

EFFEKT AV GRADERT SYKMELDING – BETYDNINGEN AV METODEVALG

Av Jon Petter Nossen og Søren Brage¹

Sammendrag

Flere studier har vist at gradering reduserer sykefraværet mer enn det som følger direkte av at restarbeidsevnen utnyttes. Enkelte studier viser at gradering fører til kraftig reduksjon i varigheten av sykefraværet og senere bruk av trygdeytelser. Formålet med denne artikkelen er å vise hvordan effektstudier kan påvirkes av datautvalg og definisjon av graderte og ikke-graderte sykefravær. Vi presenterer forslag til definisjon og utvalg som vi mener vil gi et bedre grunnlag for registerbaserte effektstudier av gradert sykmelding.

Gradert sykefravær kan defineres på forskjellige måter, og vi har prøvd ulike definisjoner der det ble tatt hensyn til når i forløpet graderingen skjer, hvilken sykmeldingsgrad som brukes og hvor lenge sykefraværet er gradert. Vi drøfter også om noen sykefravær bør utelates fra effektstudier. Grunnen til at man gjør begrensninger i utvalget er først og fremst ønsket om å utelukke sykefravær hvor gradering i praksis ikke er noe alternativ. Man kan også velge å ekskludere grupper av sykmeldte som har begrensninger på hvor lenge de kan motta sykepenger. Det ser ut til at det kan ha stor betydning for effektanalyser hvilke definisjoner som brukes, men enda mer avgjørende er hvilke sykefravær man velger å inkludere i studien.

Vi foreslår i artikkelen en definisjon av «graderte sykefravær» og spesifiserer hvilke begrensninger i utvalget vi mener bør gjøres ved effektanalyser av gradert sykefravær. Med vårt forslag til metode finner vi at graderte sykefravær har 21 dager kortere varighet enn ikke-graderte sykefravær, men det er da ikke kontrollert for seleksjon.

Det er grunn til å anta at de som får gradert sykmelding i gjennomsnitt har bedre helse og arbeidsevne enn dem som får fulltidssykmelding, slik at det ikke er tilfeldig om en person får gradert sykmelding eller ikke. Vi mener at det på bakgrunn av tidligere studier ikke kan trekkes noen sikre konklusjoner angående effekter av gradering utover utnyttelsen av restarbeidsevnen. I mange tilfeller er likevel gradert sykmelding en gunstig ordning som bør benyttes som alternativ til fulltidssykmelding når det er hensiktsmessig.

¹ Takk til Ola Thune for tilrettelegging av datagrunnlaget.

Innledning

De siste årene har gradert sykmelding fått mye oppmerksomhet som et virkemiddel for å redusere sykefraværet (Arbeidsdepartementet 2010). Det er blitt publisert forskning som tyder på at gradering fører til en markant reduksjon både i varigheten av sykefraværet og i den sykmeldtes bruk av trygdeytelser i tiden etterpå (Markussen, Mykletun og Røed 2012; Høgelund m.fl. 2009). Andre studier tyder på en mindre effekt (Andrén og Andrén 2008, Ose m.fl. 2012, Kann m.fl. 2012, Viikari-Juntura m.fl. 2012). En tidligere systematisk oversikt over effektstudier av gradert sykmelding viste blandete resultater (Kausto 2008).

Hensikten med gradert sykmelding er å utnytte restarbeidsevnen til den sykmeldte. Dette er ønskelig av økonomiske grunner, men også fordi det antas å være positivt for den sykmeldte selv å være i en viss aktivitet og å beholde kontakten med arbeidsplassen under sykefraværet. Dermed kan risikoen for varig utestengning fra arbeidslivet reduseres. Så lenge gradert sykmelding erstatter fulltidssykmelding, vil gradering medføre redusert sykefravær i den offisielle sykefraværstatistikken.

Selv om gradert sykmelding slik sett bidrar til redusert sykefravær, må det mer til enn utnyttelsen av restarbeidsevnen for at gradering skal være et kraftfullt virkemiddel mot høyt sykefravær. Hvis gradering faktisk reduserer varigheten av sykefraværet og senere bruk av trygdeytelser betydelig, vil ordningen være et slikt kraftfullt virkemiddel.

Generelt er randomiserte, kontrollerte forsøk den anerkjente metoden for å identifisere reelle effekter av et tiltak. Det er imidlertid ikke mulig å gjennomføre slike studier av gradert sykmelding i Norge, siden det er en lovbestemt ordning som gjelder for alle sykmeldte. Forskingen har derfor benyttet andre og mindre sikre analysemetoder. Dette gjør det desto viktigere at studiene som foretas blir etterprøvd og at problemstillingen belyses på flere måter før det trekkes bastante konklusjoner.

Det er usikkert om man ved bruk av alternative metoder klarer å kontrollere godt nok for seleksjon.

Seleksjon innebærer at sykmeldte personer som får gradert sykmelding på ulike vis skiller seg fra sykmeldte som gis fulltidssykmelding. I utgangspunktet er det grunn til å forvente at de som får gradert sykmelding kan gå fullt tilbake i arbeid i større grad enn dem som får fulltidssykmelding. Dessuten kan gradert sykmelding gis på ethvert tidspunkt i sykefraværet, og til tross for lovverkets aktivitetskrav etter 8 uker, gis gradering oftest i løpet av de første fraværsukene, mens andre får gradering senere enn 8 uker. På grunn av denne kompleksiteten er det viktig at datagrunnlaget er detaljert og at data tilrettelegges på en slik måte at man minimerer risikoen for skjevhet i definisjoner og utvalg. Hvis dette ikke gjøres godt nok, vil man ikke kunne trekke gyldige konklusjoner om effekter.

I denne artikkelen forsøker vi ikke å kontrollere for seleksjon, og vi gjør dermed ingen effektstudie. Formålet er i stedet å belyse hvordan datautvalg og definisjon av «graderte» versus «ikke-graderte» sykefravær kan påvirke hvilken varighet vi måler av henholdsvis graderte og ikke-graderte sykefravær, og dermed potensielt kan påvirke resultatene i en effektanalyse. Vårt ønske er altså å belyse hvor viktig gjennomtenkte definisjoner og dataavgrensninger er for å unngå skjevheter som svekker muligheten til å identifisere reelle effekter av gradering. Vi forsøker også å si noe om hvilke av de alternative metodevalgene vi mener gir et mest mulig «rettferdig» datagrunnlag for effektstudier (Lie 2012).

Til dette formålet benytter vi et datasett over legeomeldte sykefravær med detaljerte opplysninger om hvert enkelt sykefravær og på hvilke tidspunkter det er brukt gradert sykmelding. Slike detaljerte data har til nå stort sett ikke vært tilgjengelige for forskere.

