

**SINTEF Unimed**

Postadresse: 7465 Trondheim  
Besøksadresse: Olav Kyrres gt. 3  
Telefon: 73 59 25 90  
Telefaks: 73 59 63 61

Foretaksregisteret: NO 948 007 029 MVA

# SINTEF RAPPORT

TITTEL

**Er det sammenfall mellom journalopplysninger og innrapporterte data? En studie av 500 pasientopphold ved norske somatiske sykehus i 2001**

FORFATTER(E)

Linda Midttun, Erik Sverrbo, Glen Thorsen og Olafr Steinum

OPPDRAGSGIVER(E)

Avregningsutvalget

RAPPORTNR. STF78 A035504	GRADERING Åpen	OPPDRAGSGIVERS REF. Bjørn Buan	
GRADER. DENNE SIDE Åpen	ISBN 82-14-03-224-5	PROSJEKTNR. 78550606	ANTALL SIDER OG BILAG 74 inkl bilag
ELEKTRONISK ARKIVKODE I:\7855 NPR\Prosjekt\78550606\Rapport\Ferdig rapport		PROSJEKTLÉDER (NAVN, SIGN.) Linda Midttun	VERIFISERT AV (NAVN, SIGN.) Ronny Jørgenvåg
ARKIVKODE E	DATO 2003-06-05	GODKJENT AV (NAVN, STILLING, SIGN.) Jon Magnussen, prosjektdirektør	

**SAMMENDRAG**

Rapporten er basert på en gjennomgang av 500 journaler ved fem norske somatiske sykehus. Analysen av dataene viser at det var betydelig diskrepans mellom opprinnelige registreringer og reviderte registreringer. 208 av 500 opphold fikk revidert hoveddiagnose, mens 166 opphold endret DRG-plassering. Endringene i DRG var i all hovedsak et resultat av revisjon av hoveddiagnoser, men for 32 tilfeller var revisjon av bidiagnoser den direkte årsaken til endringen. Totalt sett ga den opprinnelige DRG-registreringen flere DRG-poeng enn den reviderte. Til tross for de relativt store avvikene for hoveddiagnoser, bidiagnoser og DRG, var avvikene i DRG-poeng totalt på 3,7 prosent. Variasjonene mellom de fem sykehusene var imidlertid relativt store.

Studien viser at diskrepans forekom hyppigst blant opphold i medisinske DRG-er og oftere blant akuttinnleggelser enn elektive opphold.

Resultatene avdekker et behov for oppfølgingsstudier og indikerer at det er behov for å iverksette tiltak for å forbedre kodingskvaliteten ved sykehusene.

STIKKORD	NORSK	ENGELSK
GRUPPE 1	Sykehus	Hospital
GRUPPE 2	Pasientjournaler	Medical records
EGENVALGTE	Diagnoserelaterte grupper	Diagnosis Related Groups
	Hoveddiagnose	Principal diagnosis
	Kodingskvalitet	Accuracy of coding

# Forord

Denne rapporten presenterer resultater fra en journalgjennomgang av 500 tilfeldig utvalgte journaler fra sykehusopphold ved fem norske somatiske sykehus i 2001. Datainnsamlingen ble gjennomført i henhold til forutsetningene gitt i konsesjon fra Datatilsynet og dispensasjon fra taushetsplikt gitt fra Sosial- og helsedirektoratet.

Formålet har vært å kartlegge omfanget av feilregistreringer og konsekvensene disse har. Prosjektet ble utført på oppdrag for Avregningsutvalget, og er gjennomført som et samarbeid mellom SINTEF Unimed og KITH (Kompetansesenter for IT i Helsevesenet). Prosjektet har i sin helhet vært finansiert av oppdragsgiver.

Vi ønsker å rette en takk til de fem sykehusene som har deltatt i prosjektet, og spesielt til kontaktpersonene som tilrettela for at datainnamlingsprosessen kunne gjennomføres på en effektiv måte. Vi er også takknemlige for den kvalitetssikringen av prosessen og rapporten som Hroar Piene, Jon Magnussen, Lars Rønningen og Øyvind Hope har bidratt med.

Trondheim, juni 2003.

Linda Midttun

Erik Sverrbo

Glen Thorsen

Olafr Steinum

# Innholdsfortegnelse

Forord .....	1
Innholdsfortegnelse .....	3
Liste over tabeller.....	5
1    Innledning .....	7
2    Bakgrunn.....	9
2.1    Tidligere studier.....	9
2.2    Finansiering .....	9
2.3    Utviklingen av DRG-systemet og kritikken rettet mot det .....	10
3    Metode .....	15
3.1    Utvalg .....	15
3.2    Metode for datainnsamling .....	16
3.3    Metodiske utfordringer.....	17
3.4    Reliabilitet og validitet .....	18
4    Resultater .....	19
4.1    Hoveddiagnoser .....	19
4.2    Bidiagnoser .....	21
4.3    Kirurgiske prosedyrekoder.....	22
4.4    Plassering i HDG .....	22
4.5    Plassering i DRG .....	24
4.6    Økonomiske konsekvenser .....	25
4.7    Innleggelsesmåte og behandlingstype .....	26

5	Analyse .....	29
5.1	Hoveddiagnoser .....	29
5.2	Bidiagnoser .....	30
5.3	Kirurgiske prosedyrekoder .....	31
5.4	HDG .....	31
5.5	DRG .....	31
5.6	Avvik mellom opprinnelig og revidert diagnosesetting .....	33
5.7	Praktiske konsekvenser av lav reliabilitet på data .....	33
6	Oppsummering .....	35
	Litteraturliste .....	37
	Appendix .....	39

## Liste over tabeller

Tabell 2.1	HDG for pasienter behandlet ved norske somatiske sykehus i 2001 (Sosial- og helsedepartementet, 2001) .....	11
Tabell 3.1	Prosentandel kompliserte DRG-er og gjennomsnittlig antall bidiagnoser ved de utvalgte sykehusene og i totalpopulasjonen .....	16
Tabell 3.2	Forventet standardavvik for observerte prosentandeler ved ulike utvalgsstørrelser .....	17
Tabell 4.1	Antall reviderte hoveddiagnoser per sykehus og fordeling av disse i henhold til feil-type. Kategoriens andel feil av totalt antall revisjoner per sykehus i parentes .....	20
Tabell 4.2	Antall bidiagnoser før og etter revisjon .....	21
Tabell 4.3	Antall prosedyrekoder før og etter revisjon .....	22
Tabell 4.4	Oversikt over antall opphold som fikk ny HDG-plassering etter gruppering av revidert datamateriale .....	23
Tabell 4.5	Antall reviderte DRG-plasseringer per sykehus .....	24
Tabell 4.6	Antall korrigererte og reviderte DRG-poeng samt prosentvis differanse per sykehus .....	25
Tabell 4.7	Antall endringer med betydning for hoveddiagnose etter innleggelsesmåte og behandlingstype. Andeler i parentes .....	26
Tabell 4.8	Antall endringer med betydning for DRG etter innleggelsesmåte og behandlingstype. Andeler i parentes .....	27



# 1 Innledning

Etter at den tradisjonelle rammefinansieringen ble supplert med en aktivitetsbasert del-finansiering, som tok utgangspunkt i gjennomsnittskostnadene ved behandling av ulike typer pasienter, har det stadig vært fokusert på uheldige sider ved det nye finansieringssystemet (e.g. Sosial- og helsedepartementet, 1996). Fra enkeltinstitusjonenes side har det dessuten vært gitt uttrykk for engstelse knyttet til at ens eget sykehus ville komme spesielt dårlig ut ved innføringen av denne ordningen. I media har det i tillegg den siste tiden vært fokusert på at sykehus synes å ha kodet diagnoser med den hensikt å øke egne refusjoner. Dette prosjektets målsetning har vært å gi et situasjonsbilde av kvaliteten på de innrapporterte dataene i 2001. Formålet har derfor vært å:

- Sammenligne data innrapportert til Norsk Pasientregister (NPR) om hoveddiagnoser, bidiagnoser<sup>1</sup> og prosedyrekoder med de data journalgjennomgangen ga grunnlag for å sette
- Beregne hvor omfattende eventuelle avvik er og hvilke konsekvenser revisjonen får for diagnose relatert gruppe (DRG)-gruppering og DRG-poeng

Den overordnede problemstillingen for studien har vært "Er det avvik mellom hoveddiagnoser, bidiagnoser og operasjonskoder som opprinnelig ble innrapportert til NPR og dataene journalgjennomgangen fant grunnlag for å rapportere?"

En av innvendingene som har vært rettet mot DRG-systemet har vært at det ikke er spesielt godt egnet for å klassifisere medisinske opphold. Vi vil derfor ta utgangspunkt i en hypotese om at eventuelle avvik vil være størst blant opphold innenfor medisinske DRG-er. Videre forventer vi at akuttinnleggelser generelt vil være vanskeligere å kode korrekt enn elektive opphold, og vi har derfor formulert en hypotese om at eventuelle avvik vil forekomme hyppigere blant akuttinnleggelser enn blant planlagte opphold.

Rapporten er strukturert på følgende måte: Etter innledningen gis en oversikt over bakgrunnen for denne studien med spesiell fokus på den nåværende finansieringsordningen og tidligere studier. I kapittel 3 gjør vi rede for logikken bak trekkingen av utvalget samt metoden for datainnsamling som er benyttet. Kapittel fire presenterer resultatene fra journalgjennomgangen, mens vi i kapittel fem analyserer disse resultatene og peker på spesielle tendenser i datamaterialet. I kapittel 6 oppsummeres hovedfunnene fra undersøkelsen.

---

<sup>1</sup> Betegnelsen *hoveddiagnose* i DRG-systemet kan misforstås. DRG-systemet forholder seg bare til *diagnosekoder*, og bruker betegnelsen om den International Classification of Diseases (ICD)-10-koden som er oppført først i registreringen for det aktuelle kasus. Norsk ICD-10 opererer ikke med *diagnoser*, bare med *helsetilstander* og *beslektede helseproblemer*. Dette er en erkjennelse av at kodeverket jo ikke inneholder diagnoser, bare *diagnosekategorier* som for å sortere diagnostisk informasjon til kategorier i virksomhetsregistrering og statistikk. Enda mer forvirrende blir det fordi ICD-10 har en rekke regler som krever eller tillater at det mange ganger anvendes to eller flere koder for å kategorisere én tilstand. Da blir *hovedkoden* i ICD-10 *hovedtilstand* DRG *hoveddiagnose* og så vel *tilleggskode* for *hovedtilstand* som alle koder for det ICD-10 kaller *andre tilstander*, blir til *bidiagnoser* i DRG-systemet. Betegnelsene *hoveddiagnose* og *bidiagnose* i denne rapporten gjelder således *diagnosekoder*, ikke diagnoser som sådan. Det som revideres er ikke diagnoser, men hvordan diagnoser er representert og avrapportert med koder.



## 2 Bakgrunn

I dette kapitlet skal vi i den første delen kort presentere hovedfunnene fra noen tidligere studier og journalgjennomganger som har vært gjennomført, mens vi i del to gir en oversikt over implementeringen og organiseringen av det nåværende finansieringssystemet, og kritikken som har vært rettet mot det.

