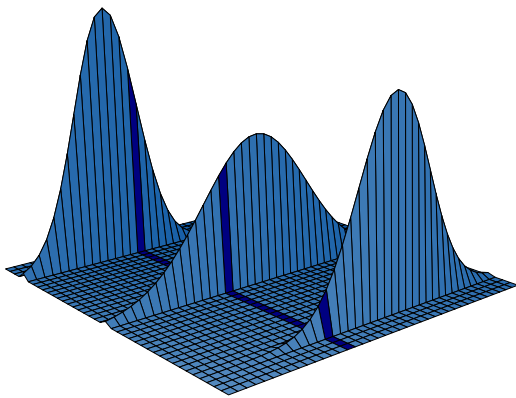


## Forskjeller i helsetjenesten

Hva kan vi lære av variasjonen  
i medisinske kvalitetsregistre?



Dokumentet er en oversettelse, videreutvikling og tilpassning av:  
*Ännu bättre vård – Vad kan vi lära från variationen i Öppna Jämförelser?*  
Sveriges kommuner och landsting, 2012, ISBN: 978-91-7164-896-9

Nasjonalt servicemiljø for medisinske kvalitetsregistre  
Senter for klinisk dokumentasjon og evaluering  
Hovedforfatter: Alexander Walnum  
**ISBN: 978-82-93141-12-9**

# Forord

Arbeidet med å vise åpenhet i framlegging av resultater fra helsevesenet har en lang historie, men det er først i den senere tid at mer omfattende og offentlig tilgjengelig resultatrapportering har vært aktuelt. I Sverige har *Öppna jämförelser*, en årlig offentliggjøringen av resultater primært hentet fra medisinske kvalitetsregistre, vært en viktig del av denne utviklingen. I Norge arbeides det nå med å utvikle et lignende system for systematisk rapportering og offentliggjøring av resultater fra kvalitetsregistrene, og det er i denne sammenheng at dette dokumentet er utarbeidet.

Dokumentet handler om hvordan man ut fra ulike perspektiver skal forstå, håndtere og trekke konklusjoner fra variasjonen som forekommer i kvalitetsregistrene. Primært vil et overordnet perspektiv legges til grunn og teksten tar for seg hvordan man med et slikt (ledelses-)perspektiv kan stimulere til enda bedre helsetjenester. Det gjennomgående mantra er: «Den som ikke forbedrer seg, slutter snart å være god.»

Dokumentet er i bunnen en oversettelse av den svenske rapporten *Ännu bättre vård – Vad kan vi lära från variationen i Öppna Jämförelser?* utgitt av Sveriges Kommuner och Landsting i 2012, men både tekst, eksempler og figurer har vært gjennom en vesentlig videreutvikling i tillegg til en tilrettelegging til norske forhold. Forfatter av den svenske versjonen er Bo Bergman, professor i kvalitetsutvikling ved Chalmers Tekniska Högskola, mens videreutvikling, tilpasningen til norske forhold, figurer og layout er gjort av Alexander Walnum, statistiker ved Nasjonalt servicemiljø for medisinske kvalitetsregistre.

For mer informasjon om arbeidet med resultatrapportering og offentliggjøring fra medisinske kvalitetsregistre samt informasjon om kvalitetsregisterfeltet generelt henvises det til nettsiden til Nasjonalt servicemiljø for medisinske kvalitetsregistre på [www.kvalitetsregistre.no](http://www.kvalitetsregistre.no).

*Tromsø, september 2013*

# Innhold

<b>1</b>	<b>Innledning</b>	<b>1</b>
1.1	Enda bedre helsetjenester . . . . .	1
1.2	Formål . . . . .	2
1.3	Dokumentets oppbygning . . . . .	2
1.4	Advarsel . . . . .	3
<b>2</b>	<b>Om medisinske kvalitetsregistre</b>	<b>5</b>
2.1	Helseregistre . . . . .	5
2.2	Medisinske kvalitetsregistre . . . . .	5
2.2.1	Innhold og oppbygging . . . . .	5
2.2.2	Formål . . . . .	6
2.2.3	Nasjonal status . . . . .	6
2.3	Mer informasjon . . . . .	6
<b>3</b>	<b>Kilder til variasjon</b>	<b>7</b>
3.1	Resultat, prosess, struktur og system . . . . .	7
3.1.1	Resultater . . . . .	7
3.1.2	Proessen skaper resultatene . . . . .	7
3.1.3	Systemet og dets struktur bestemmer prosessen . . . . .	8
3.2	Resultat- og prosessmåling . . . . .	9
3.2.1	Resultatmåling for enkeltpasienter . . . . .	9
3.2.2	Prosessmåling for enkeltpasienter . . . . .	9
3.3	Måling av den overordnede prosessen . . . . .	10
3.4	Måling over tid . . . . .	11
3.5	Måleproblematikk . . . . .	11
3.6	Oppsummering . . . . .	11
<b>4</b>	<b>Variasjon i et ledelsesperspektiv</b>	<b>13</b>
4.1	Grunnleggende forutsetninger . . . . .	13
4.1.1	Visjon og mål . . . . .	13
4.1.2	Forbedringsstrategi . . . . .	13
4.2	Illustrasjon av prosessforbedringer . . . . .	15
4.3	Illustrasjon av læring fra kvalitetsregistre . . . . .	16
4.3.1	Er variasjonen naturlig? . . . . .	17
4.4	Å håndtere variasjon . . . . .	17