Definisjon av «graderte» og «ikke-graderte» sykefravær

Det finnes ingen klar og utvetydig definisjon av henholdsvis «graderte sykefravær» og «ikke-graderte sykefravær». Det er bare den enkelte sykmelding

Datagrunnlag

Datagrunnlaget for analysen er hentet fra NAVs sykmeldingsregister. En sykmelding er en attestert erklæring fra lege eller annen sykmelder. De fleste gjelder for en kort periode, ofte 14 dager. Sykmeldingen er enten fulltids (100 %) eller gradert med en sykmeldingsgrad fra 20 til 99 %. Utgangspunktet var alle sykmeldinger for arbeidstakere i 2001–2012.

Ved lengre fravær skrives det vanligvis ut flere sykmeldinger etter hverandre. Hele perioden som en person er sammenhengende borte fra arbeidet med én eller flere sykmeldinger kalles et «sykefraværstilfelle». Vi konstruerte sammenhengende sykefraværstilfeller ved å slå sammen alle sykmeldinger for samme person, så lenge de hadde samme dato for første fraværsdag. Siden vi ikke kunne følge alle sykefraværstilfeller som startet i 2012 fram til avslutning, ble grunnlaget for analysen alle legemeldte sykefraværstilfeller som startet i 2001–2011.

I hoveddelen av analysen ser vi bare på sykefraværstilfellene som startet i 2011, men i siste del ser vi på 2001–2005. Kun arbeidstakere er med i datamaterialet. Tilgjengelige kjennetegn var blant annet alder, kjønn, diagnose på siste sykmelding, start- og sluttdato for sykefraværet, samt sykmeldingsgrad og start- og sluttdato for perioder med gradering. Gravide ble identifisert ved hjelp av opplysninger om nyfødte fra Folkeregisteret.

I artikkelen har vi brukt betegnelsen «sykefravær» i samme betydning som «sykefraværstilfelle».

som enten er en fulltidssykmelding eller gradert.¹ Langvarige sykefravær består gjerne av en rekke påfølgende sykmeldinger hvor sykmeldingsgraden kan variere. Slike sykefravær er det ikke like enkelt å klassifisere som graderte eller ikke-graderte. Vi tar utgangspunkt i sykefravær som startet i 2011 og ser på noen mulige definisjoner, samt hvilke konsekvenser det kan ha for analysen hvilket av alternativene man velger.

En enkel definisjon av gradering

En enkel definisjon er at man regner et sykefravær som gradert hvis gradert sykmelding forekommer på et eller annet tidspunkt i løpet av sykefraværet. Med denne definisjonen hadde de graderte sykefraværene som startet i 2011 gjennomsnittlig 92 kalenderdager *lengre varighet* enn de ikke-graderte (tabell 1, første rad).

Den enkle definisjonen tar ikke hensyn til verken hvor tidlig i sykefraværet graderingen starter, den nøyaktige sykmeldingsgraden eller varigheten av graderingen. En slik definisjon kan egne seg til utarbeiding av enkel statistikk, men når formålet er å analysere mulige effekter av gradering på sykefraværets varighet eller på senere bruk av trygdeytelser, er den ikke hensiktsmessig. Vi ser derfor på andre mulige definisjoner.

¹ I praksis er det også mulig å sette flere ulike sykmeldingsgrader på én og samme sykmelding. For eksempel kan det på en 14-dagers sykmelding oppgis at sykmeldingsgraden skal være 100 % den første uken og deretter 50 %. Men dette gjøres sjelden og antas ikke å ha noen betydning for analysen.

Tabell 1. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær. Enkel definisjon og ulike tidspunktdefinisjoner. Sykefravær som startet i 2011

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
Alle sykefravær	1 400 094	17,9	115	23	92
Gradert i løpet av første 12 uker	1 400 094	16,2	99	28	71
Gradert i løpet av første 8 uker	1 400 094	15,2	95	30	65
Gradert i løpet av første 4 uker	1 400 094	13,1	90	32	58
Gradert i løpet av første 2 uker	1 400 094	10,9	91	33	58
Gradert fra 1. fraværsdag	1 400 094	8,3	101	34	67

Kilde: NAV

Tidspunktdefinisjon

Først definerer vi «graderte sykefravær» ut fra tidspunktet for graderingen. Dette har betydning fordi gradert sykmelding iblant brukes etter en lengre periode med fulltidssykmelding. Når vi sammenligner varigheten av graderte og ikke-graderte sykefravær, vil slike tilfeller trekke opp gjennomsnittsvarigheten av graderte sykefravær, men graderingen kan ikke være årsaken til dette. Det er derfor viktig å skille mellom gradering tidlig og sent i sykefraværet. Hva som bør regnes som «tidlig» og «sent» er likevel ikke opplagt, her kan det argumenteres for ulike valg. Markussen m.fl. (2012) har valgt å definere graderte sykefravær ut fra om det ble gitt gradert sykmelding i løpet av de første 8 fraværsukene. En slik definisjon kan vi kalle en tidspunktdefinisjon (tabell 1).

I tabell 1 viser vi konsekvensen av valg av ulike tidspunkt. Hvis vi begrenser de graderte sykefraværene til bare å omfatte gradering i løpet av de første 12 ukene, synker andelen graderte sykefravær og gjennomsnittsvarigheten av graderte sykefravær, mens varigheten av ikke-graderte sykefravær øker. Når vi strammer inn definisjonen av gradert sykefravær til bare å gjelde fravær med gradering i løpet av de første 8 eller 4 ukene synker andelen gradert fravær ytterligere, sammen med varigheten av graderte sykefravær. Men hvis vi strammer inn definisjonen ytterligere, ser vi at varigheten av graderte sykefravær øker igjen – og særlig når vi bare regner med gradering fra første fraværsdag. Dette kan tyde på at det er noe spesielt med dem som får gradert sykmelding fra første fraværsdag.

Gradsdefinisjon

I epidemiologisk forskning sier man at eksponeringen for en viss faktor må ha et minste omfang for at

man kan påregne en effekt. Man kan argumentere for samme tankegang når det gjelder gradering, det vil si at graderingen må ha en viss intensitet (grad) og en viss varighet for å kunne ha en effekt på helse og arbeidsevne. Et enkelt mål som fanger opp både intensitet og varighet er sykefraværets gjennomsnittsgard. Vi kan for eksempel kreve at gjennomsnittsgarden skal være lavere enn 90 prosent, som tilsvarer to uker med 50 prosent sykmelding i løpet av et 10 ukers sykefravær. Dette kan vi kalle en gradsdefinisjon (tabell 2). Vi velger da å utelate fra analysen sykefravær hvor gjennomsnittsgarden er på 90–99 prosent, siden det ikke er naturlig å regne disse som ikke-graderte. Som forventet synker andelen graderte sykefravær, og gjennomsnittsvarigheten av de graderte sykefraværene reduseres også noe. Det samme er tilfelle når vi skjerper definisjonen ytterligere til sykefravær med gjennomsnittsgard under 80 prosent.