### 2.1 Tidligere studier

KITH gjennomførte i 1999 en journalgjennomgang av 336 pasientjournaler ved tre norske somatiske sykehus. Materialet var utplukket etter DRG, slik at kun opphold for pasienter med hjerte- og lungesykdommer samt pasienter som fikk utført keisersnitt ble inkludert i undersøkelsen. Gjennomgangen viste at 18 prosent av oppholdene var kommet i feil DRG.

Videre viste resultatene at nesten halvparten av de oppholdene som var feilklassifiserte skulle vært i en DRG hvor refusjonsbeløpet avvek med mer enn 50 prosent. Ved korrekt koding ville samtlige av de tre sykehusene mottatt større refusjoner. En tilsvarende studie fra 2000-2001 hvor KITH gjennomgikk 466 pasientjournaler fra opphold avviklet i 2000, viste også at korrekt koding ville gitt en samlet gevinst for sykehusene.

KITHs studie fra 1999, møtte en viss skepsis fra Sosial- og helsedepartementet på grunn av særegenheter ved utvalget. For det første påpekte departementet at studien var fra 1999, samme år som ICD-10 ble tatt i bruk. På bakgrunn av dette ble det antatt at risikoen for feilkoding ville være større for dette året enn andre år. For det andre kommenterte Sosial- og helsedepartementet at valg av sykdomsområder var uheldig, ettersom både kronisk obstruktiv lungesykdom (KOLS) og hjerteinfarkt hadde vært i søkelyset i Avregningsutvalget tidligere grunnet antakelser om utbredte unøyaktigheter ved kodingen. I vår studie har vi derfor søkt å ikke konsentrere oss om bestemte sykdomsgrupper, men istedenfor trekke opphold fritt blant alle opphold i 2001. Forøvrig benyttes diagnose- og prosedyrekodeverket ICD-10 fortsatt i 2001, slik at tilvenning til nytt kodeverk ikke vil kunne antas å prege resultatene i vår studie.

Flere internasjonale studier har også evaluert datakvaliteten på innrapporterte hoveddiagnoser (for eksempel Fisher et al, 1992; Hsia et al, 1988; Corn, 1981; Doremus & Michenzi, 1983; Johnson & Appel, 1984; Barnard & Esmond, 1981; Nilsson et al, 1992; Nilsson et al, 1994; Steinum, 1995). Ettersom kontekstuelle faktorer som tidspunktet for undersøkelsen, det geografiske området undersøkelsen har funnet sted i, utvalgsseleksjonen, variablene som har blitt undersøkt samt analysemetodene som har vært brukt har vært svært forskjellige, vil vi imidlertid ikke gå nærmere inn på innholdet i disse arbeidene.

### 2.2 Finansiering

Finansieringen av sykehusdriften i Norge har historisk sett vært gjennomført etter flere ulike prinsipper. Fram til 1980 fikk sykehusene tilskudd i henhold til en kurpengeordning, der grunnlaget for utbetalingene var antallet pasientdøgn sykehusene hadde. I 1980 ble denne ordningen erstattet av et rammefinansieringssystem der fylkeskommunene ble tildelt en ramme som var beregnet ut i fra et sett av demografiske kriterier. Fra 1. juli 1997 ble innsatsstyrt finansiering (ISF) innført som oppgjørsform for sykehusbehandling. Ramme-

finansieringen ble dermed supplert av et aktivitetsbasert system bygget på klassifisering av inneliggende pasienter etter DRG-systemet. 79 prosent av de somatiske sykehusene ble omfattet av det nye finansieringssystemet fra 1. juli 1997, mens de resterende institusjonene ble integrert i løpet av den påfølgende fireårs-perioden (Kjekshus, 2002). I 1997 ble derfor fortsatt 70 prosent av budsjettet tildelt gjennom rammefinansiering, mens kun 30 prosent var aktivitetsbasert. Siden har ISF-andelen gradvis økt til den i 2003 er 60 prosent.

## 2.3 Utviklingen av DRG-systemet og kritikken rettet mot det

DRG-systemet er bygget på en logikk der alle opphold ved somatiske sykehus grupperes til ressursmessig mest mulig homogene og medisinsk sett meningsfulle grupper. I 1996 utarbeidet Sosial- og helsedepartementet et skriv som gjorde rede for det overordnede utgangspunktet for bruk av ISF ved somatiske sykehus, samtidig som kritikken mot dette klassifiserings- og finansieringssystemet ble diskutert (Sosial- og helsedepartementet, 1996). Ettersom disse aspektene er av stor betydning for denne rapportens tema og problemstilling, vil vi kort gjengi noen av hovedpunktene.

### 2.3.1 Bakgrunnen for gruppering i DRG-system

Prinsippet med inndeling av grupper i henhold til et prissystem har sin opprinnelse innenfor industrien. Professor Robert B. Fetter ved Yale University og professor og kirurg John D. Thompson ble i 1967 bedt om å tilpasse dette systemet til behovet for problemløsning relatert til kvalitet, effektivitet og kostnader innenfor helsesektoren. Arbeidsgruppen som utarbeidet tilpasningen kom fram til en liste med krav som et slikt system måtte tilfredsstille. For det første måtte pasientene innenfor hver gruppe gjenspeile et ensartet ressursbehov. For det andre skulle inndelingen medisinsk sett være meningsfull, slik at hver gruppering klinisk inneholdt ensartede pasientgrupper. Dernest ble det vektlagt at klassifiseringen skulle bygge på eksisterende journalopplysninger, og for det fjerde at systemet skulle omfatte et overskuelig antall grupper og et tilstrekkelig antall pasienter i hver av disse. Videre skulle grupperingen være gjensidig utelukkende, slik at enhver pasient bare skulle være mulig å plassere i én gruppe. Det sjette kravet var at grupperingen måtte ivareta behovet for innbyrdes sammenligninger.

I den første utgaven av DRG-systemet for helsesektoren ble disse seks kravene søkt ivare tatt gjennom en inndeling i 383 grupper spesifiserte i henhold til pasientens hoveddiagnose, bidiagnoser og alder. Denne inndelingen var basert på World Health Organisations (WHO) sykdomsklassifisering, ICD-8.

Ved innføringen av ISF i Norge i 1997 ble sykehusoppholdene DRG-gruppert basert på diagnoser fra ICD-9 og kirurgiske prosedyrer fra The NOMESCO Classification of Surgical Procedures (NCSP). Sykdomsklassifikasjonen ICD-9 ble benyttet som basis for grupperingen til og med 1998. I 1999 ble den foreløpige siste revisjonen, ICD-10, innført. Det er diagnoser fra denne klassifikasjonen som sammen med registrerte kirurgiske inngrep (NCSP) brukes til å gruppere sykehusoppholdene.

Også i denne studien har ICD-10 og NCSP vært benyttet ved journalgjennomgangen. Dersom grupperingslogikken skal fremstilles noe forenklet, så bestemmes DRG-plasseringen av et sykehusopphold i utgangspunktet av hoveddiagnosen (hovedregelen). Unntaket fra denne regelen er opphold med svært ressurskrevende prosedyrer, der grupperingsforløpet er noe annerledes. Dersom for eksempel prosedyrer som levertransplantasjon, lungetransplantasjon eller benmargstransplantasjon har blitt utført, grupperes oppholdet direkte til spesielle DRG-er. Det samme er tilfelle for opphold som har hoveddiagnose multitraume eller hoved- og/eller bidiagnose HIV.

For majoriteten av oppholdene, som ikke faller innenfor unntakene nevnt ovenfor, skjer grupperingen ved at avdelingsoppholdene DRG-grupperes. Dette plasserer avdelingsoppholdet i en av de 25 hoveddiagnosegruppene (HDG).

Tabell 2.1 HDG for pasienter behandlet ved norske somatiske sykehus i 2001 (Sosial- og helsedepartementet, 2001)

HDG	Tekst
0	Tilleggskategori
1	Sykdommer i nervesystemet
2	Øyesykdommer
3	Øre-, nese- og halssykdommer
4	Sykdommer i åndedretsorganene
5	Sykdommer i sirkulasjonsorganene
6	Sykdommer i fordøyelsesorganene
7	Sykdommer i lever, galleveier og bukspyttkjertel
8	Sykdommer i muskel-skjelettsystemet og bindevev
9	Sykdommer i hud, underhud og bryst
10	Indresekretoriske, ernærings- og stoffskiftesykdommer
11	Nyre- og urinveissykdommer
12	Sykdommer i mannlige kjønnsorganer
13	Sykdommer i kvinnelige kjønnsorganer
14	Sykdommer under svangerskap, fødsel og barseltid
15	Nyfødte med tilstander som har oppstått i perinatalperioden
16	Sykdommer i blod, bloddannende organer og immunapparatet
17	Myeloproliferative sykdommer og lite differensierte svulster
18	Infeksiøse og parasittære sykdommer
19	Psykiske lidelser og atferdsforstyrrelser
20	Alkohol-, stoff- og medikamentmisbruk og organiske sinnslidelser framkalt av disse
21	Skade, forgiftninger og toksiske effekter av medikamenter/andre stoffer medikamentmisbruk og organiske sinnslidelser framkalt av disse
22	Forbrenninger
23	Faktorer som påvirker helsetilstanden – andre kontakter med helsetjenesten
24	Signifikant multitraume
25	HIV-infeksjon

Deretter blir avdelingsoppholdene aggregert til sykehusopphold, som igjen grupperes videre til ulike DRG-er basert på opplysninger om bidiagnose(r), kirurgiske prosedyrer, pasientens alder og kjønn samt utskrivingsstatus (som levende eller død). Det er denne DRG-grupperingen som sammen med visse prisregler bestemmer hvilken ISF-refusjon det enkelte sykehusoppholdene skal utløse, siden alle DRG-er er tilordnet en gitt kostnadsvekt som benyttes til å beregne ISF-refusjonen.

I Norge har det til nå blitt beregnet norske kostnadsvekter til DRG-systemet ved fire anledninger. Første gang var i 1990 da HCFA-3 versjonen av DRG-systemet ble utarbeidet av SINTEF NIS. I 1992 ble kostnadsvekter til HCFA-8 versjonen av DRG-systemet laget- denne gangen i regi av Sosial- og helsedepartementet. Den tredje gangen arbeidet ble utført, var det til HCFA-12 versjonen av DRG-systemet i 1999. Utarbeidelsen ble gjort av SINTEF Uni-

med. Den fjerde, og foreløpig siste, revisjonen av kostnadsvektene ble utført i 2001- også denne gang av SINTEF Unimed. Ved den fjerde revisjonen ble det beregnet kostnadsvekter til NorskDRG-versjonen av DRG-systemet.<sup>2</sup> Disse kostnadsvektene ble tatt i bruk fra og med 2002, og erstattet dermed kostnadsvektene til HCFA-12 versjonen (Pedersen & Solstad, 2002). En femte revisjon av kostnadsvektene til NorskDRG-versjonen i regi av SINTEF Unimed vil finne sted i 2003, og disse nye kostnadsvektene er planlagt implementert i 2004.