<b>5</b>	<b>Fenomenet variasjon</b>	<b>19</b>
5.1	Variasjon	19
5.1.1	Synkron og diakron variasjon	19
5.2	To typer studier	21
5.3	En analytisk tilnærming	23
5.4	To typer årsaker til variasjon	23
5.5	Når har noe hendt?	26
<b>6</b>	<b>Å beskrive variasjon</b>	<b>29</b>
6.1	Typer variasjon	29
6.2	Prosesser i statistisk likevekt	30
6.3	Usikkerhet, sannsynlighet og middelverdi	30
6.4	Å illustrere variasjon	32
<b>7</b>	<b>Fallgruver</b>	<b>33</b>
7.1	Sentralitets- og spredningsmål	33
7.2	Faren med sentralitetsmål	33
7.3	Statistisk skjevhet	34
7.3.1	Utvalgsskjevhet	34
7.3.2	Den glemte nevneren	35
7.4	Årsakssammenheng	36
7.5	Måling av andeler som innfrir et kriterium	37
7.5.1	Betydningen av operative definisjoner	37
7.6	Datakvalitet og tillit til data	38
7.7	Mer enn numerisk variasjon	39
7.8	Et motstykke til fallgruvene	39
<b>8</b>	<b>Refleksjoner og anbefalinger</b>	<b>41</b>
8.1	Muligheter	41
8.2	Resultatbasering og forbedringsarbeid	41
8.3	Supplering av rapporteringen	43
8.3.1	Forbedringstakt	43
8.3.2	Spredningen har betydning	43
8.3.3	Konfidensgrader	43
8.3.4	Variasjon og pasientsammensetning	43
8.4	Datakvalitet	44
8.5	Læring og forbedringsmuligheter	44
8.6	Spørsmål til refleksjon	44
8.7	Avslutning	45
<b>A.</b>	<b>Historisk tilbakeblikk</b>	<b>47</b>
<b>B.</b>	<b>Sluttnoter og forslag til videre lesning</b>	<b>49</b>
	<b>Bibliografi</b>	<b>54</b>
	<b>Register</b>	<b>57</b>



# Kapittel 1

## Innledning

Systematisk bruk av medisinske kvalitetsregistre gir mulighet til langsgående evaluering av helsetjenesten, men generelt er det grunn til å tro at det hersker usikkerhet, og sannsynligvis også uvitenhet, omkring hvordan variasjonen i resultatene fra kvalitetsregistrene skal tolkes. Temaet har rett og slett ikke vært viet mye oppmerksomhet. I det innledende kapitlet vil kunnskapsbehovet for å tolke variasjonen i medisinske kvalitetsregistre drøftes.

### 1.1 Enda bedre helsetjenester

Muligheten for tilgang på oppdaterte og aktuelle resultater fra medisinske kvalitetsregistre er i stor grad med på å påvirke tankesettet og ledelsesfunksjonene i helsetjenesten. Resultatrapportering kan fungere som en bro mellom ledernivåene og de mer operative delene av helsetjenesten. Hvis man kan sikre at resultatene fra medisinske kvalitetsregistre og andre relevante registre fremstår som troverdige, rettferdige og nyanserte, kan dette bidra til å skape et felles utgangspunkt for viktig dialog om ressursfordeling og forbedringsarbeid.