I alle alternativene vi har sett på til nå har graderte sykefravær klart lengre varighet enn ikke-graderte. Dette indikerer at det er hvilke sykefravær som tas med i datautvalget som først og fremst har betydning for hvilke sammenhenger man finner mellom gradering og varighet av sykefraværet.

Å gjøre begrensninger i datautvalget

I tillegg til at definisjonen av graderte og ikke-graderte sykefravær har betydning for analysen, spiller også utvalget av sykefravær en viktig rolle. I utgangspunktet ønsker vi å utelate fra analysen sykefravær hvor gradering ikke er et relevant alternativ. Det samme gjelder spesielle grupper av sykmeld-

Tabell 2. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær. Enkel definisjon og ulike gradsdefinisjoner. Sykefravær som startet i 2011

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
Alle sykefravær	1 400 094	17,9	115	23	92
Gjennomsnittsgard <90 %	1 377 749	16,6	107	23	84
Gjennomsnittsgard <80 %	1 338 035	14,1	102	23	79

Kilde: NAV

te som har begrensede rettigheter med hensyn til varigheten av et sykefravær eller som av andre grunner kan påvirke resultatene på en uheldig måte. For å forsøke å utelate sykefravær hvor gradering ikke er relevant, er en mulighet å utelate visse diagnoser der sykefraværet kan forventes å bli kortvarig eller der gradering av andre grunner ikke anses aktuelt, eller rett og slett å utelate alle sykefravær med kort varighet. Vi ser først på begrensninger ut fra varigheten.

Å utelate korte fravær

I tabell 1 og 2 viste vi at gradert sykefravær i gjennomsnitt varer mye *lenger* enn fulltidsfravær. Det skyldes at gradering sjelden er aktuelt ved kortvarige sykefravær, for eksempel på grunn av forkjølelse eller influensa. Så lenge det er klart fra starten at sykefraværet vil bli kortvarig, er det rimelig å anse at gradering ikke er aktuelt og dermed utelate alle disse. Dette kan vi regne med vil være tilfelle i hvert fall når varigheten ikke overstiger 14 dager. Vi begrenser derfor datautvalget til sykefravær med varighet over 14 dager og utelater alle med kortere varighet (tabell 3). Siden det er mange kortvarige sykefravær, fører dette til en sterk nedgang i antall sykefravær i utvalget. I tillegg fører det som ventet til en økning i andelen graderte sykefravær, samt i gjennomsnittsvarigheten av både graderte og ikke-graderte sykefravær. Økningen er imidlertid klart størst for ikke-graderte sykefravær, slik at forskjellen mellom graderte og ikke-graderte sykefravær blir markant mindre enn før. Dersom vi begrenser utvalget gradvis opp til 12 ukers varighet, fortsetter differansen å synke og blir til slutt negativ. Det er derfor tydelig at

de betingelsene man eventuelt velger å sette på varigheten påvirker analysen i sterk grad.

Å utelate visse diagnoser

En annen måte å begrense hvilke sykefravær som skal være med i analysen, er ved å bruke diagnoseopplysninger til å definere hvilke sykefravær som skal regnes som «uaktuelle å gradere». Vi har gjort dette med utgangspunkt i medianvarigheten av hver enkelt diagnose,² altså om flertallet av sykefraværene med en bestemt diagnose har varighet over eller under et visst antall dager. En lignende fremgangsmåte ble brukt i en studie av Markussen m.fl. (2009), hvor de delte inn sykefraværene i «major» og «minor» ut fra om diagnosen hadde en gjennomsnittlig varighet over 16 dager eller ikke. Vi regner altså diagnoser med kortere medianvarighet som «mindre alvorlige tilfeller», hvor vi antar gradering er lite aktuelt og som derfor utelates, mens de med lengre medianvarighet regnes som «mer alvorlige tilfeller» hvor gradering er en aktuell problemstilling. Dette kan begrunnes med at det ved mer langvarige tilstander i de fleste tilfeller er naturlig å vurdere tilrettelegging på arbeidsplassen, redusert arbeidsinnsats og gradert sykmelding. Hvordan denne metoden slår ut er vist i tabell 4. Vi viser tre ulike grader av «antatt alvorlige tilfeller», der den første som er definert ut fra 10 dagers varighet inkluderer flere sykefravær, mens den strengeste er definert ut fra en varighet på 28 dager.

.....

² Alle diagnoser er omkodet til diagnosestandarden ICPC-2.

Tabell 3. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær. Enkel definisjon. Ulike begrensninger av utvalget ut fra varigheten. Sykefravær som startet i 2011

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
Alle sykefravær	1 400 094	17,9	115	23	92
Varighet >2 uker	572 817	40,1	125	64	61
Varighet >4 uker	379 159	51,6	143	102	41
Varighet >8 uker	233 999	61,3	179	167	12
Varighet >12 uker	171 054	64,6	212	216	-4

Kilde: NAV

Tabell 4. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær. Enkel definisjon. Ulike begrensninger av utvalget ut fra diagnosens medianvarighet. Sykefravær som startet i 2011

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
Alle sykefravær	1 400 094	17,9	115	23	92
Medianvarighet >10 dager	694 517	29,0	122	37	85
Medianvarighet >14 dager	475 966	32,2	126	43	83
Medianvarighet >28 dager	129 079	39,6	131	65	66

Kilde: NAV

Alle tre alternativene gir en kraftig reduksjon i antall sykefravær og øker andelen graderte sykefravær. Mest uttalt er reduksjonen når 28 dager brukes som grenseverdi. Som vi også så i tabell 3, øker varigheten av ikke-graderte sykefravær mer enn av graderte, slik at differansen i gjennomsnittlig varighet reduseres noe. Men effekten er mindre enn ved å utelate alle sykefravær under en viss varighet.

Å utelate visse grupper

En tredje begrensning i utvalget som bør vurderes er om bestemte grupper av sykmeldte skal utelates. Gravide er en slik gruppe, siden varigheten av sykefraværet for de fleste gravide er begrenset til maksimalt ni måneder, og siden gravide har et spesielt fraværsmønster. Siden sykepengeperioden begrenses etter fylte 67 år, er det også aktuelt å utelate denne aldersgruppen. Av 1,4 millioner sykefravær som startet i 2011 kan det dermed vurderes å utelate 66 000 på grunn av graviditet og 13 000 på grunn av alder over 66 år.

Personer som har brukt opp en del av sykepenge-rettighetene på grunn av tidligere sykefravær vil ha kortere maksimal varighet av sykefraværet enn 365 dager. For å unngå at dette påvirker analysen på en uheldig måte, kan man også vurdere å utelate disse. Siden det er komplisert å beregne dette presist, valgte vi å utelate personer som startet et sykefravær av minst 60 dagers varighet i 2010, altså året før analyseåret. Dette er en gruppe sykmeldte som har en betydelig avkortning av sine sykepenge-rettigheter i 2011. Antallet sykefravær reduseres da med drøyt

200 000, og andelen graderte sykefravær reduseres fra 17,9 prosent til 17,0 prosent. Vi finner imidlertid ingen merkbar effekt på gjennomsnittlig varighet verken for graderte eller ikke-graderte sykefravær.