### 2.3.2 Kritikk av DRG-systemet

Et argument som ofte brukes mot anvendelsen av et DRG-system innenfor helsesektoren er at det ikke er i stand til å fange opp alvorlighetsgraden i pasientens tilstand. Hoved- og bidiagnose skal etter intensjonen gi et bilde av dette, men studier som har blitt gjort (for eksempel Horn et al, 1986) viser at indikatorene ikke er presise nok for dette formålet. Videre har DRG-systemet blitt kritisert for at det ikke har inneholdt noe effektmål på pasientbehandlingen, og for at det har eksistert for dårlige rutiner i forhold til oppfølging og vedlikehold av systemet. Sistnevnte innvending har særlig vært rettet mot manglende tilpasning av systemet (blant annet kodeverk og kostnadsvekter) til den medisinske og teknologiske utviklingen. I de senere årene har det derfor blitt lagt ned en betydelig innsats for å forbedre kodeverket og beregne nye kostnadsvekter til NorskDRG. Det sistnevnte arbeidet har blant annet blitt gjennomført ved SINTEF Unimed. Kostnadsvektene oppgraderes nå jevnlig med utgangspunkt i gjennomsnittskostnadene for behandling av pasienter innenfor de ulike pasientgruppene ved et utvalg av norske somatiske sykehus.

Det har også blitt reist kritikk mot mulighetene for misbruk av DRG-systemet i en aktivitetsbasert sammenheng. Innvendingene kan deles inn i tre hovedgrupper:

- For det første kan DRG-systemet bidra til at sykehusene prioriterer å behandle de til enhver tid mest økonomisk lønnsomme pasientene. Dette fenomenet omtales gjerne som *fløteskumming*. Som en konsekvens av dette kan det være at *terskelen for behandling* settes lavere for de populære og mest lønnsomme DRG-ene enn de faglige vurderingene skulle tilsi.
- En annen form for pervertering av systemet omtales som *patient-shifting*. Dette er en samlebetegnelse på en utviklingstendens der sykehusene går bort fra å behandle pasienter med faktiske kostnader over gjennomsnittet. På denne måten vil pasientene som gis en behandling koste mindre enn den refusjonen som sykehuset mottar for oppholdet.
- Den tredje måten DRG-systemet kan misbrukes på er gjennom det som kalles *DRG-kryp*. I studien gjennomført av Hsia et al (1988) refereres tre former for slik virksomhet. Først nevnes mis-specification som innebærer at feil diagnose har blitt valgt som hoveddiagnose, eller at diagnosen som har blitt satt ikke understøttes av journalopplysningene. En annen måte å øke inntektene på uten å øke aktiviteten, er ved miscoding. Ved miscoding fører sykehuset opp sykdommer eller prosedyrer som ikke har vært utført under oppholdet, bryter kodingsforskriftene eller setter unødvendig vag diagnosekode. Den tredje formen for kryp er resequencing, hvor rekkefølgen på diagnosene endres, for eksempel ved å føre opp en av bidiagnosene som hoveddiagnose, med den effekt at refusjonsstørrelsen øker.

I kapitlene 4 og 5 vil vi komme nærmere tilbake til omfanget av feilkodinger på hoveddiagnose- og bidiagnosenivå, samt i hvor stor utstrekning dette påvirker DRG-plasseringen.

---

<sup>2</sup> Nord DRG er grunnlag for den norske versjonen, NorskDRG. Endringer i NordDRG videreføres som oftest i NorskDRG.

Først skal vi imidlertid presentere utvalget som har blitt undersøkt, metoden for data-innsamling som har blitt benyttet og diskutere de metodiske utfordringene ved studien.



## 3 Metode

I dette kapitlet skal vi redegjøre for sentrale aspekter ved forskningsdesignet. Vi vil hovedsakelig fokusere på framgangsmåten ved trekkingen av utvalget, metodene for datainnsamling og dataanalyse som ble benyttet samt de metodiske utfordringene studien har bydd på.

### 3.1 Utvalg

Den totale populasjonen som vårt utvalg til studien skulle trekkes fra besto av 799 309 avdelingsopphold ved 73 somatiske sykehus i 2001.<sup>3</sup> Ved trekkingen av utvalget benyttet vi datafiler fra NPR, og utvalget ble trukket i en totrinns-prosess der vi først trakk ut institusjonene og deretter selekterte enkelte avdelingsopphold ved disse institusjonene. Forut for uttrekningen av sykehus stratifiserte vi populasjonen i to kategorier, avhengig av hvilken sykehustype institusjonen tilhørte. Den første kategorien besto av 20 store sykehus, og inkluderte samtlige universitetsklinikker og sentralsykehus. Den andre kategorien inkluderte de resterende 53 sykehusene, og omfattet dermed lokal- og fylkessykehusene. To sykehus ble trukket tilfeldig fra gruppen av store sykehus og tre ble tilfeldig utvalgt fra gruppen av små sykehus. Sykehusene som refereres til som A og B ble trukket fra gruppen med store sykehus, mens sykehusene C, D og E ble trukket fra den andre kategorien.

I tabell 2 har vi satt opp hvor stor prosentandel av oppholdene ved de fem utvalgte sykehusene som var gruppert til den kompliserte DRG-en i et DRG-par, og hvor mange bidiagnoser som gjennomsnittlig er registrert per opphold. Begge disse målene kan si oss noe om kodefrekvensen ved de utvalgte sykehusene, og hvorvidt disse sykehusene skiller seg vesentlig fra gjennomsnittet i den sykehusgruppen de representerer.

---

<sup>3</sup> Kjemoterapi, dialyse, rehabilitering og fleravdelingsopphold er ekskludert. De to førstnevnte behandlingstypene ble holdt utenfor fordi de omfattet et svært høyt antall opphold der kodingen bør være relativt ukomplisert. Rehabiliteringsopphold ble ekskludert pga. de spesielle refusjonsreglene som ikke egner seg så godt for denne typen koderevisjon. Fleravdelingsoppholdene ble ikke inkludert ettersom undersøkelsen ikke tok sikte på å avklare virkningen av aggregering fra avdelingsopphold til sykehusopphold

Tabell 3.1 Prosentandel kompliserte DRG-er og gjennomsnittlig antall bidiagnoser ved de utvalgte sykehusene og i totalpopulasjonen

	Prosentandel kompliserte DRG-er	Gjennomsnittlig antall bidiagnoser
Sykehus A	37,3	1,26
Sykehus B	37,5	1,33
Sykehus C	26,0	1,08
Sykehus D	30,3	1,24
Sykehus E	37,5	1,28
Store sykehus	33,4	1,26
Små sykehus	29,0	1,08
Alle sykehus	32,0	1,21

Både sykehus A og B har en andel kompliserte opphold på rett over 37 prosent, en andel som ligger noe over det samlede gjennomsnittet for de store sykehusene, som er 33,4 prosent. Gjennomsnittlig antall opphold i kompliserte DRG-er er 29 prosent for de små sykehusene. Av tabellen ser vi at sykehus C ligger under dette gjennomsnittet, mens de to øvrige små sykehusene ligger over. Sykehus E har en prosentandel kompliserte DRG-er som er like høy som de to store sykehusene, og skiller seg dermed en del fra gjennomsnittet i gruppen med små sykehus.

Den andre kolonnen i tabell 2 viser det gjennomsnittlige antallet registrerte bidiagnoser ved avdelingsoppholdene ved de fem sykehusene, beregnet ut fra samtlige opphold. Både sykehus A og C har samme gjennomsnittlige antall bidiagnoser per opphold som den sykehuskategorien de representerer, mens de tre øvrige sykehusene har et noe høyere gjennomsnitt. Igjen er det sykehus E, som også hadde en høy andel opphold i komplisert DRG, som skiller seg mest fra gjennomsnittet med 1,28 bidiagnoser per opphold, mens gjennomsnittet for de små sykehusene er 1,08.

I det neste trinnet av utvelgelsesprosessen plukket vi tilfeldig ut 110 avdelingsopphold ved hvert av de fem sykehusene, uten annen seleksjon av avdelingsoppholdene enn nevnt i fotnote på forrige side, verken i henhold til DRG, hoveddiagnoser, bidiagnoser, pasientsammensetning eller andre pasientspesifikke eller administrative kjennetegn. Totalt ble det trukket 550 avdelingsopphold, hvorav 500 skulle undersøkes nærmere i journalgjennomgangen ved hvert enkelt sykehus. De 50 resterende avdelingsoppholdene utgjorde en reserve i tilfelle enkeltjournaler av ulike årsaker ikke kunne skaffes til veie og framlegges for gjennomgang.

## 3.2 Metode for datainnsamling

Datainnsamlingen ble organisert slik at to leger fra KITH gikk gjennom 500 av de selekterte journalene ute ved de enkelte sykehusene. Ved denne gjennomgangen ble det kontrollert om de diagnose- og prosedyrekoder som var registrert i det pasientadministrative systemet stemte overens med data innrapportert til NPR. Videre ble det undersøkt hvorvidt de koder som stod oppført på journalenes oversiktsark stemte overens med registreringene i det pasientadministrative systemet. Den mest sentrale delen av journalgjennomgangen var imidlertid sammenholdingen av selve journaldokumentasjonen med de diagnose- og prosedyrekoder som var registrert på diagnosearket. I forbindelse med sistnevnte del, ble et revidert kodesett registrert i en database. I denne databasen ble hoveddiagnoser, bidiagnoser og prosedyrekoder som ikke skulle vært registrert eller som ikke hadde dekning i journalopplysninger strøket, mens manglende koder ble tilføyet og feilaktige koder korrigert.

For å finne ut om de revideringene som ble foretatt på bakgrunn av journalgjennomgangen medførte endringer av oppholdenes DRG-plassering, ble alle de undersøkte oppholdene gruppert på nytt ved hjelp av grupperingsprogrammet NorskDRG versjon 1.7. Denne re-grupperingen av de undersøkte avdelingsoppholdene gjorde det mulig å analysere avvikene mellom opprinnelig og ny DRG-plassering.

I de tilfeller hvor det ble avdekket feil av praktisk betydning mellom det pasientadministrative systemet og NPR-data, eller mellom journalomslag og PAS-system, ble dette registrert. Disse avvikene var ikke sentrale under det videre arbeidet med data og i forbindelse med selve dataanalysen. Forøvrig ble kun avvik mellom journalenes innhold og NPR-data registrert. Forhold ved kodingen som kun gjaldt årsaks-, steds- og aktivitetskoder for skadetilfeller, ble ikke vektlagt ved gjennomgangen. Manglende årsakskoding ble imidlertid notert, men ikke kodet eller omkodet ved revisjonen av kasusene. Manglende koder som det pasientadministrative systemet og NPR ikke kan håndtere, for eksempel ATC-koder, ble heller ikke notert. Det samme gjaldt manglende bruk av frivillig tilleggs-koding uten betydning for DRG-gruppering.

### 3.3 Metodiske utfordringer

Utvalget som er benyttet i undersøkelsen er trukket ved sannsynlighetsutvelgning. Dette reduserer sannsynligheten for at enhetene i utvalget skiller seg systematisk fra populasjonen. Eventuelle avvik mellom universet og utvalget vil altså dermed være et resultat av tilfeldige variasjoner. Som i alle andre utvalgsundersøkelser er resultatene fra denne undersøkelse beheftet med usikkerhet. Den usikkerheten man får i resultatene fordi man bygger på opplysninger om en del av de avdelingsoppholdene som undersøkelsen dekker, kalles utvalgsvarians. Standardavviket er et mål på denne usikkerheten. Størrelsen på standardavviket kan anslås ved hjelp av antallet observasjoner i utvalget. Tabell 3 viser størrelsen på standardavviket for observerte prosentandeler ved ulike utvalgstørrelser.