Ønsket om å forbedre helsetjenestene har vært en viktig drivkraft ved opprettelsen av de medisinske kvalitetsregistrene<sup>1</sup>. Norge har internasjonalt sett en god helsetjeneste i dag, i det minste i et medisinsk perspektiv, men den kan bli enda bedre. Spørsmålet er hvordan man

kan anvende de opplysningene og den variasjon som framkommer i de medisinske kvalitetsregistrene i arbeidet med å skape enda bedre helsetjenester. Man ønsker et helsevesen som er sikrere, mer kunnskapsbasert, mer effektivt, mer pasientfokusert, mer tilgjengelig og mer likestilt for alle pasientgrupper.

Man ser en hel del variasjon i resultatene fra de medisinske kvalitetsregistrene, men hvordan skal den tolkes? Hva er bra og hva er mindre bra, hva skal man reagere på, og hvilke beslutninger kan treffes på grunnlag av variasjonen? Deler av helsetjenesten kan ha funnet bedre måter å håndtere visse pasientgrupper på eller funnet nye organisasjonsformer. Andre har kanskje ennå ikke tatt inn over seg eller tatt i bruk nye behandlingsmetoder, eller har organisasjonsformer som påvirker resultatene negativt. En del variasjon kan bero på tilfeldigheter, mens andre variasjoner kan skyldes at ulike deler av helsetjenesten har pasienter med sykdommer, som til tross for samme diagnose, har ulik alvorlighetsgrad (eller ulik «**case-mix**»<sup>2</sup>). Kildene til variasjon er mange. For å oppnå læring og forbedring i helsetjenestene er det viktig å ha forståelse for variasjon og hvordan den kan anvendes konstruktivt som grunnlag for forbedringer.

Å møte framtidens utfordringer uten store kostnadsøkninger forutsetter et systematisk forbedringsarbeid, som dels handler om forbed-

ringer i lokal sammenheng og dels om å ta til seg nye arbeidsmetoder, altså å trekke lærdom fra forskning og andre virksomheter. Her spiller kvalitetsregistre og andre helseregistre en sentral rolle, og det vesentlige er å ha den organisatoriske evnen til å ta vare på disse mulighetene. Blant annet trengs det systematisk kunnskap om hvordan data fra kvalitetsregistre og andre relevante registre, kan brukes i forbedringsarbeid. Medisinske kvalitetsregistre, spesielt de som er nettbaserte, kan utgjøre en del av det grunnlaget som trengs for å stimulere til og initiere aktiviteter som kan gi enda bedre helsetjenester.

En viktig oppgave for ledere i helsesektoren er å legge grunnlaget for en hensiktsmessig infrastruktur for informasjonshåndtering. At den er hensiktsmessig må innebære at det skal være lett å benytte data lokalt i eget utviklingsarbeid. Nytteverdien av kvalitetsregistre i forbedringsarbeidet er tett knyttet til dette.

Ledere på høyt nivå påvirker strukturene og systemene i helsetjenestene og dermed også forutsetningene for forbedring. Forbedringstakt blir derfor en viktig verdimåler for hvordan de øverste lederne fungerer<sup>3</sup>. På lokalt nivå står de enkelte prosessene og tilhørende aktiviteter i sentrum, i tillegg til pasientene. Dermed blir de absolutte nivåene på resultatene og en direkte tilbakemelding viktigst. For en person som skal velge behandlingssted, kan både de absolutte resultatene og forbedringstakten være av interesse.

## 1.2 Formål

Dette dokumentet henvender seg først og fremst til personer i ulike lederfunksjoner i helsetjenesten. Formålet er å øke forståelsen for hvordan man i et ledelsesperspektiv kan forholde seg til variasjon når man skal tolke og prioritere områder for videre analyse- og utviklingsarbeid basert på resultatene av medisinske kva-

litetsregistre. Dokumentet er likevel ikke utelukkende for de i lederstillinger i helsevesenet ettersom ledelsesperspektivet, eventuelt tolket som et fugleperspektiv, er en egnet innfallsvinkel for alle som ønsker å lære om medisinske kvalitetsregistre og variasjonen som observeres. Spesielt bør media og aktører som skal formidle resultater fra kvalitetsregistrene sørge for å ha en grunnleggende forståelse (tilsvarende stoffet i dette dokumentet) for kvalitetsregistre og variasjon.

## 1.3 Dokumentets oppbygning

Etter litt overordnet informasjon om – eller en kort introduksjon til medisinske kvalitetsregistre – i kapittel 2, drøftes i kapittel 3 generelle kilder til variasjon i kvalitetsregistre, og i kapittel 4 hvordan registrene kan trekkes inn i ledelsesarbeidet. I kapittel 5 og 6 utdypes fenomenet variasjon – hvordan variasjon kan illustreres og måles samt hvordan man kan omtale variasjon. Da det er lett å gå seg vill, vil noen fallgruver bli adressert i kapittel 7.