Forslag til definisjon og utvalg

Som vi har sett, har valgene man gjør vedrørende graderingsdefinisjon og utvalg ulik påvirkning på gjennomsnittsvarigheten av graderte og ikke-graderte sykefravær. Hvis man ikke er oppmerksom på hvordan de ulike valgene slår ut og ikke begrunner valgene godt, er det derfor en risiko for at datautvalget blir skjevt og uegnet til å identifisere effekter.

Hvilket datautvalg bør legges til grunn?

For å identifisere effekter av gradert sykmelding som alternativ til fulltidssykmelding, er det generelt ønskelig å utelate sykefravær hvor gradering ikke er mulig eller hensiktsmessig. Dette kan være på grunn av smitterisiko, behov for hvile og/eller behandling, svært nedsatt arbeidsevne, eller forventning om rask friskmelding. Det vi derfor ønsker er å sammenligne graderte sykefravær med ikke-graderte sykefravær som *burde vært graderte*, gitt lovverkets sterke føringer i retning gradering. At slike sykefravær likevel ikke graderes kan ha flere årsaker, som lav motivasjon hos den sykmeldte, liten interesse for gradering hos arbeidsgiver, vanskeligheter med tilrettelegging på arbeidsplassen, konflikter eller overtallighet på arbeidsplassen, eller liten interesse for gradering hos lege/sykmelder.

Siden korttidsfravær på inntil 14 dager sjelden graderes (Brage m.fl. 2011), og sykmeldende lege sannsynligvis i stor grad vet hvilke sykefravær som blir så korte, mener vi det er fornuftig å utelate alle fravær på inntil 14 dager i effektstudier av gradert sykmelding. Ved sykefravær på 3–8 uker burde det imidlertid være mulighet for å benytte gradering. Vi har tidligere vist at 70 prosent av de graderte sykefraværene ble gradert i løpet av de første 4 ukene og 82 prosent i løpet av 8 uker (Brage m.fl. 2011). Vi kan derfor ikke se noen sterke argumenter for å utelate alle sykefravær med varighet 3–8 uker, slik det er gjort i noen studier (Markussen m.fl. 2012; Høgelund m.fl. 2009).³ Ved å utelate alle sykefravær på inntil 8 uker er det sannsynlig at man utelater mange ikke-graderte sykefravær som har relativt kort varighet og som *kunne ha vært graderte*. Vi mener altså det er bedre å utelate bare de helt korte tilfellene, for eksempel på inntil 14 dager, siden disse domineres av akutte lidelser med kraftig redusert arbeidsevne eller høy smitterisiko.

Et annet alternativ er å benytte diagnoseopplysninger for å identifisere typisk kortvarige tilfeller, og utelate alle sykefravær med disse diagnosene. Med denne metoden forsøker vi å utelate de mindre alvorlige sykefraværene. Som vist i tabell 4 har dette mindre konsekvenser for forholdet mellom graderte og ikke-graderte sykefravær enn å utelate alle med varighet inntil 14 dager. Dette kan tyde på at en diagnosebasert inndeling i mer og mindre alvorlige sykefravær er beheftet med stor usikkerhet. En mulig årsak kan være at spredningen i alvorlighetsgrad innenfor hver diagnose er stor, og at man i betydelig grad inkluderer mer bagatellmessige tilfeller av tilstander som regnes som alvorlige (eksempelvis mild depresjon) eller ekskluderer alvorlige tilfeller av tilstander som oftest er mindre alvorlige (eksempelvis alvorlig fibromyalgi). Det kan også være andre forklaringer, for eksempel underrapportering av diagnoseendringer i løpet av langvarige sykefravær. Vi synes man bør forsøke å kombinere disse to metodene, det vil si begrense datautvalget ved å utelate både sykefravær med varighet 1–14 dager og

sykefravær med diagnoser som har en medianvarighet på 1–14 dager.⁴

I valget av hvilke sykefravær som skal inkluderes i effektstudier mener vi som nevnt at gravide og personer som er 66 år og eldre bør utelates. Vi ser ingen grunn til å utelate flere på grunnlag av alder. I tillegg vil vi utelate sykmeldte med sykefravær på over 60 dager året før. Vårt forslag til utvalg er således å utelate korttidsfraværet (inntil 14 dager), mindre alvorlige tilfeller (diagnoser med medianvarighet inntil 14 dager), personer med høyt sykefravær året før, gravide, samt personer over 66 år.

Hvordan bør graderte sykefravær defineres?

Vi har sett at den enkle definisjonen av graderte sykefravær ikke egner seg for effektstudier, og at en definisjon ut fra tidspunktet for gradering og/eller den gjennomsnittlige sykmeldingsgraden er bedre alternativer. Siden den førstnevnte metoden fanger opp tidspunktet for gradering, mens den andre fanger opp selve sykmeldingsgraden og varigheten av graderingen, synes vi det virker fornuftig å kombinere dem. Vi velger altså å regne et sykefravær som gradert hvis det graderes i løpet av de første 14 fraværsdagene og har en gjennomsnittlig sykmeldingsgrad på under 90 prosent. Vi velger 90 prosent som terskel for ikke å utelate altfor mange sykefravær med et visst innslag av gradering.

Når det gjelder tidspunktet for gradering, velger vi samme tilnærming som Markussen m.fl. (2012), det vil si at når vi varierer hvordan vi begrenser data ut fra varighet, fastsetter vi tidspunktdefinisjonen av et gradert sykefravær ut fra samme tidspunkt. Når sykefravær med varighet inntil to uker utelates, definerer vi dermed graderte sykefravær ut fra om de ble gradert i løpet av de to første ukene, og tilsvarende når vi velger 4, 8 og 12 uker (tabell 5).

Tabell 5 viser resultatene med denne metoden for sykefravær som startet i 2011. Vi mener alternativet

³ Et argument for dette valget kan være at lovverket krever at aktivitet skal vurderes senest etter 8 ukers sykefravær. Gradering kan imidlertid være aktuelt langt tidligere.

⁴ Selv om vi mener at varighetsbegrensningen bare bør være på 14 dager, ønsker vi også å se på hvordan det vil slå ut om vi gjør andre valg, siden flere studier som nevnt har utelatt alle med varighet inntil 8 uker.