Tabell 3.2 Forventet standardavvik for observerte prosentandeler ved ulike utvalgstørrelser

Antall obs (n)	Prosentandeler (p)					
	95/5	90/10	80/20	70/30	60/40	50/50
25	4,36	6,00	8,00	9,17	9,80	10,00
50	3,08	4,24	5,66	6,48	6,93	7,07
100	2,18	3,00	4,00	4,58	4,90	5,00
250	1,38	1,90	2,53	2,90	3,10	3,16
400	1,09	1,50	2,00	2,29	2,45	2,50
500	0,97	1,34	1,79	2,05	2,19	2,24
750	0,80	1,10	1,46	1,67	1,79	1,83
1000	0,69	0,95	1,26	1,45	1,55	1,58
1500	0,56	0,77	1,03	1,18	1,26	1,29
2000	0,49	0,67	0,89	1,02	1,10	1,12

Følgende eksempel viser hvordan man kan bruke tabell 3 for å finne konfidensintervaller: Anslaget på standardavviket til et observert prosenttall på 80 (f.eks 80 prosent ja og 20 prosent nei) er 1,79 når antall observasjoner er 500. Konfidensintervallet for den sanne verdi får grensene  $80 \pm 2 \times 1,79$ , dvs at det strekker seg fra 76,4 til 83,6. Det betyr at dette intervallet med 95 prosent sikkerhet vil inneholde den tallstørrelsen en ville ha fått om alle enheter i populasjonene hadde vært med i undersøkelsen. Som det også fremgår av tabellen

øker størrelsen på standardavviket når antallet observasjoner synker, og når prosentfordelingen er jevn. Det høyeste forventede standardavviket befinner seg derfor i et utvalg med få observasjoner, der prosentfordelingen er 50/50.

Til tross for at vi i utgangspunktet ikke har grunn til å forvente noen systematiske avvik mellom populasjonen og utvalget, så består utvalget av kun om lag 0,07 prosent av den totale utvalgspopulasjonen<sup>4</sup>. Ettersom vi kun undersøkte journaler ved fem sykehus, var det naturligvis av stor betydning at de institusjonene som ble utvalgt ikke avvek systematisk fra de øvrige sykehusene i utvalget. Vi vurderte derfor å kontrollere for også andre egenskaper ved sykehusene og/eller de enkelte avdelingsoppholdene, men ettersom det også er stor usikkerhet knyttet til hvilke dimensjoner det er mest hensiktsmessige å stratifisere i forhold til, valgte vi til slutt å ikke kontrollere for ytterligere forhold.

En annen utfordring i forbindelse med gjennomføringen av studien var knyttet til faren for lav interreliabilitet. Dette kan påvirke resultatene fordi tolkningen av opplysningene gitt i pasientjournaler ofte varierer fra lege til lege. Vi forsøkte å korrigere for slike problemer ved at to leger samarbeidet om å gjennomgå de enkelte journalene. Denne prosessen åpnet dermed for at legene hadde mulighet til å overprøve hverandres konklusjoner, og diskutere seg imellom dersom de var uenige. I tillegg hadde det vært ønskelig å inkludere et større antall leger i arbeidet med journalgjennomgangen. Dette kunne for eksempel vært organisert som en for-undersøkelse der grad av konsensus om koding innad i en større gruppe av leger hadde blitt undersøkt. Innenfor rammene for dette prosjektet, var det imidlertid ikke mulig å gjennomføre en slik studie. Øvrige utfordringen knyttet til datainnsamling gjennom journalgjennomgang kommer vi nærmere tilbake til i kapittel 5.

### 3.4 Reliabilitet og validitet

Dataenes validitet og reliabilitet er forenklet forklart et mål på hvorvidt det innsamlede materialet gir et riktig bilde av de egenskapene vi ønsker å undersøke. Reliabiliteten bestemmes av hvordan datainnsamlingen er gjennomført, og betegnelsen sikter til nøyaktigheten i denne prosessen. Validiteten avhenger av at datas reliabilitet er høy, og fungerer som en indikator på hvorvidt vi har undersøkt eller målt det vi intenderte å måle. Validiteten betegner altså dataenes relevans for den problemstillingen som skal undersøkes.

Siden denne undersøkelsen analyserer de samme dataene som vi på forhånd hadde tilgjengelig fra NPR, gir det seg selv at dataene utvilsomt er relevante i forhold til problemstillingen. Videre er denne studien en undersøkelse som faktisk tar sikte på å vurdere reliabiliteten til de dataene som alt befinner seg i det norske pasientregisteret. Kan man stole på at de opplysningene som fins i registeret stemmer med de medisinske opplysningene som er registrert i pasientenes journaler ute på sykehusene?

Videre kan vi ikke helt utelukke at forskjeller i pasienttilbudet ved de fem sykehusene som deltar i denne undersøkelsen, kan påvirke reliabiliteten. Blant annet har ikke sykehus E noen fødeavdeling, noe de øvrige sykehusene har. Dersom det forekommer et større omfang av medisinske feilkodinger innenfor visse medisinske fagområder, kan resultatene slik påvirkes av at de fem sykehusene har et ulikt pasienttilbud. De to store sykehusene, A og B, har et mer spesialisert og bredere tilbud enn de tre små sykehusene, C, D og E, og det kan også være forskjeller mellom sykehusene innenfor de to kategoriene. Tabell 2 viste tidligere at det eksisterer variasjoner mellom sykehusene både med hensyn til prosentandelen opphold i kompliserte DRG-er og gjennomsnittlig antall registrerte bidiagnoser per opphold.

---

<sup>4</sup> Dersom man antar at prosentfordelingen er 50/50 i datamaterialet, vil en undersøkelse bestående av 1 500 opphold ha et forventet standardavvik på 1,29 prosent. Det tilsvarende standardavviket i et utvalg på 2000 opphold vil være 1,12 prosent.

## 4 Resultater

I dette kapitlet vil vi presentere de viktigste funnene fra den gjennomførte journalgjennomgangen, med spesiell vekt på revisjon av hoveddiagnoser, bidiagnoser og prosedyrekoder, samt avvikene journalgjennomgangen ga mellom opprinnelig og revidert DRG-gruppering. Vi vil også drøfte hvilke konsekvenser den relativt høye diskrepansen mellom opprinnelig og revidert koding har for beregningen av DRG-poeng og for den påfølgende ISF-refusjonen som i 2001 ble utbetalt til institusjonenes hjemmehørende fylkeskommuner. Vi har dessuten undersøkt om de observerte avvikene samvarierer med innleggelsesmåte og behandlingstype.

### 4.1 Hoveddiagnoser

I forbindelse med journalgjennomgangen ble en del av de opprinnelige hoveddiagnosene revidert. Revisjonene kan deles inn i følgende fire hovedgrupper:

- Feil kun i kodens fjerde posisjon
- Kode i feil tretetnsgruppe
- Kode i feil kategoriblokk
- Kode i feil kapittel

Tabell 4 gir en totaloversikt over antall revisjoner av hoveddiagnoser og en frekvensfordeling over utbredelsen av disse i henhold til de fire hovedgruppene, fordelt per sykehus.

Tabell 4.1 Antall reviderte hoveddiagnoser per sykehus og fordeling av disse i henhold til feil-type. Kategoriens andel feil av totalt antall revisjoner per sykehus i parentes<sup>5</sup>

Sykehus	Kapittel	Kat. blokk	3. tegn	4. tegn	Totalt
A	9 (28)	5 (16)	10 (31)	8 (25)	32
B	13 (39)	13 (39)	2 (6)	5 (15)	33
C	20 (41)	13 (27)	8 (16)	8 (16)	49
D	20 (41)	10 (20)	6 (12)	13 (27)	49
E	18 (40)	7 (16)	8 (18)	12 (27)	45
Totalt	80 (38)	48 (23)	34 (16)	46 (22)	208

Tabellen viser at hele 208 av de 500 undersøkte avdelingsoppholdene fikk revidert den opprinnelige hoveddiagnose ved journalgjennomgangen. Dette utgjør i overkant av 41 prosent av oppholdene. Mellom sykehusene varierer feilandelen fra 32 prosent av de undersøkte oppholdene for sykehus A til 49 prosent for sykehus C og D. Dermed fikk mer enn hvert tredje opphold ved samtlige fem sykehus revidert hoveddiagnosen, mens nesten halvparten av alle oppholdene ved sykehus C og D har fått revidert den opprinnelig satte hoveddiagnosen.

Den fire-delte kategoriseringen gir et grovt mål på *graden* av avvik i kodingen, og som tabellen viser, er avvik på kapitteinivå totalt sett den hyppigst forekommende formen for feil. De tre øvrige kategoriene fanger opp henholdsvis 48, 34 og 46 feil i hoveddiagnosene totalt. Mellom de fem sykehusene i utvalget varierer det imidlertid hvilken feiltipe som forekommer hyppigst. Sykehus A, som hadde det laveste antallet avvik, hadde flest revisjoner på tre-tegnnivå. For sykehus C og D var revisjon på kapitteinivå mest utbredt, mens sykehus B hadde 13 avvik både på kapittel- og kategori-bloknivå.

Når det gjelder kapitelfeil, fikk totalt 22 av de 80 oppholdene i denne gruppen erstattet den opprinnelige hoveddiagnosen med en av de tilstandene som alt var registrert som en bidiagnose. For 16 av disse 22 oppholdene var det den første bidiagnosen som ble satt som ny hoveddiagnose, mens det for de seks øvrige oppholdene enten var bidiagnose nummer to (fem opphold) eller tre (ett opphold) som ble valgt. Slik sett tilsvarer ikke alltid graden av systematiske feil graden av logiske feil. For eksempel vil denne typen revidering av rekkefølgen, der en bidiagnose erstatter den opprinnelige hoveddiagnosen, ofte bety at den nye hoveddiagnosen stammer fra et annet kapittel. Antallet opphold med revisjon av denne typen feil varierte fra fire opphold ved sykehus B, til sju ved sykehus E. Sykehus A hadde ingen feilkodinger av denne typen og samlet sett var det dermed ved de tre små sykehusene i utvalget at dette var mest utbredt.

Blant oppholdene som fikk endret den opprinnelige hoveddiagnosen, ble den opprinnelige diagnosen i flere tilfeller ført opp som bidiagnose ved revisjonen. Dette skjedde for over en fjerdedel av de 208 oppholdene. Ved sykehus C ble 16 av de 49 hoveddiagnosene som ble endret føyd til som bidiagnose. For sykehus B var det tilsvarende tallet 13 av 33, mens sykehus D fikk ført opp 11 av 49 opprinnelige hoveddiagnoser som bidiagnoser og sykehus E 11 av 45. Sykehus A, som forøvrig hadde det minste antallet endrede hoveddiagnoser i utgangspunktet, fikk kun fem av 32 hoveddiagnoser ført opp som bidiagnoser ved revisjonen.

