silt tiltaks personer etter kriterier som ble nevnt ovenfor. De holder fast ved at de har fulgt de prosedyrer for klientutvelgelse som det var enighet om på forhånd.

Det er to forhold som taler for at de forholdsvis svake effekter som avdekkes i et kontrollert forsøk vil være større i det praktiske liv. For det første foregår ikke utvelgingen av tiltaks personer i trygdekontorenes hverdag etter tilfeldighetsprinsippet. I hverdagen er det ikke de «håpløse» som velges ut, men personer trygdekontoret tror kan hjelpes: utvelgingen av personer til tiltak vil være strategisk ut fra et håp om at tiltaket vil hjelpe. For det andre har sykmeldte i forsøksgruppa ikke vært upåvirket mens forsøket har gått. I sosialpolitiske forsøk av denne typen, vil ikke tiltakene være uten spredningseffekter.

⁴ Vi har brukt tosidige signifikanstester. Vi kunne ha bedret den statistiske sikkerheten ved å bruke ensidige tester siden signifikanssannsynligheten for en ensidig test er halvparten av en tosidig test. I motsetning til en tosidig test, er en ensidig test basert på den forutsetning at tiltaket ikke gir dårligere resultat enn ikke-tiltaket. En ensidig test gir dermed holdepunkter for å vurdere om tiltaket gir bedre, eller minst like bra, resultat som ikke-tiltaket. Denne forutsetningen er forholdsvis streng og kan ofte være tvilsom. I litteraturen om klinisk kontrollerte forsøk regnes det derfor ikke som god tone å bruke ensidige tester (Pocock 1983:206).

Vi må for eksempel spørre om det spredte seg mer restriktiv sykmeldingspraksis blant sykmeldende leger når de fikk høre om trygdekontorets forsøksvirksomhet. Ble noen leger flinkere til å henvise pasienter, eller til å få sine pasienter raskere til spesialistutredning- og behandling? Fikk andre langtidssykmeldte høre om tiltaket og ble påvirket av det? Var det noen som friskmeldte seg fordi de ikke «hadde rent mel i posen»? Var det noen som oppsøkte trygdekontoret på eget initiativ for en samtale, eller henvendte seg til andre behandlingsinstanser? Vi har selvsagt ikke hatt mulighet til å forhindre at andre har satt i gang tiltak overfor sykmeldte i kontrollgruppa som har vært effektive. Slike problemer er en uløselig del av denne typen sosiale eksperimenter. Andre sykmeldte, herunder sykmeldte i tiltaksgruppa, kan ikke hindres i å få hjelp som de har rett til, men det vil selvsagt medvirke til at den målte effekten av forsøket blir mindre enn den ellers ville ha vært. I slike sosialpolitiske eksperimenter som vi har å gjøre med her, er slike problemer ikke til å unngå. Dette skiller sosialpolitiske eksperimenter fra rene kliniske eksperimenter som langt på vei kan garantere at kontrollgruppa ikke utsettes for de samme stimuli som tiltaksgruppa.

Å bruke strategien med å velge personer til tiltak som en tror lar seg hjelpe er imidlertid ikke uproblematisk. Den medfører en viss risiko for at trygdekontoret velger ut personer som ellers ville klart seg på egen hånd om det ikke var grepet inn. Hvis trygdekontoret ofte setter i verk tiltak overfor personer som ellers ville ha klart seg selv, vil knappe ressurser brukes ineffektivt. Dette spørsmålet om treffsikkerhet er selvsagt vanskelig. I praksis løses dette antakelig ved å bruke opparbeidet kompetanse, erfaring, ferdigheter og sunn fornuft.

Konklusjon

Nytter det å følge opp langtidssykmeldte ved å innkalle dem til personlig samtale på trygdekontoret? Svaret er et betinget ja, og forutsetter at vi aksepterer noe større statistisk usikkerhet enn det som er vanlig. Effekten er trygdepolitisk betydningsfull i den forstand at sykefraværet er 17–19 dager kortere i tiltaksgruppa enn i kontrollgruppa. Fordi sykmeldte i tiltaksgruppa med lange fravær, det vil si over 150 dager, har utbetalt mer per dag enn tilsvarende sykmeldte i kontrollgruppa, er imidlertid forskjellen i utbetalte sykepenger ikke så imponerende, vel 4300 kroner i gjennomsnitt.

Litteratur

- Gueron J.M. og E. Pauly (1991), *From welfare to work*. New York: Russel Sage Foundation
- Hippe J. og K. Hagen, red. (1991), *Alle vet jo at...* Rapport om sykelønn og sykefravær. Fafo-rapport nr. 119. Oslo: Fafo
- Melsom, H. og Noreik (1988), «Gjengangersykmeldte og bruk av rådgivende lege ved Skedsmo trygdekontor.» *Tidsskrift for den Norske Lægeforening*, 108:1013-5
- NOU 1986:22. Oppfølging av langtidssykmeldte. Oslo: Sosialdepartementet
- Pedersen J.T. (1993), *Arbeid eller tvangsarbeid. Resultater av arbeidsmarkedstiltak for unge sosialklienter og rusmiddelmissbrukere*. Oslo: Universitetsforlaget
- Pocock S.J. (1983), *Clinical trials: A practical approach*. Chichester: John Wiley & Sons
- Røed, M. og S. Try (1993), *Evaluering av mobilitetsfremmende stonader*. ISF-rapport 93:8 Oslo: Institutt for samfunnsforskning
- Raaum, O. og H. Torp (1993), *Evaluering av AMO-kurs: Sysselsettingseffekter og seleksjon til kurs*. SNF-rapport 72/93. Oslo: SNF
- Raaum, O., H. Torp og H. Goldstein (1994), *Experiments in manpower policy evaluation: The case for simple estimators? Experiences from a Norwegian study of labour market training*. Oslo: SNF
- Tellnes, G. m. fl. (1992), *Tiltak for sykmeldte*. Oslo: Universitetet i Oslo og Rikstrygdeverket
- Try, S. (1992), *Arbeidsmarkedet og arbeidsmarkedspolitikken for unge i Norden i 1980-årene*. ISF-rapport 92:8 Oslo: Institutt for samfunnsforskning
- Try S. (1993a), *Effects of a youth labour market programme in Norway*. ISF-rapport 93:7 Oslo: Institutt for samfunnsforskning
- Try, S. (1993b), *Evaluering av arbeidsmarkedsrettet mot unge ledige*. ISF-rapport 93:17 Oslo: Institutt for samfunnsforskning
- Wiseman, M. (1995), *Welfare reform in the United States: A background paper*. Madison: The University of Wisconsin-Madison



Fafo

Forskningsstiftelsen Fafo
Borggata 2B
Postboks 2947 Tøyen
N-0608 Oslo

Fafo-notat 1996:3
Bestillingsnummer 861