Tabell 5. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær. Ulike tidspunkt for definisjon av graderte sykefravær og hvilke varigheter som inkluderes. Sykefravær som startet i 2011

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
>2 uker	189 370	17,5	109	96	13
>4 uker	130 623	26,6	124	136	-12
>8 uker	82 744	37,0	166	199	-33
>12 uker	61 627	43,0	205	244	-39
Utelater sykefravær med gradert sykmelding fra første dag					
>2 uker	165 507	5,6	64	96	-32
>8 uker	69 669	25,2	138	199	-61

Note: Uten gravide, personer over 66 år, personer med 60 dagers sykefravær eller mer påbegynt i 2010, sykefravær der diagnosen har medianvarighet < 15 dager og sykefravær med gjennomsnittlig sykmeldingsgrad på 90–99 %.

Kilde: NAV

med varighet over 14 dager er å foretrekke; dette alternativet er derfor skrevet med fete typer. Differansen mellom gjennomsnittsvarigheten av graderte og ikke-graderte sykefravær er kraftig redusert fra det vi har sett tidligere, men fortsatt positiv (13 kalenderdager lengre varighet for graderte sykefravær). Fortegnet snur dersom vi betinger på 4 ukers varighet, men det er først når vi betinger på 8 ukers varighet at de graderte sykefraværene blir tydelig kortere enn de ikke-graderte.

At vårt foretrukne valg gir *lengre* varighet av de graderte sykefraværene er overraskende, da det som nevnt er grunn til å anta at graderte sykefravær som følge av seleksjon uansett vil ha kortere gjennomsnittsvarighet enn ikke-graderte sykefravær. Siden vi så i forbindelse med tabell 1 at personer som får gradert sykmelding fra første dag har spesielt lange sykefravær, forsøkte vi å utelate disse. Dette er vist i de to nederste radene i tabell 5. De graderte sykefraværene får da en markant reduksjon i gjennomsnittsvarigheten. Når vi velger 2-ukersalternativet, blir de graderte sykefraværene nå i gjennomsnitt 32 dager kortere enn de ikke-graderte.

Det kan diskuteres hvor heldig det er å fastsette tidspunktdefinisjonen av et gradert sykefravær ut fra samme tidspunkt som vi bruker til å begrense data ut fra varighet. Metoden avviker fra tankegangen med

randomiserte kontrollerte forsøk, der tiltaksgruppen og kontrollgruppen skal være mest mulig like før tidspunktet da de enten får behandling eller ikke. Avviket skyldes at man inkluderer sykefravær som har pågått i 14 dager, og samtidig inndeler disse ut fra om de hadde gradert sykmelding eller ikke *i den samme perioden*. Vi prøvde derfor også en alternativ metode der vi kun inkluderte sykefravær med fulltidssykmelding i minst 14 dager. Dette vil da gi et likt utgangspunkt, samtidig som det fjerner sykefraværene med gradering tidlig i forløpet, som har et spesielt fraværsmønster. Vi definerte gradering ut fra om det ble gitt gradert sykmelding *neste dag*, som i vår foretrukne spesifisering betyr den 15. fraværsdagen. Vi inkluderte bare sykefravær med minst 15 dagers varighet.⁵ Med denne metoden fikk vi en langt lavere andel graderte sykefravær. Vi fant at graderte sykefravær i gjennomsnitt var 21 kalenderdager kortere enn ikke-graderte sykefravær (tabell 6). I tabellen introduseres de ulike begrensningene som er nevnt tidligere trinnvis.⁶ Uten noen ekstra betingelser finner vi at de graderte fraværene er

⁵ Med denne metoden er det ikke nødvendig å ta utgangspunkt i akkurat 14 dager med fulltidssykmelding, men vi valgte å holde oss til dette for at det skal være mest mulig sammenlignbart med resultatene i tabell 5.

⁶ Betingelsen vedrørende gjennomsnittlig sykmeldingsgrad er her ikke relevant.

6 dager kortere enn de ikke-graderte i gjennomsnitt, men forskjellen øker altså til 21 dager når vi iverksetter alle betingelsene.

Vi mener denne metoden er den som egner seg best ved effektstudier, og bør kombineres med en god metode for å kontrollere for seleksjon. Vårt forslag til definisjon av «graderte sykefravær» ved effektstudier er altså å inkludere sykefravær med minst 15 dagers varighet, med fulltidssykmelding de første 14 dagene og der gradering defineres ut fra den 15. fraværsdagen.

I neste avsnitt sammenholder vi dette med en norsk effektstudie som har vist et ganske annerledes resultat.

Den norske effektstudien av gradert sykmelding

Den viktigste norske effektstudien av gradert sykmelding er gjort av Markussen, Mykletun og Røed (2012). Studien benytter individdata og viser sterke effekter av gradering både i form av redusert varighet av sykefraværet, redusert senere bruk av trygdeytelser og økt framtidig yrkesaktivitet. Den er basert på registerdata over legemeldt sykefravær for perioden 2001–2005, inkluderer bare sykefravær med varighet over 8 uker, og definerer graderte sykefravær ut fra om gradering ble gitt innen utgangen av den 8. fraværsuken. Sykefravær som graderes senere enn dette eller ikke graderes, defineres som ikke-graderte. De begrenser sitt utvalg ytterligere til aktive arbeidsforhold ifølge arbeidsgiver-/arbeidstakerregisteret, personer som hadde en inntekt

over 2G året før sykefraværet startet, ikke mottok sykepenger de siste 6 måneder før og ikke mottok andre helserelaterte ytelser. De utelater gravide og sykefravær hvor verken den første eller siste sykmeldingen ble skrevet ut av den sykmeldtes egen fastlege. 37 prosent av datautvalget fjernes som følge av den sistnevnte betingelsen. De sitter da igjen med et utvalg på 339 000 personer, og de graderte sykefraværene har i gjennomsnitt hele 60 dager kortere varighet enn de ikke-graderte (se tabell 7, rad D). Dette er langt større forskjell enn det vi har sett i vår analyse basert på 2011-data. Etter å ha forsøkt å kontrollere for seleksjon, blir forskjellen enda litt større.

En medvirkende forklaring på forskjellen (før kontroll for seleksjon) er at Markussen og medforfattere har utelatt alle sykefravær som varer inntil 8 uker. Dette er begrunnet med at kortvarige sykefravær domineres av influensa o.l., at graderingsandelen stiger med økende varighet på sykefraværet, samt at gradert sykmelding fra 2004 av skal være hovedalternativet etter 8 ukers sykefravær. På bakgrunn av gjennomgangen over, er det imidlertid vår oppfatning at forfatterens valg av 8 uker vil ha som konsekvens at de overvurderer den gjennomsnittlige varigheten, og særlig av de ikke-graderte sykefraværene. Årsaken til at graderingsandelen øker ved økende varighet er at mange av de ikke-graderte sykefraværene avsluttes før 8 uker. Ved å utelate alle sykefravær med varighet 8 uker eller mindre, skapes dermed en skjevhet i datamaterialet til fordel for gradert sykefravær.