Avvikene i hoveddiagnosene vil bli drøftet mer inngående i kapittel 5. Det er imidlertid verdt å merke seg at revisjon av hoveddiagnose ikke nødvendigvis medfører endret DRG. Ofte vil revisjon av hoveddiagnosen ikke ha betydning for DRG-plasseringen, for eksempel ved re-

<sup>5</sup> På grunn av desimalavrunding blir ikke summen av prosentandelene 100 for sykehus B og E.

visjon av fjerdesiffer. Hoveddiagnoser som revideres til et annet kapittel vil imidlertid som oftest medføre endret DRG gruppering, og i noen tilfeller også endret HDG.

## 4.2 Bidiagnoser

Feilregistrering av bidiagnoser har ofte blitt sett i sammenheng med DRG-kryp, ettersom bidiagnoser kan føre til at oppholdet havner i en komplisert DRG. Forutsetningene for å sette slik aktivitet i sammenheng med DRG-kryp er enten at det tidligere har vært underregistrering av bidiagnoser, eller at sykehusene for det første innrapporterer bidiagnoser som ICD-10 ikke gir grunnlag for, og for det andre, at disse bidiagnosene fører til at oppholdet havner i en DRG som gir en høyere DRG-vekt enn korrekt koding ville utløst.

Med utgangspunkt i dataene fra journalgjennomgangen har vi sett nærmere etter tegn på slik registreringspraksis. Totalt sett har sykehus A, D og E ført opp for mange bidiagnoser, mens de to øvrige har ført opp for få. Samlet sett hadde sykehusene registrert nesten fem prosent flere bidiagnoser før revisjon enn etter. Tabell 5 viser hvordan avvikene fordelte seg mellom institusjonene.

Tabell 4.2 Antall bidiagnoser før og etter revisjon<sup>6</sup>

Sykehus	Opprinnelig ant bidiagnoser	Ant bidiagnoser lagt til	Ant bidiagnoser fjernet	Ant bidiagnoser etter revisjon
A	111	14	30	95
B	106	25	21	110
C	105	30	25	110
D	129	26	42	113
E	156	18	40	134
Totalt	607	113	158	562

Som vi ser av tabellen var avvikene i antall bidiagnoser størst for de sykehusene som i utgangspunktet hadde registrert flest bidiagnoser, og som fikk en *nedgang* i antallet etter revisjonen. Endringer i det totale antallet bidiagnoser gir imidlertid ikke et uttømmende bilde av hva som skjer ved revisjonen. En nærmere analyse av dataene viser at en svært stor andel av bidiagnosene ble revidert ved samtlige fem sykehus. Dette bidrar til å forklare 38 revisjoner innenfor par-DRG-er. Dette var tilfelle for ni opphold ved sykehus A og B, sju opphold ved sykehus C og D og for seks opphold ved sykehus E. Blant disse 38 revisjonene var det imidlertid 12 opphold som også fikk endret hoveddiagnose. For disse oppholdene blir det dermed noe vanskelig å fastslå hvorvidt det var revisjon av bidiagnoser eller hoveddiagnosen som forårsaket DRG-endringen.

Det vi imidlertid med sikkerhet kan fastslå er at det var 32 av oppholdene i datasettet som endret DRG-gruppering som en direkte konsekvens av endringer i bidiagnose, uten at hoveddiagnosen ble revidert. Dette gjaldt sju opphold ved sykehus A, åtte ved sykehus B, fire ved sykehus C, seks ved sykehus D og sju ved sykehus E. En oversikt over hvilken betydning disse revisjonene hadde for opptjeningen av DRG-poeng, gis i delkapittel 4.6.

<sup>6</sup> Se tabell B-F i appendix for en mer detaljert oversikt over endringene.

### 4.3 Kirurgiske prosedyrekoder

I tillegg til hoved- og bidiagnoser, kodes også opphold i henhold til de kirurgiske prosedyrene (NCSP) som har blitt utført under oppholdet. Prosedyrekoder kan for eksempel inneholde viktig tilleggsinformasjon om operasjonsteknikker eller operasjonsvarigheten (Sosial- og helsedepartementet, 2001). De formelle kravene til koding av prosedyrer er for det første at samtlige fem tegn skal benyttes, og dernest at alle utførte prosedyrer skal registreres.

Nedenfor er en sykehusspesifikk oversikt over avvikene mellom opprinnelige og reviderte prosedyrekoder.

Tabell 4.3 Antall prosedyrekoder før og etter revisjon

Sykehus	Opprinnelig ant prosedyrekoder	Ant prosedyrekoder lagt til	Ant prosedyrekoder fjernet	Ant prosedyrekoder etter revisjon
A	53	7	0	60
B	161	1	7	155
C	70	8	1	77
D	56	11	1	66
E	41	11	0	52
Totalt	381	38	9	410

Som vi ser av tabellen er det kun sykehus B som totalt sett står oppført med flere prosedyrekoder før revisjonen enn etter. Forøvrig registrerer sykehusene generelt for få prosedyrekoder.

I ti tilfeller hadde revisjon av prosedyrekoder direkte innvirkning på endring av DRG-gruppering (hoveddiagnosen ble ikke revidert og eventuelle endringer av bidiagnose hadde ikke innvirkning på DRG-plasseringen). Sykehusene A og D hadde tre slike opphold hver, mens sykehus B og C hadde to hver. Hvilken betydning dette hadde omregnet til DRG-poeng, er kommentert i delkapittel 4.6.

### 4.4 Plassering i HDG

I kapittel 2 gjorde vi rede for hvordan de ulike DRG-ene er tilordnet til en av de 25 hoveddiagnosegruppene som DRG-systemet er oppbygd av. Hoveddiagnosegruppene følger i hovedsak spesialitet og/eller organsystem, og det synes rimelig å forvente at de fleste oppholdene som har blitt revidert til en ny DRG fortsatt vil befinne seg innenfor samme HDG. Dersom oppholdene havner i en ny HDG vil det imidlertid gi et bilde på *graden* av endring. En gjennomgang av dataene viste at 69 av oppholdene som ble revidert havnet innenfor en ny HDG på bakgrunn av de revisjonene som var gjort. For å avdekke om noen av de 25 gruppene utmerket seg spesielt, lagde vi en frekvensoversikt over revisjonen av DRG-er med betydning for HDG-plassering. Oversikten er gjengitt i tabell 7.

Tabell 4.4 Oversikt over antall opphold som fikk ny HDG-plassering etter gruppering av revidert datamateriale

Opprinnelig HDG	Totalt antall opphold før revisjon	Antall reviderte opphold
0 Tilleggskategori	2	2
1 Sykdommer i nervesystemet	39	11
2 Øyesykdommer	14	0
3 Øre-, nese- og halssykdommer	22	1
4 Sykdommer i åndedretsorganene	47	5
5 Sykdommer i sirkulasjonsorganene	80	10
6 Sykdommer i fordøyelsesorganene	61	2
7 Sykdommer i lever, galleveier og bukspyttkjertel	16	4
8 Sykdom i muskel- skjelettsystem og bindevev	59	3
9 Sykdommer i hud, underhud og bryst	20	8
10 Indreseretoriske ernærings- og stoffskiftesykdommer	10	3
11 Nyre- og urinveissykdommer	20	2
12 Sykdommer i mannlige kjønnsorganer	13	4
13 Sykdommer i kvinnelige kjønnsorganer	12	2
14 Sykdom under svangerskap, fødsel og barseltid	40	0
15 Nyfødte med tilstander som har oppstått i perinatalperioden	3	0
16 Sykdommer i blod, bloddannende organer og immunapparatet	3	0
17 Myeloproliferative sykdommer og lite differensierte svulster	8	4
18 Infeksiøse og parasittære sykdommer	10	0
19 Psykiske lidelser og adferdsforstyrrelser	3	1
20 Alkohol-, stoff- og medikamentmisbruk og organiske sinnslidelser framkalt av disse	3	1
21 Skade, forgiftninger og toksiske effekter av medikamenter/andre stoffer medikamentmisbruk og organiske sinnslidelser framkalt av disse	8	2
22 Forbrenninger	0	0
23 Faktorer som påvirker helsetilstanden – andre kontakter med helsetjenesten	7	4
24 Signifikant multitraume	0	0
25 HIV-infeksjon	0	0
Totalt antall opphold	500	69

Som vi ser av tabellen var antallet opphold som ble revidert til ny HDG spesielt høyt innenfor HDG 1, 5 og 9, med henholdsvis elleve, ti og åtte opphold. Blant oppholdene som endret plassering fra hoveddiagnosegruppen som omfatter sykdommer i nervesystemet, havnet fire

i gruppe 23.<sup>7</sup> Blant de åtte oppholdene som endret plassering fra HDG 9, ble tre omplasserte til HDG 21. Fra HDG 5 ble to opphold revidert til HDG 1 og to til HDG 8. Tolkningen av disse resultatene vil vi komme nærmere tilbake til i delkapittel 5.4.

## 4.5 Plassering i DRG

Som nevnt i kapittel 2.3.1, bestemmes grupperingen av oppholdene i DRG-er av hoveddiagnosen, eventuelle bidiagnoser og/eller kirurgiske inngrep som er relevante for oppholdet samt pasientens kjønn, alder og utskrivningsstatus.

Resultatene fra undersøkelsen viser at totalt 166 av 500 opphold, altså en tredjedel, ble revidert til ny DRG. Dette var et resultat av en av følgende faktorer:

- Revisjon av hoveddiagnosesetting
- Riktig hoveddiagnose, men revisjon av bidiagnose fører til forskyvning fra DRG med bidiagnoser/kompliserende sykdom (b/k) til uten b/k, eller omvendt
- Riktig hoveddiagnosesetting, men revisjon av prosedyrekode forårsaker endring av DRG, for eksempel medisinsk DRG istedenfor kirurgisk, eller omvendt

I 124 av de 166 tilfellene der DRG ble revidert, var også hoveddiagnosen endret.<sup>8</sup> I 42 tilfeller ble imidlertid oppholdene gruppert til en annen DRG uten at dette var årsaken. Dette skyldtes for 32 opphold revisjon av bidiagnosesetting og for ti opphold revisjon av kirurgiske prosedyrekoder. Tabell 8 viser hvordan revisjon med konsekvenser for DRG-grupperingen fordelte seg mellom de fem sykehusene.

Tabell 4.5 Antall reviderte DRG-plasseringer per sykehus<sup>9</sup>

Sykehus	Antall opphold med revidert DRG-plassering	Antall DRG-poeng	
		Før revisjon	Etter revisjon
A	31	104	91
B	26	135	125
C	33	86	91
D	40	97	102
E	36	100	94
Totalt	166	522	503

Sykehus B hadde 26 avdelingsopphold som ble gruppert til ny DRG etter revisjonen. Dette er den laveste andelen av slik revisjon blant de fem sykehusene. Sykehus D hadde hele 40 opphold som ble korrigerende til ny DRG, og er dermed den institusjonen som har den høyeste andelen av feil DRG-plassering blant de undersøkte institusjonene. Til tross for at revisjonsandelen varierer mellom 26 og 40 prosent, er disse resultatene oppsiktsvekkende høye sammenlignet med resultatene fra undersøkelsen KITH gjorde på oppdrag for Riks-

<sup>7</sup> For en mer detaljert beskrivelse av hvilke HDG-er oppholdene ble revidert til, se tabell A i appendix.

<sup>8</sup> I disse tilfellene kan også endringer av bidiagnoser og/eller prosedyrekoder ha hatt betydning for DRG-revisjonen.