I kapittel 8 reflekteres det over variasjon og hvordan variasjon mest hensiktsmessig kan tolkes og illustreres for å utgjøre et godt grunnlag for framtidige beslutninger.

For å øke lesbarheten er henvisninger til litteratur ikke satt direkte inn teksten. I stedet anvendes sluttnoter og forslag til videre lesning, i eget vedlegg. Det er også inkludert et vedlegg med et historisk tilbakeblikk, i tillegg til bibliografi/ referanseliste og stikkordsregister. Sluttnoter i teksten er markert med hevet skrift (slik<sup>x</sup>) mens kildehenvisninger er angitt med klammeparanteser (slik [x]). Videre er sentrale konsepter og begreper uthevet med **fet skrift**.



## 1.4 Advarsel

Uansett hvor viktig det er å måle og sikre tilbakeføring av resultater, er det et par ting man alltid bør ha i mente: Det er ikke alt som enkelt lar seg måle, eller som i det hele tatt lar seg måle, og målingene man bruker kan være beheftet med feil av forskjellig slag. Hva er det egentlig målestørrelsene representerer? Hvor mye av et sykdomsforløp kan man faktisk fange opp i numeriske termer? Som Albert Einstein uttrykte det: «Not everything that is counted counts, and not everything that counts can be counted»<sup>4</sup>.

En mulig komplikasjon ved alle former for måling er at selve måleprosessen kan gi opphav til feil og avvik av mer eller mindre systematisk art, slik at det fenomenet man egentlig vil måle, ikke gjenspeiles i målestørrelsene. Det hviler et stort ansvar på ledelsen på ulike nivåer for å håndtere feilrisiko som skyldes svikt i organisasjonen.



## Kapittel 2

# Om medisinske kvalitetsregistre

Et **medisinsk kvalitetsregister** er kort oppsummert en strukturering av klinisk relevante opplysninger i pasientforløp. Målsetningen er å utgjøre et kunnskapssystem for bruk i alle nivåer i helsetjenesten for fortløpende læring, forbedring, prioritering og styring. I dette kapittelet gis en skjematisk fremstilling av kvalitetsregistrenes rolle i helsetjenesten.

### 2.1 Helseregistre

Det finnes ulike typer **helseregistre**<sup>5</sup> som alle har et overordnet formål om å bedre helsetjenesten. Registerne dekker ulike medisinske fagområder og kan deles inn etter mer spesifikke formål – som administrasjon, styring, kvalitetsutvikling og forskning – eller juridisk hjemmel i lov, forskrift eller konsesjon. I tillegg til en del større *sentrale* helseregistre, som har som formål å ivareta landsomfattende funksjoner og oppgaver, finnes det en rekke *lokale* og *regionale* helseregistre. Disse er i utgangspunktet opprettet av helseforetak (inkludert regionale), men kan inneholde opplysninger av landsdekkende karakter. Helseregistrene omfatter både sykdoms-, helsetjeneste- og kvalitetsregistre samt andre registre som ikke har pasientbehandling som sitt primære formål.

### 2.2 Medisinske kvalitetsregistre

De medisinske kvalitetsregistrene, som altså er en delmengde av helseregistrene, er registre hvor det fortløpende registreres opplysninger for en avgrenset pasientgruppe med utgangspunkt i individuelle behandlingsforløp, hvor opplysningene skal kunne fungere som en **kvalitetsindikator**<sup>6</sup> for de behandelende enhetene og helsetjenesten generelt.

#### 2.2.1 Innhold og oppbygging

For å tilfredsstille kravene til et medisinsk kvalitetsregister og fungere som grunnlag for lokalt og nasjonalt **kvalitetsforbedringsarbeid**, må opplysningene om behandlingsopplegg være både strukturerte og detaljerte med tanke på diagnostikk og behandling. Med dagens utforming av pasientjournaler er det ikke mulig å hente opplysninger til kvalitetsregistre direkte derfra. Opplysninger til kvalitetsregistre må derfor inntil videre registreres og behandles separat.

Juridisk er de medisinske kvalitetsregistre som hovedregel basert på informert **samtykke** fra den registrerte, men enkelte kvalitetsregistre er del av de sentrale helseregistrene og er dermed unntatt fra kravet om samtykke.

Innholdsmessig dekker kvalitetsregistrene

typisk detaljer om diagnostikk og behandling for én bestemt pasientgruppe. En «hendelse» i registeret oppstår ved kontakt med helsevesenet i forbindelse den spesifikke sykdommen eller behandling av denne.