Tabell 6. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær. Fulltidssykmelding første 14 dager og varighet minst 15 dager. Et sykefravær defineres som gradert hvis gradering ble gitt den 15. fraværsdagen. Gradvis innfasing av øvrige betingelser. Sykefravær som startet i 2011

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
Ingen ekstra betingelser	440 632	1,1	77	83	-6
Uten gravide	417 027	1,1	76	84	-8
..og 66 år+	410 939	1,1	77	84	-7
....og >60 dager fravær 2010	342 715	1,1	75	82	-7
.....og medianvarighet <15 dager	166 449	1,3	82	103	-21

Kilde: NAV

Tabell 7. Gjennomsnittlig varighet av graderte og ikke-graderte sykefravær på over 8 uker. Utvalg og definisjoner tilnærmet som i Markussen m.fl. (2012). Uten gravide. Sykefravær som startet i mai 2001 – desember 2005

	Antall	Andel graderte (%)	Gjennomsnittlig varighet (dager)		
			Graderte	Ikke-graderte	Differanse
A. Våre data – definerer gradering på samme måte som i Markussen m.fl. (2012), men mer nøyaktig	899 156	26,1	171	194	-23
B. Våre data – definerer gradering på samme måte som i Markussen m.fl. (2012), med samme estimat av graderingstidspunkt*	899 156	21,9	161	196	-35
C. Som B. men utelater sykefravær hvor sist registrerte diagnose er en ICD-10 diagnose	784 753	23,7	162	203	-41
D. Markussen m.fl. (2012)	339 251	22,9	138	198	-60

*: Markussen m.fl. (2012) kjente ikke det nøyaktige graderingstidspunktet, og deres inndeling i graderte og ikke-graderte sykefravær er derfor unøyaktig. Blant annet måtte de regne sykefravær som startet med fulltidssykmelding, fortsatte med gradert sykmelding og endte med fulltidssykmelding, som ikke-graderte uansett når de ble graderte.

Kilde: NAV

For nærmere å undersøke hva forskjellene ellers kan skyldes, prøvde vi å gjenskape deres utvalg. Vi brukte samme tidsperiode og samme definisjon av gradering som Markussen og medforfattere, men det var en del andre forskjeller i materialet. Vi fant da en differanse mellom gjennomsnittsvarigheten av graderte og ikke-graderte fravær på 23 dager (tabell 7, rad A), sammenlignet med 60 dager hos Markussen og medforfattere.

En del av forskjellen kan forklares med at Markussen og medforfattere hadde mindre detaljerte opplysninger om gradering, noe som gjorde at de ikke kunne identifisere graderingstidspunktet nøyaktig. Når vi beregner tidspunktet på samme måte, reduseres forskjellen med en tredjedel (tabell 7, rad B).

Selv om vi ikke kan utelukke at andre forskjeller i datagrunnlaget har betydning, skyldes den resterende forskjellen trolig i hovedsak begrensningen av utvalget til sykefravær hvor første og/eller siste sykmelding ble skrevet av den sykmeldtes fastlege. Hvis studien hadde inkludert også de som er sykmeldt av andre enn fastlegen, hadde forskjellene mellom de graderte og ikke-graderte sykefraværene

vært mindre.⁷ Det ser ut til at denne begrensningen i datautvalget fører til at man utelater en del langvarige sykefravær som ble gradert innen 8 uker.

Uansett viser også våre tall at sykefravær med gradering før den 8. fraværsuken i gjennomsnitt har kortere varighet enn sykefravær som ikke graderes i den perioden. Dette er som nevnt ikke overraskende i lys av at sykmeldte som får gradering må antas å være nærmere å komme fullt tilbake i jobb enn dem som ikke får gradering. Med andre ord kan forskjellen skyldes seleksjon. Markussen og medforfattere forsøker å kontrollere for seleksjon med en instrumentvariabel. De estimerer en indikator for den enkelte fastleges graderingstilbøyelighet, bruker denne til å forutsi hvem av de sykmeldte som får gradering, og bruker den predikerte graderingsvariabelen i analysen i stedet for den faktiske. Vi er skeptiske til denne metoden. For det første fordi vi mener det er sannsynlig at en leges graderingstilbøyelighet er positivt korrelert med den samme legens friskmeldingstilbøyelighet, altså at en lege som graderer ofte også er raskere til å friskmelde. For det

.....
⁷ Dette undersøkte vi ved å utelate sykefravær hvor den siste registrerte diagnosen er en ICD-10 diagnose, altså en spesialistdiagnose, noe som sjelden brukes av fastleger (tabell 8, rad C). Differansen øker da til 41 dager.

andre synes vi ut fra analysen over at det er uheldig å beregne indikatoren med utgangspunkt i data med begrensning på 8 ukers varighet og gradering i løpet av de første 8 ukene. Siden forsøket på å kontrollere for seleksjon med instrumentvariabelen heller ikke ser ut til å påvirke resultatene av studien, mener vi det er grunn til å vurdere hvor godt metoden egner seg til effektanalyser av gradert sykmelding.

Gradert sykmeldte har bedre helse og arbeidsevne

Gradert sykmelding er en gammel ordning. Muligheten til å motta halve sykepenger ved delvis arbeidsuførhet ble innført med Lov om syketrygd i 1930 (Brage m.fl. 2011). Til tross for at ordningen har blitt brukt i betydelig grad, ikke minst etter sykmeldingsreformen i 2004, er sykefraværet i Norge fortsatt høyt ut fra internasjonale sammenligninger.

Vi har i denne artikkelen kun undersøkt konsekvensene av ulike valg av definisjoner og utvalg. Når man skal forsøke å identifisere mulige effekter av gradert sykmelding som alternativ til fulltidssykmelding, er det også nødvendig med en god metode for å skille mellom reelle årsakssammenhenger mellom gradering og varighet og sammenhenger som bare skyldes seleksjon. I denne forbindelse innebærer seleksjon at sykefravær hvor det blir gitt gradert sykmelding på ulike vis skiller seg fra sykefravær hvor det blir gitt fulltidssykmelding. Det kan dreie seg om observerbare kjennetegn ved individet som alder, kjønn, utdanning, yrke og familiesituasjon, eller mer uobserverbare kjennetegn som generell helsetilstand, alvorlighetsgrad av sykdommen, arbeidsevne og motivasjon. Det kan også dreie seg om forhold ved arbeidet, som i hvilken grad arbeidsgiveren er i stand til eller villig til å tilrettelegge for at de skal komme fullt tilbake i arbeid. Det kan også tenkes at det er forskjeller knyttet til legens sykmeldingspraksis. Med tanke på disse forholdene – ikke minst helsetilstand og arbeidsevne – og når vi ser bort fra korttids-sykefravær hvor gradering er lite aktuelt, skulle man tro at de som får gradert sykmelding jevnt over vil være klare for å komme fullt tilbake i arbeid i større grad enn dem som får fulltidssykmelding. Personer med gradert sykmelding, alt annet likt, kan derfor

forventes å ha kortere varighet av sine fravær enn de fulltidssykmeldte. Hvis det ikke kontrolleres godt nok for slik seleksjon, vil det som framstår som en effekt ikke nødvendigvis være det.