<sup>9</sup> Se tabell B-F i appendix for detaljert oversikt over revisjonene gjort for opphold med revidert DRG.

trygdeverket. I sistnevnte undersøkelse varierte de tilsvarende andelene mellom 13 og 24 prosent, med et totalt avvik på 18 prosent.

Den relativt høye forekomsten av opphold med avvik i DRG kan i all hovedsak forklares av avvikene i hoveddiagnoser og bidiagnoser. Ettersom svært mange opphold fikk revidert hoveddiagnosen til en ny diagnose i et annet kapittel i ICD 10, bidrar dette i mange tilfeller til endret DRG-plassering. Vi vil komme nærmere tilbake til dette i kapittel 5.5.1.

## 4.6 Økonomiske konsekvenser

Så langt har vi sett at avvikene mellom den opprinnelige og reviderte DRG-plasseringen er relativt store. Hvilken økonomisk effekt omfanget av dette faktisk har, framgår av tabell 8 og 9 som viser endringene i korrigerede DRG-poeng. Sykehus A fikk 12,7 prosent færre DRG-poeng etter revisjon. For sykehus B og E er de tilsvarende andelene henholdsvis 7,3 og 5,8 prosent. Sykehusene C og D hadde henholdsvis 5,9 og 4,4 prosent økning i DRG-poeng etter revisjonen.

Tabell 4.6 Antall korrigerede og reviderte DRG-poeng samt prosentvis differanse per sykehus

Sykehus	Opprinnelig ant DRG-poeng	Ant DRG-poeng etter revisjon	Prosentvis endring
A	104,3	91,1	-12,7
B	134,8	125,0	-7,3
C	85,8	90,9	5,9
D	97,5	101,8	4,4
E	99,6	93,8	-5,8
Totalt	522,0	502,6	-3,7

De prosentvise avvikene i DRG-poeng framstilt i tabell 9 er altså et resultat av endringer i DRG-grupperingen. 124 av oppholdene som har fått endret DRG har også fått revidert hoveddiagnose og 38 opphold ble omkodet innad i et DRG-par. Målt i DRG-poeng medførte revisjonen innenfor par-DRG-ene en nedgang på 3,3 poeng for sykehus A, 2,10 poeng for sykehus B, 2,34 poeng for sykehus D, 2,35 poeng for sykehus E og en økning på 2,47 poeng for sykehus C.

Blant de 42 revisjonene der DRG ble endret uten at hoveddiagnosen ble revidert, var 32 av endringene forårsaket utelukkende av revisjon av bidiagnoser<sup>10</sup>. I DRG-poeng innebar disse endringene en nedgang etter revisjon på 4,62 poeng for sykehus A, 2,03 poeng for sykehus B, 1,25 poeng for sykehus D og en økning på 1,41 poeng for sykehus C og 1,28 poeng for sykehus E. Ti opphold endret DRG utelukkende på grunn av revisjon av prosedyrekoder. I DRG-poeng medførte disse revisjonene en nedgang på 0,37 poeng for sykehus A og 0,87 poeng for sykehus B, mens sykehus C og D fikk en økning på henholdsvis 0,62 og 1,16 poeng etter revisjon.

Totalt sett ser vi dermed at utbetalingene ville vært 3,7 prosent høyere før revisjonen enn etter. Det ser dermed ut til at det har vært en overkodning blant sykehusene. Variasjonene mellom sykehus A, som ville fått 12,7 prosent færre DRG-poeng dersom opplysningene fra journalgjennomgangen hadde blitt lagt til grunn, og sykehus C, som ville fått nesten seks prosent flere DRG-poeng, er imidlertid påfallende. Dette kan tyde på at det også på nasjonalt nivå vil forekomme relativt store variasjoner mellom enkeltinstitusjoner.

<sup>10</sup> Se tabell B-F i appendix for detaljert oversikt over revisjonene gjort for opphold med revidert DRG.

Det totale resultatet på 3,7 prosent avvik i DRG-poeng er imidlertid overraskende lavt i forhold til hva avvikene i hoved- og bidiagnoser ga grunnlag for å forvente. Dette tyder på at DRG-grupperingen utbetalingsmessig gir et relativt robust og godt tilpasset mål for beregning av utbetalinger.

## 4.7 Innleggelsesmåte og behandlingstype

Noe av kritikken som tradisjonelt har blitt rettet mot DRG-systemet er at det egner seg best til å klassifisere elektive kirurgiske innleggelser. Som nevnt i innledningen, forventet vi derfor å finne større avvik i medisinske enn kirurgiske DRG-er, og større avvik blant akuttinnleggelser enn blant elektive opphold. For å undersøke om påstandene stemte, delte vi inn endringene av henholdsvis hoveddiagnose og DRG etter innleggelsesmåte og typen behandling som ble gitt.

Tabell 4.7 Antall endringer med betydning for hoveddiagnose etter innleggelsesmåte og behandlingstype.<sup>11</sup> Andeler i parentes

Sykehus	Øyeblikkelig hjelp		Elektiv	
	Kirurgi	Medisin	Kirurgi	Medisin
A	2 (22)	20 (37)	3 (18)	6 (33)
B	2 (13)	20 (42)	3 (19)	8 (40)
C	2 (25)	33 (58)	7 (30)	5 (63)
D	1 (25)	39 (49)	4 (50)	5 (63)
E	3 (38)	38 (50)	2 (18)	2 (40)
Totalt	10 (23)	150 (48)	19 (25)	26 (44)

<sup>11</sup> Opphold registrert uten verken kirurgisk eller medisinsk DRG-type er holdt utenfor. Tre opphold med revidert DRG er derfor utelatt.

Tabell 4.8 Antall endringer med betydning for DRG etter innleggelsesmåte og behandlingstype.<sup>12</sup> Andeler i parentes

Sykehus	Øyeblikkelig hjelp		Elektiv	
	Kirurgi	Medisin	Kirurgi	Medisin
A	2 (22)	17 (32)	3 (18)	8 (44)
B	2 (13)	12 (25)	3 (19)	8 (40)
C	2 (25)	19 (33)	6 (26)	4 (50)
D	2 (50)	34 (43)	2 (25)	2 (25)
E	2 (25)	31 (41)	1 (9)	2 (25)
Totalt	10 (23)	113 (36)	15 (20)	24 (41)

Både når det gjelder revisjon av hoveddiagnose og DRG er forekomsten av avvik utvilsomt størst innenfor medisinske akuttinnleggelser. Totalt 113 opphold av denne typen fikk endret DRG og 150 endret hoveddiagnose. Vi ønsket å kontrollere om disse resultatene skyldtes at et svært høyt antall av de undersøkte oppholdene i utgangspunktet faktisk var øyeblikkelig hjelp innleggelser innenfor medisinske DRG-er, og vektet derfor antallet i forhold til antall opphold av denne typen i utvalget totalt. Tendensen gjorde seg imidlertid fortsatt like sterkt gjeldende blant de medisinske oppholdene. Skillet mellom øyeblikkelig hjelp og elektive innleggelser var også fortsatt tydelig, selv om det ikke lenger var like markant.

Resultatene gir dermed støtte til antagelsen om at kirurgiske innleggelser er enklere å plassere på en entydig måte enn medisinske opphold. Det ser også ut til at akuttinnleggelser er vanskeligere å plassere korrekt enn elektive opphold, noe som muligens kan forklares av at pasienter som innlegges som øyeblikkelig hjelp og får medisinsk behandling gjennomgående har et komplisert sykdomsbilde og svært sammensatte lidelser som det er vanskelig å kode på en entydig måte. I tillegg vil fraværet av et medisinsk prosedyrekodeverk og mangelfull opplæring i bruk av det eksisterende kodeverket kunne bidra til å forklare avvikene.

<sup>12</sup> Opphold registrert uten verken kirurgisk eller medisinsk DRG-type er holdt utenfor. Fire opphold med revidert DRG er derfor utelatt.



## 5 Analyse

Resultatene som ble presentert i forrige kapittel avdekket store avvik mellom opprinnelige og reviderte journaldata. I denne delen av rapporten vil vi drøfte disse funnene nærmere. Avslutningsvis vil vi se nærmere på mulige årsaker til, og konsekvenser av, denne disprepsen.

### 5.1 Hoveddiagnoser

Resultatene fra journalgjennomgangen viste at over 40 prosent av oppholdene fikk revidert hoveddiagnose. Hva kan forklare denne relativt høye andelen, og hvilke konsekvenser har dette i praksis?

#### 5.1.1 Tendenser i datamaterialet

For å kunne trekke tentative konklusjoner om årsaker til den høye revisjonsandelen, er det nødvendig å analysere nærmere hvilke hoveddiagnoser som hyppigst ble revidert, og om spesielle tendenser i datamaterialet eventuelt bør belyses nærmere.

Når det gjelder revisjon på kapittelnivå har kapitlene om svulster (II) og symptomer, tegn, unormale kliniske funn og laboratoriefunn (XVIII) i ICD-10 størst avvik målt i *absolutte tall*, med henholdsvis ti og ni opphold hver. Kapitlene som omhandler sykdommer i sirkulasjonssystemet (IX), sykdommer i fordøyelsessystemet (XI), skader, forgiftninger og visse andre konsekvenser av ytre årsaker (XIX) og faktorer som har betydning for helsetilstand og kontakt med helsetjenesten (XXI) markerer seg også med sju opphold hver. Bildet nyanseres imidlertid noe når avvikene måles som prosentandeler av det totale antallet opphold innenfor hvert kapittel. Da utgjør de *prosentvise* andelene for de to kapitlene endokrine sykdommer, ernærings sykdommer og metabolske forstyrrelser (IV) og symptomer, tegn, unormale kliniske funn og laboratoriefunn (XVIII) 20 og 21 prosent, mens kapitlene sykdommer i sirkulasjonssystemet (IX), sykdommer i fordøyelsessystemet (XI), skader, forgiftninger og visse andre konsekvenser av ytre årsaker (XIX) og faktorer som har betydning for helsetilstand og kontakt med helsetjenesten (XXI) har andeler som tilsvarer henholdsvis 9 prosent, 15 prosent, 19 prosent og 47 prosent. Ved å legge vekt på andelene, peker dermed sistnevnte kapittel seg ut med den klart høyeste revisjonsandelen. Videre ble fem av elleve opphold (45,5 prosent) innenfor kapitlet for sykdommer i blod og bloddannende organer og visse tilstander som angår immunsystemet (III) revidert og innenfor kapitlet skader, forgiftninger og visse andre konsekvenser av ytre årsaker (XIX) ble fem av 13 opphold (38,5 prosent) revidert.

Når avvikene måles i absolutte tall er det altså diagnosekapitlene i ICD-10 knyttet til svulster og unormale kliniske funn/laboratoriefunn som utmerker seg. Målt som prosentandeler er det imidlertid blodsykdommer og hoveddiagnoser relaterte til skader og forgiftninger som er mest utbredt. Revisjonen på kapittelnivå avdekker derfor i all hovedsak at hoveddiagnosene som ble revidert kom fra et stort utvalg av hovedkapitlene i ICD-10 kodeverket. Likevel er det interessant å merke seg at hoveddiagnoser innenfor kapitlet for faktorer som har betydning for helsetilstand og kontakt med helsetjenesten (XXI) utmerker seg med svært høy

forekomst av endringer ved revisjon uavhengig av om man ser på absolutte tall eller prosentandeler.