Generelt er randomiserte, kontrollerte forsøk den anerkjente metoden for å avdekke reelle årsakseffekter (Viikari-Juntura 2012). Siden gradert sykmelding er innført i det norske lovverket og gjelder for alle sykmeldte, er det ikke mulig å foreta randomiserte forsøk av dette. Alternative analysemetoder kan benyttes, men det er da viktig at datautvalget foretas på en slik måte at det ikke oppstår utvalgsskjevhet.

Det finnes åpenbart flere utfordringer knyttet til bruken av gradert sykmelding. For det første vil det ofte oppstå praktiske problemer på arbeidsplassene når de delvis sykmeldtes arbeidskraft skal erstattes (Ose m.fl. 2012). For det andre er det en risiko for at økt bruk av gradert sykmelding kan bidra til å senke terskelen for sykmelding. Denne risikoen vil særlig være til stede hvis det er vanlig med sykmeldingsgrader under 50 prosent. Vi har tidligere vist at slike sykmeldingsgrader utgjør mer enn en fjerdedel av de graderte sykmeldingene (Brage m.fl. 2011), men i en senere analyse kunne vi ikke påvise noen terskeleffekt som følge av gradering (Kann m.fl. 2012). For det tredje kan det for arbeidsgiveren være lønnsomt at en arbeidstaker på gradert sykmelding fortsetter å være sykmeldt framfor å komme fullt tilbake i arbeid med risiko for tilbakefall, slik at arbeidsgiver må dekke sykepenger i en ny arbeidsgiverperiode. Slik er det også ved fulltids sykefravær, men problemstillingen synliggjøres kanskje enda mer når arbeidstakeren er delvis i arbeid og delvis sykmeldt. Det er også en risiko for at det gis mindre oppfølging til personer på gradert sykmelding enn til fulltidssykmeldte; blant annet var det før reglene ble endret i 2011 ikke obligatorisk med dialogmøte for personer på gradert sykmelding.

Selv om det altså er flere grunner til at man bør unngå overutnyttelse av gradert sykmelding, er det viktigste spørsmålet om gradert sykmelding har gunstige effekter på varigheten av det enkelte sykefravær og på framtidig yrkesaktivitet og mottak av trygdeytelser. Vi mener kunnskapsgrunnlaget for

øyeblikket ikke er sikkert nok til å konkludere med at gradert sykmelding reduserer varigheten av sykefraværene, eller har andre effekter utover den direkte utnyttelsen av restarbeidsevnen som ordningen i utgangspunktet innebærer. Siden randomiserte kontrollerte forsøk ikke er mulig å gjennomføre, og det er vanskelig å kontrollere for seleksjon med rene registerdata, kan en mulig vei å gå være å koble registerdata med mer kvalitative data fra for eksempel Levekårsundersøkelsen (Statistisk sentralbyrå) eller mer omfattende helseundersøkelser (HUNT).

Vår konklusjon er at gradert sykmelding er en gunstig ordning som bør benyttes som alternativ til fulltidssykmelding når det er hensiktsmessig. Det er likevel en risiko for at ordningen brukes mer enn det som er forsvarlig, og det er grunn til å advare mot å tro at gradert sykmelding kan løse alle utfordringene knyttet til det høye sykefraværet i Norge.

Referanser

Andrén, Daniela og Thomas Andrén (2008). *Part-time sick leave as a treatment method?* Working paper in economics 320. Göteborg Universitet.

Arbeidsdepartementet (2010). *Tiltak for reduksjon i sykefravær: Aktiviserings- og nærværsreform*. Ekspertgrupperapport til Arbeidsdepartementet 01.02.10 ifølge mandat av 27.11.09.

Brage, Søren, Inger Cathrine Kann, Arne Kolstad, Jon Petter Nossen og Ola Thune (2011). «Gradert sykmelding – omfang, utvikling og bruk». *Arbeid og velferd*, 3/2011, 32–44.

Høgelund, Jan, Anders Holm og James McIntosh (2009). *Does graded return to work improve disabled workers' labor market attachment?* Working paper no. 2009–06. København: Centre for Applied Microeconometrics

Kann, Inger Cathrine, Søren Brage, Arne Kolstad, Jon Petter Nossen og Ola Thune (2012). «Har gradert sykmelding effekt på sykefraværet?». *Arbeid og velferd*, 2/2012, 60–70.

Kausto, Johanna, Helena Miranda, Kari-Pekka Martimo og Eira Viikari-Juntura (2008). «Partial sick leave – review of its use, effects and feasibility in the Nordic countries». *Scandinavian Journal of Work Environment & Health*, 34 (4), 239–249.

Lie, Stein Atle (2012). *Forløp av sykefravær og retur til jobb. Sluttrapport til FARVE*. Bergen: Uni Helse

Markussen, Simen, Knut Røed, Ole J. Røgeberg og Simen Gaure (2009). *The anatomy of absenteeism*. IZA Discussion paper no. 4240. Oslo: Frischsenteret

Markussen, Simen, Arnstein Mykletun og Knut Røed (2012). «The case for presenteeism – evidence from Norway's sickness insurance program». *Journal of Public Economics*, 96 (11–12), 959–972.

Ose, Solveig Osborg, Silje L. Karspersen, Silje Haus Reve, Roland Mandal, Heidi Jensberg og Jan Lippestad (2012). *Sykefravær – gradering og tilrettelegging*. Rapport A22 397. Trondheim: SINTEF.

Viikari-Juntura, Eira, Johanna Kausto, Rahman Shiri, Leena Kaila-Kangas, Esa-Pekka Takala, Jaro Karppinen, Helena Miranda, Ritva Luukkonen og Kari-Pekka Martimo (2012). «Return to work after early part-time sick leave due to musculoskeletal disorders: a randomized controlled trial». *Scandinavian Journal of Work, Environment & Health*, 38 (2), 134–143.

REPLIKK TIL ARTIKKELEN «EFFEKT AV GRADERT SYKMELDING»

Simen Markussen (Frischsenteret), Arnstein Mykletun (Nasjonalt folkehelseinstitutt), Knut Røed (Frischsenteret)

I artikkelen «The case for presenteeism — Evidence from Norway's sickness insurance program» (Markussen *et. al.*, 2012) har vi vist at det er svært positive effekter forbundet med bruk av gradert sykmelding for langtids sykmeldte arbeidstakere i Norge. Ved å erstatte fulltids sykmelding med gradert sykmelding oppnås en reduksjon i varigheten av sykefraværet, en reduksjon i bruken av trygdeytelser i de nærmeste årene etter at sykefraværet er avsluttet, og (ikke minst) en markert økning i sannsynligheten for å forbli sysselsatt etterpå.