Revisjon der hoveddiagnosen er innenfor riktig kapittel, men feil kategoriblokk forekom for 48 opphold. En nærmere analyse av dataene viser at denne typen diskrepans mellom opprinnelig koding og revidert koding var svært framtrædende for hoveddiagnoser i kapittel XV i ICD-10 kodeverket. Dette kapitlet kategoriserer diagnoser relaterte til svangerskap, fødsel og barseltid. For sykehus B skyldtes hele ni av 13 avvik i kategoriblokk slik ukorrekt koding. Sykehus A hadde kun fem slike avvik, men av disse hadde hele tre opphold fødsels- eller svangerskapsdiagnoser. Sykehus C hadde fire slike diskrepanser mellom opprinnelig og revidert koding, mens sykehus D hadde tre. Sykehus E har ikke fødeavdeling, og hadde derfor naturlig nok ingen avvik innenfor disse diagnosekodene. Totalt sett, både målt i absolutte tall og som prosentvise andeler av alle opphold, har diagnoser innenfor kapittel XV den klart hyppigste forekomsten av denne typen feilkoding.

På tredje- og fjerdetegnsnivå er det ingen spesielle kapitler i ICD-10 som skiller seg ut med større revisjonsandeler enn de andre. Selv om sirkulasjonssykdommer og sykdommer i fordøyelsessystemet har relativt høyt antall endringer, utmerker ikke disse gruppene seg når avvikene måles som andeler.

Kort oppsummert ser det derfor ut til at enkeltkapitler innenfor ICD-10 kodeverket kun skiller seg ut når det gjelder revisjon på første- og andretegnnivå.

### 5.1.2 Konsekvenser av revisjon

En av de direkte konsekvensene av revisjonen er at den til en stor grad påvirker DRG-grupperingen og dermed har økonomisk betydning. Dernest vil datakvaliteten på de innrapporterte opplysningene ha betydning for hva dataene forskningsmessig kan brukes til. Dersom hoveddiagnosekoden ikke gir et riktig bilde av hva som faktisk feiler pasienten, vil det for eksempel være lite hensiktsmessig å benytte hoveddiagnosen som utgangspunkt for å sortere pasientdata i henhold til sykdomsgrupper. Begge disse aspektene vil imidlertid bli nærmere diskutert under 5.7.

## 5.2 Bidiagnoser

Som vi så i kapittel 4.2, var det relativt mange opphold som fikk endret bidiagnosene etter revisjonen. Dette hadde for flere opphold også direkte betydning for DRG-plasseringen. I dette delkapitlet vil vi se nærmere på om det var noen bidiagnoser som hyppigere ble revidert enn andre.

### 5.2.1 Tendenser i datamaterialet

En sammenstilling av de originale og reviderte bidiagnosene viste at noen diagnosekategorier oftere ble revidert enn andre. Med utgangspunkt i de diagnosene som opprinnelig var satt, utmerket diagnoser fra kapitlet om sykdommer i sirkulasjonssystemet (IX) i ICD-10 seg spesielt. Hele 78 av bidiagnosene innenfor dette kapitlet ble revidert ved journalgjennomgangen. Innenfor kapitlet om svulster (II) ble 29 bidiagnoser revidert, innen kapitlet som omfatter sykdommer i muskel- og skjelettsystemet og bindevev (XIII) var det tilsvarende antallet 26. 25 bidiagnoser innenfor kapitlet "faktorer som har betydning for helse-tilstand og kontakt med helsetjenesten" (XXI) ble revidert, mens det ble revidert 22 diagnoser innenfor kapitlet som omfatter endokrine sykdommer, ernærings sykdommer og metabolske forstyrrelser (IV). Forøvrig fordelte revisjonene seg nokså jevnt.

### 5.2.2 Konsekvenser av revisjon

Den store diskrepansen mellom opprinnelig og revidert koding for bidiagnoser har naturlig nok ofte betydning for hvorvidt opphold blir gruppert i en komplisert eller ukomplisert DRG. Dataene avslørte i tillegg at det var et avvik mellom antallet og rekkefølgen på diagnosene for flere enkeltopphold, noe som også kan ha betydning for DRG-plasseringen.

I forskningsøyemed vil det også være problematisk at det er store avvik mellom innrapporterte opplysninger til NPR og de opplysningene journalgjennomgangen ga grunnlag for å rapportere. For eksempel impliserer resultatene fra denne studien at det ved bruk av bidiagnoser som mål på komplikasjonsgraden ved opphold, bør tas forbehold om datakvaliteten.

## 5.3 Kirurgiske prosedyrekoder

Sammenlignet med revisjonen av hoveddiagnosene og bidiagnosene, hadde endringene av prosedyrekodene relativt lite omfang, og kun i få tilfeller direkte effekt på oppholdenes DRG-plassering.

### 5.3.1 Tendenser i datamaterialet

En sammenligning av opprinnelige og reviderte prosedyrekoder viser altså at avvikene er relativt små, og i den grad det forekommer diskrepans, er det stort sett fordi det ved journalgjennomgangen ble funnet grunnlag for å *tilføye* koder.

Korrigerende avvik forekom i sju tilfeller innenfor kapitlet om bevegelsesapparatet i NCSP. Videre ble fire endringer gjort innenfor både kapitlet om transluminal endoskopi og kapitlet som omhandler fordøyelsesorganer og milt. Forøvrig fordelte endringene seg nokså tilfeldig. Vi kan derfor konkludere med at avvikene mellom opprinnelig og revidert koding av NCSP-koder var relativt små, men at det ved revisjon ble funnet grunnlag for å gjøre enkelte tilføyelser.

## 5.4 HDG

### 5.4.1 Tendenser i datamaterialet

Bakgrunnen for å se spesielt på omfanget av revisjon med betydning for HDG-plassering, var først og fremst at dette gir et inntrykk av *graden* av avvik i DRG. Omfanget av revisjonen er riktignok, og som forventet, mindre omfattende på HDG-nivå enn på DRG-nivå, men likevel endrer hele 13,8 prosent av oppholdene HDG ved revisjon. Som andel av alle opphold med endret DRG, utgjør dette nesten 42 prosent. Denne relativt høye andelen kan igjen ses i sammenheng med andelen av hoveddiagnoserevisjoner på kapittelnivå.

Når HDG-revisjonene fordeles på enkeltinstitusjonene ser vi at 19 av revisjonene skjedde ved sykehus D og 17 ved sykehus C. Institusjon A, B og E hadde lavere antall med henholdsvis ti, ti og 13 revisjoner av denne typen. Dette harmonerer med mønstret fra hoveddiagnoserevisjonen der sykehus C og D begge hadde høy grad av diskrepans mellom opprinnelig koding og revidert koding generelt, og på kapittelnivå spesielt.

## 5.5 DRG

Kapittel 4 viste at en tredjedel av oppholdene fikk revidert DRG ved journalgjennomgangen. Vi vil nå drøfte nærmere hvorvidt enkelt DRG-er skilte seg ut, og hvilke konsekvenser revisjonen av DRG-ene hadde.

### 5.5.1 Tendenser i datamaterialet

I datamaterialet var 207 forskjellige DRG-er representerte. Innenfor 108 av disse forekom det revisjon, og for størstedelen av disse var det kun ett opphold som ble revidert. Følgende tre DRG-er fikk revidert fem opphold:

- DRG 12: Degenerative sykdommer i nervesystemet
- DRG 89: Lungebetennelse & pleuritt for pasienter eldre enn 17 år med b/k
- DRG 467: Påvirkning på helsetilstanden (ikke tildelt annen DRG)

Tre DRG-er fikk revidert fire opphold:

- DRG 141: Synkope og kollaps med b/k
- DRG 281: Skade av hud, underhud og bryst for pasienter eldre enn 17 år uten b/k
- DRG 346: Ondartet svulst i mannlige kjønnsorganer med b/k

Andelsmessig utgjorde revisjonsandelen for hver av disse seks DRG-ene henholdsvis 83 prosent, 46 prosent, 83 prosent, 80 prosent, 100 prosent og 100 prosent. Blant gruppene som utmerker seg, er det fire opphold innenfor par-DRG-er og én som har tilføyelsen "ikke tildelt annen DRG". Blant oppholdene innenfor par-DRG 89 ble to tilfeller revidert til den tilhørende par-DRG-en uten b/k, mens ett opphold ble revidert til KOLS, ett til en ryggslidelse-DRG og ett til sykdommer i åndedretsorganer. For DRG 141 ble ett opphold revidert til den tilhørende par-DRG-en uten b/k, mens de tre øvrige ble endret til øsofagitt, hjertesvikt & ikke-traumatisk sjokk og organisk betingede sinnslidelser & mental retardasjon. DRG 281, ble i tre tilfeller omkodet til traume for pasienter eldre enn 17 år uten b/k og i ett tilfelle til en DRG for neseskade & medfødt nesedeformitet. For DRG 346 ble ett opphold revidert til den tilhørende par-DRG-en uten b/k, mens de tre øvrige endringene var til sykdommer i bukspyttkjertelen, patologiske brudd & ondartede svulster i HDG 8 og øsofagitt, gastroenteritt & diverse med b/k.

Totalt var det 38 revisjoner *innenfor* par-DRG-er i utvalget. Blant disse ble 24 opphold revidert fra en komplisert til en ukomplisert DRG, mens endringen gikk i motsatt retning for de 14 resterende oppholdene. Dette tyder på at en stor del av avvikene skyldes uklarheter i bruk av kompliserende hoveddiagnoser.

Det store avviket i DRG er i all hovedsak et resultat av revisjonene av hoved- og bidiagnoser. Dette er tydelig ettersom bortimot 60 prosent av oppholdene med revidert hoveddiagnose også havnet i en ny DRG.

### 5.5.2 Konsekvenser av revisjon

Endring av DRG vil nesten alltid medføre endret DRG-vekt, og dermed korrigert refusjonsstørrelse for oppholdet. For de 166 oppholdene der DRG ble endret, utgjorde avviket totalt 19,4 DRG-poeng. Prosentvis utgjør denne differansen 3,7 prosent av den opprinnelige poengsummen.

Som vi så i kapittel 4, varierte differansene betydelig mellom enkeltinstitusjonene, med de klart største avvikene hos sykehusene som fikk redusert refusjon etter revisjon. Sykehus A, D og E utmerket seg spesielt, og det er derfor interessant å se nærmere på hvilke faktorer som utløste vektøkninger. Sykehus A hadde seks opphold som ble revidert fra par-DRG med b/k til ukomplisert DRG. I tillegg ble ett opphold revidert fra DRG 81A til DRG 91B, med den konsekvens at antall DRG-poeng ble redusert fra 4,61 til 0,5. Også sykehus D hadde seks

revisjoner fra komplisert DRG til tilsvarende DRG uten b/k. I tillegg til de økonomiske konsekvensene av endret DRG, gir revisjonen en indikasjon på validiteten til DRG- og vektvariablene i de nasjonale pasientdatafilene.

Som vi også framhevet i kapittel 4, var de totale avvikene i DRG-poeng betraktelig lavere enn de tilsvarende avvikene i hoveddiagnoser og bidiagnoser. Selv om journalgjennomgangen indikerer at det samlet sett har vært en overkoding som har gitt grunnlag for høye utbetalinger til sykehusene, resulterer altså ikke avvikene i særlig store endringer refusjonsmessig.

## 5.6 Avvik mellom opprinnelig og revidert diagnosesetting

Ettersom diagnosesetting og pasientklassifisering ikke er en presis eller eksakt vitenskap, vil journalgjennomganger høyst sannsynlig avdekke et visst avvik. I tillegg vil forskjellige systemusikkerheter kunne forklare den lite tilfredsstillende kvaliteten på journaldata. Et av aspektene knyttet til dette er varierende kompetansenivå på diagnosesetting og journalføring ved sykehusene. Legene som samlet inn våre data har svært god oversikt over det eksisterende kodeverket og praksis for korrekt diagnosesetting. Noe av avviket som ble avdekket kan derfor skyldes behov for oppdatert kompetanse om klassifisering av sykdommer blant sykehuslegene. På den andre siden er diagnosene som opprinnelig ble satt basert på kunnskapene hos, og skjønnet utøvd av, den behandelende lege omkring hva som er den dominerende tilstanden hos pasienten. Denne kunnskapen og skjønnet kan være vanskelig å overprøve ved en journalgjennomgang. Videre er det også mulig at diskrepansen i kodesettingen er preget av at legene som reviderte journalene hadde tilgang til mindre informasjon enn legene som satte den opprinnelige diagnosen. Mens sistnevnte hadde observert pasienten under oppholdet, og muligens også hadde tilgang til opplysninger som ikke ble journalført, var dette ukjent for legene som overprøvde dataene ved journalgjennomgangen. Dette indikerer at det kan være behov for innskjerping av plikten til å journalføre opplysninger om diagnoser og behandling etter Journalforskriften.<sup>13</sup> I tillegg er det grunn til å anta at en forbedring og konkretisering av de eksisterende kodeveiledningene, og da særlig innenfor det medisinske kodeverket (jamfør tabell 10 og 11), vil bidra til at de nevnte systemusikkerhetene kan reduseres. På problemområder kan det også være grunn til konsensuskonferanser om en felles nasjonal forståelse av uklare regler, hvor det påvises avvikende praksis sykehusene imellom.

Resultatene fra denne studien indikerer dessuten at undersøkelsen bør følges opp for å generere en større oversikt over, og innsikt i, problemfeltet. En slik oppfølging kan primært gjennomføres på to måter. Den ene framgangsmåten innebærer at studien gjøres bredere ved å inkludere data fra flere opphold og eventuelt også flere sykehus i datamaterialet. En inkludering av opphold fra 2002, kan videre muliggjøre analyser av utviklingen over tid. Den andre tilnærmingen vil være en dybdeanalyse, der flere opplysninger om de 500 oppholdene samles inn. Slike opplysninger kan for eksempel være knyttet til hvordan den primære kodingen ved sykehuset ble gjort, hvilke rutiner for kvalitetssikring som eksisterer internt ved sykehuset, hvor lenge etter utskrivningen kodingen ble foretatt og om kodingsansvarlige leger opplever deler av kodeverket som spesielt problematisk å anvende.

## 5.7 Praktiske konsekvenser av lav reliabilitet på data

De relativt store avvikene som ble presentert i forrige kapittel har noen praktiske konsekvenser. For det første gir dataene grunnlag for å predikere store variasjoner i registrerte DRG-poeng mellom sykehusopphold der diagnoser og sykdomsbilde er relativt ensartet. Framtidige analyser vil kunne avsløre hvorvidt det eksisterer systematiske variasjoner mellom enkeltinstitusjoner og sågar også mellom enkeltavdelinger ved samme sykehus, når det

---

<sup>13</sup> Se FOR 2000-12-21 nr 1385: *Forskrift om pasientjournal*, § 8, litra f).

gjelder kodingspraksis. Den andre sentrale praktiske konsekvensen av dårlig kodingskvalitet er knyttet til de forskningsmessige implikasjonene. Analyser der informasjon om diagnoser benyttes som utgangspunkt for studier av medisinsk praksis vil på samme måte være preget av stor usikkerhet. I forskningsøyemed er det spesielt alarmerende at ingen studier så langt har bidratt med omfattende opplysninger om hvorvidt de eksisterende registreringene er preget av systematiske feil, eller om forekomsten er av mer tilfeldig art.

Framtidig forskning vil forhåpentligvis gi svaret på dette og andre sentrale spørsmål knyttet til dagens praksis innenfor diagnosesetting og pasientklassifisering ved norske somatiske sykehus. Både undersøkelser som fokuserer på den dynamiske utviklingen i kodekvalitet over tid etter innføringen av ISF i 1997, og studier som inkluderer et bredt utvalg av opphold innenfor en avgrenset tidsperiode vil kunne kaste lys over de eksisterende usikkerhetene knyttet til kodekvalitet.

## 6 Oppsummering

I rapportens første kapittel skisserte vi formålet for studien. Neste kapittel oppsummerte hovedkonklusjonene fra et utvalg av tidligere studier og skisserte den norske historiske konteksten for innføringen av det aktivitetsbaserte inntektssystemet, måten systemet er bygd opp på samt sentrale angrepspunkter som har vært rettet mot det. Kapittel 3 ga en oversikt over den metodiske tilnærmingen som har vært benyttet, mens kapittel 4 presenterte hovedfunnene fra journalgjennomgangen. I forrige kapittel drøftet vi disse funnene nærmere gjennom å påpeke spesielle mønstre i datamaterialet og ved å diskutere konsekvensene av de relativt store avvikene mellom opprinnelige og reviderte data.

Journalgjennomgangen avdekket et forholdsvis stort avvik mellom dataene før og etter revisjon. Ettersom utvalget som ble studert omfattet kun 500 case fordelt på fem sykehus, er det forbundet med statistiske problemer å benytte funnene som grunnlag for å generalisere utover oppholdene ved de fem undersøkte institusjonene. Likevel er det ingen grunn til å anta at funnene ikke gir et realistisk tverrsnittsbilde av dagens status på nasjonalt nivå. Journalgjennomgangen avdekket avvik som antyder at en mer korrekt koding ville gitt reduserte refusjoner for tre av sykehusene, og for de 500 oppholdene totalt. Undersøkelsen viser at avvik i hoveddiagnosekoding var hovedårsaken til det store avviket i DRG. Våre data gir imidlertid ikke grunnlag for å trekke entydige konklusjoner omkring hvilke årsaker avvikene har. Videre er det også viktig å tolke resultatene fra undersøkelsen i lys av kompleksiteten som er knyttet til diagnosesetting. Dette innebærer blant annet at revisjon med utgangspunkt i journaldokumentasjon, ikke bare vil føre til en mer korrekt koding, men også til bedre journalføring. Likevel fremhever vi i kapittel 5 behovet for mer omfattende kodeveiledninger og større fokus på riktig koding ved de enkelte sykehusene. Siden en majoritet av avvikene forekom for medisinske opphold, ser det ut til at forbedringspotensialet er størst innenfor denne typen opphold. Videre vil det være svært nødvendig å gjennomføre flere omfattende studier av kodingskvaliteten, både ved enkeltinstitusjoner, helseregioner og på nasjonalt nivå, med sikte på å avdekke hvilke forhold som kan forklare diskrepans mellom opplysninger dokumenterte i journal og data innrapporterte til NPR. Den sentrale utfordringen videre vil så bli å finne kontrollmekanismer for å nøytralisere eller balansere de eksisterende systemusikkerhetene samt å utvikle kodeveiledninger som det er etablert konsensus om i fagmiljøene.



# Litteraturliste

Barnard, C. & T. Esmond (1981). DRG-based reimbursement: the use of concurrent and retrospective clinical data. *Medical care* 19: 1071-1082.

Corn, R. F. (1981). The sensitivity of prospective hospital reimbursement to errors in patient data. *Inquiry* 18: 351-360.

Doremus, H. D. & Michenzi, E. M. (1983). Data quality: an illustration of its potential impact upon a diagnosis-related group's case mix index and reimbursement. *Medical Care* 21: 1001-1011.

Fisher, E. S., F. S. Whaley, W. M. Krushat, D. J. Malenka, C. Fleming, J. A. Baron & D. C. Hsia (1992). The Accuracy of Medicare's Hospital Claims Data: Progress Has Been Made, but Problems Remain. *American Journal of Public Health* 82: 243-248.

Forskrift om pasientjournal.

Horn, S. D., R. A. Horn, P. D. Sharkey & A. F. Chambers (1986). Severity of illness with DRGs. *Medical Care* 24: 225-235.

Hsia, D. C., W. M. Krushat, A. B. Fagan, J. A. Tebbutt & R. P. Kusserow (1988). Accuracy of diagnostic coding for medicare patients under the prospective-payment system. *The New England Journal of Medicine* 318: 352-355.

Johnson, A. N. & G. L. Appel (1984). DRGs and hospital case records: implications for Medical case mix accuracy. *Inquiry* 21: 128-134.

KITH (2000). ICD-10 kodeveiledning 2000. Trondheim.

Kjekshus, L. E. (2002). "Ring fencing" of Elective Surgery: Does it affect Hospital Efficiency? Artikkel under publisering. Senter for Helseadministrasjon, Medisinsk fakultet, Universitetet i Oslo.

Nilsson, A. C., C-L Spetz, K. Carsjö, R. Nightingale & B. Smedby (1994). Slutenvårdsregistrets tillförlitlighet. Diagnosuppgifterna bättre än sitt rykte. *Läkertidningen* 91: 598-605.

Nilsson; C. A, S. Gårdmark, H. Johnell & S. Montan (1992). Fel diagnos, fel operation, fel DRG. Stockholm: Spri rapport 341.

Pedersen, M. & K. Solstad (2002). Kostnadsvekter til NorskDRG-versjonen av DRG-systemet, Trondheim: SINTEF Unimed Pasientklassifisering og finansiering.

Sosial- og helsedepartementet (1996). Informasjon om ny finansieringsordning for somatiske sykehus, Oslo.

Sosial- og helsedepartementet (2001). Innsatsstyrt finansiering 2001 med prisliste, Oslo.

Statens Helsetilsyn (2000). ICD-10. Den internasjonale klassifikasjonen av sykdommer og beslektede helseproblemer, 10. revisjon. Oslo: Elanders Publishing AS.

Statens Helsetilsyn og KITH (2001). Klassifikasjon av kirurgiske inngrep 2001. Klinisk prosedyrekodeverk, Trondheim.

Steinum, O (1995). Korrekt diagnos ger rätt ersättning. DRG-systemet förutsätter intern kontroll. *Läkertidningen* 92: 1715-1717.

# Appendix

Tabell A Opphold med revidert HDG. Opprinnelig HDG i venstre kolonne

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0															1											1
1				1	1					2	1						1				1					4
2																										
3										1																
4		1					1	2										1								
5		2	1			1	2	1				1									1					1
6					1			1																		
7				1	1						1									1						
8										1									2							
9				1	1		1	1												1			3			
10						2																				1
11		1					1																			
12							1	1	1			1														
13											1								1							
14																										
15																										
16																										
17							1				1												2			
18																										
19							1																			
20		1																								
21										2																
22																										
23					1	2																	1			
24																										
25																										