På grunn av de store seleksjonsproblemene forbundet med å anslå effekter av gradert sykmelding benyttet vi i vår analyse en såkalt instrumentvariabel-metode, der effektene estimeres på bakgrunn av fastlegenes tilbøyelighet til å gradere *andre pasienters* sykmelding – og ikke på bakgrunn av hvorvidt egen sykmelding ble gradert eller ikke. Dermed påvirkes ikke våre resultater av at de som får gradert sykmelding eventuelt har bedre (eller dårligere) arbeidsmarkedsutsikter enn de som får fulltids sykmelding.

I artikkelen «Effekt av gradert sykmelding – betydningen av metodevalg» i dette nummeret av *Arbeid og Velferd*, uttrykker Jon Petter Nossen og Søren Brage skepsis til vår studie. Deres hovedinnvending er at vi har valgt å analysere fravær som varer minst 8 uker. De påpeker at svært mange av de ikke-graderte fraværene avsluttes før 8 uker, og at vi ved å utelate kortere fravær dermed skaper en skjevhet i datamaterialet til fordel for gradert fravær. Men etter som den statistiske metoden vi benytter kontrollerer for seleksjonen inn i gradert fravær er ikke denne innvendingen relevant for vår analyse. Det kan naturligvis godt tenkes at *effektene* av gradert sykmelding er mer gunstig for lange enn for korte fravær. Hvorvidt korte fravær bør inkluderes eller ikke avhenger i så fall av hvilket forskningsspørsmål

man ønsker å besvare. Det finnes ikke noe «riktig» datautvalg, slik Nossen og Brage synes å mene. Vi ønsket å fokusere på fravær over 8 uker fordi det er disse fraværene som i henhold til det norske regelverket som hovedregel skal være graderte. Langtids-sykefravær er også relevant fordi risikoen for aldri å returnere til arbeidslivet øker med økende varighet av sykefravær. Derfor benyttet vi en 8-ukersbegrensning i hoveddelen av vår analyse. Nossen og Brage unnlater imidlertid å nevne at vi i vår artikkel også gjennomfører en full statistisk analyse av fravær som har vart minst 4 uker og av fravær som har vart minst 12 uker. De gunstige effektene er ganske riktig større jo lengre vi betinger på at fraværene har vart, men selv med en varighetsbegrensning på bare 4 uker finner vi svært positive effekter av gradert sykmelding.

Nossen og Brage uttrykker også skepsis til vår statistiske metode. De mener at leger som har høy tilbøyelighet til å gradere sykmeldinger også kan antas å være raskere til å friskmelde sine pasienter. I så fall kan det i noen grad være dette vi fanger opp effekten av – og ikke bare graderingen. Dette er en relevant innvending, spesielt når varigheten av sykefraværet benyttes som utfallsmål. I vår studie finner vi imidlertid ikke bare effekter på varigheten av sykefraværet. Den kanskje viktigste effekten av gradert sykmelding som avdekkes i vår analyse er at sannsynligheten for å falle helt ut av arbeidslivet (to år senere) reduseres med mer en 10 prosentpoeng (det gjelder også analysen der vi benytter alle fravær som varer minst 4 uker).

Det er for øvrig tenkelig at tilbud om gradert sykmelding i en del tilfeller kan «presse fram» en full friskmelding framfor faktisk realisering av en gradert sykmelding, ettersom noen arbeidstakere kan oppfatte gradert sykmelding som et lite attraktivt

alternativ. I så fall vil vår metode overvurdere de gunstige effektene for de personene som faktisk får en gradert sykmelding, men vi vil fortsatt ha riktig anslag på de totale effektene av å satse på graderte sykmeldinger framfor fulltidssykmeldinger.

Spørsmålet om vår «graderingsindikator» kan være korrelert med uobserverte kjennetegn ved legenes atferd diskuteres for øvrig grundig i vår artikkel, og i de statistiske analysene forsøker vi å kontrollere for andre trekk ved legenes praksisstil (bl.a. basert på deres observerte atferd overfor andre pasientgrupper).

Så lenge vi må forholde oss til ikke-eksperimentelle data er det naturligvis vanskelig å gi et eksakt og udiskutabelt svar på hva som er effektene av gradert sykmelding. På grunn av de store seleksjonsproblemer har det imidlertid liten eller ingen hensikt å granske den observerte samvariasjonen mellom gradering og varighet av sykefravær slik Nossen og Brage her gjør. For å kunne si noe om årsakssammenhenger må seleksjonsproblemet tas på alvor. Vår studie er et av nokså få forsøk på å gjøre nettopp dette. Det er for øvrig verd å merke seg at også de andre publiserte (fagfellevurderte) studiene som har gjort dette kommer til lignende resultater som oss, selv om de bygger på helt andre metodiske tilnærminger. Det gjelder både den finske studien til Viikari-Juntura *et al* (2012), som er basert på et randomisert kontrollert forsøk, og den danske ikke-eksperimentelle studien til Høgelund *et al* (2010), som bygger på forløpsanalyse (såkalt timing-of-events approach). Det gjelder for øvrig også den norske studien til Kann *et al* (2012), som Brage og Nossen var medforfattere på, og som basert på registerdata på kommunenivå viser en negativ sammenheng mellom omfanget av gradert sykmelding og totalt sykefravær.

Nossen og Brage redegjør innledningsvis helt korrekt for at deres deskriptive analyser ikke inneholder noe forsøk på å håndtere seleksjonseffekter og dermed heller ikke er egnet til å trekke slutninger om kausale effekter av gradert sykmelding. Det framstår derfor litt underlig når Nossen og Brage – med henvisning til en blanding av deskriptiv statistikk og nokså løs argumentasjon – mener å kunne fastslå at kunnskapsgrunnlaget i dag er for dårlig til å kunne si om gradert sykmelding reduserer varigheten av sykefraværene eller ikke.

Referanser

- Høgelund, J., Holm, A., McIntosh, J. (2010) «Does Graded Return-to-Work Improve Sick-Listed Workers' Chance of Returning to Regular Working Hours?» *Journal of Health Economics*, Vol. 29, 158-169.
- Kann, I.C., Brage, S., Kolstad, A., Nossen, J.P., & Thune, O. (2012). Har gradert sykmelding effekt på sykefraværet? *Arbeid og velferd* nr. 2, 60-70.
- Markussen, S., Mykletun, A., Røed, K. (2012) «The case for presenteeism — Evidence from Norway's sickness insurance program». *Journal of Public Economics*, Vol. 96, 959-972
- Viikari-Juntura E., Kausto J., Shiri R., Kaila-Kangas L., Takala E.-P., Karppinen J., Miranda H., Luukkonen R., Martimo K.-P. (2012) «Return to Work after Early Part-Time Sick Leave Due to Musculoskeletal Disorders: A Randomized Controlled Trial». *Scandinavian Journal of Work, Environment & Health*, Vol. 38, No. 2, 134-143.