Man er helt avhengig av helseinstitusjonenes og helsepersonellens engasjement og innsats i forbindelse med både registrering og anvendelse av resultater/registerdata for at et kvalitetsregister skal fungere hensiktsmessig. Kvalitetsregistrene er i hovedsak bygget opp av helsepersonell, ofte med en akademisk tilknytning, som skal rapportere data og bruke resultatene. Registrene er med andre ord etablert og drevet av de medisinske fagmiljøene selv, og den nære tilknytning til de kliniske miljøene er viktig for god innregistrering og målrettet bruk av data.

Det **nasjonale servicemiljøet** for medisinske kvalitetsregistre er opprettet for å bidra til at kvalitetsregistrene i enda større grad enn i dag kan oppfylle sine formål.

### 2.2.2 Formål

De medisinske kvalitetsregistrene har som hovedformål å følge med på og forbedre *kvaliteten* på behandling av sykdom, derav navnet. Systematisk bruk av medisinske kvalitetsregistre gir muligheten å evaluere prosesser og resultater i pasientbehandlingen over tid og i sanntid. Dokumentasjon fra kvalitetsregistre skal gi beslutningsgrunnlag for bedre kvalitet og riktige prioritering i helsetjenesten, og slik bidra til en bedre helsetjeneste til befolkningen.

I tillegg til kartlegging av sykdom gir kvalitetsregistrene mulighet til å avdekke eventuelle forskjeller i kvalitet mellom sykehus eller ulike sosiale, medisinske, kulturelle og geografiske grupper. På denne måten er registrene et viktig redskap for å sikre et likeverdig helsetilbud av god kvalitet uavhengig av kjønn, bosted og sosial status. Helsetjenesten skal drive kunnskapsbasert og faglig forsvarlig, og alle skal ha

lik tilgang til medisinsk behandling, pleie og oppfølging.

Registerdata kan også brukes til *forskning* på behandlingsmetoder og resultater samt *oppfølging* av klinisk praksis (avdekke om faglige retningslinjer blir fulgt og gir ønsket effekt). Man kan også sammenligne effektene av ulike tiltak og behandlingsopplegg, på både kort og lang sikt, og anvende dette i utformingen av anbefalinger og faglige retningslinjer. Kvalitetsregistrene gir med andre ord grunnlag for, og unike muligheter til, å finne gode eller bedre behandlingsmetoder og forebyggende tiltak. I mange tilfeller er randomiserte studier umulige å gjennomføre av etiske, praktiske eller økonomiske hensyn. Nyttens av registerbaserte studier kommer både fellesskapet og ulike pasientgrupper til gode (utover anvendelse av registerdata i behandlingen av enkeltpasienter).

### 2.2.3 Nasjonal status

Et medisinsk kvalitetsregister kan bli tildelt **nasjonal status**, noe som betyr at registeret innfrir visse kriterier<sup>7</sup> og at et regionalt helseforetak har påtatt seg databehandlingsansvaret, som – i tillegg til å være et juridisk begrep – inkluderer ansvar for oppfølging, videreutvikling og drift av registeret. Nasjonal status beslattes av Helsedirektoratet, og per 2013 er det 45 registre som har fått denne statusen. Disse er ikke alle nødvendigvis landsdekkende, men har i det minste som mål å bli det.

## 2.3 Mer informasjon

For mer informasjon om medisinske kvalitetsregistre og arbeidet med disse henvises det til nettsiden: [www.kvalitetsregistre.no](http://www.kvalitetsregistre.no) hvor spesielt dokumentet **Håndbok for medisinske kvalitetsregistre**<sup>8</sup> kan anbefales for de som ønsker en fordypning.

## Kapittel 3

# Kilder til variasjon

For å se hvordan variasjon i helsetjenestene og kvalitetsregistre har betydning for ledere på ulike nivåer – både politisk og administrativt – bør man reflektere over de ulike kildene til variasjon og hva en leder kan påvirke. Som en introduksjon presenteres et enkelt bilde av hvordan et resultat og variasjonen i dette resultatet skapes i en organisasjon i helsesektoren, et tema som bør interessere de som ønsker å betrakte og forstå variasjonen i helsetjenestene. Vi tar utgangspunkt i en enkel modell som opprinnelig ble utarbeidet av Avedis Donabedian [7], men som er videreutviklet i denne teksten. Beskrivelsen blir nødvendigvis forenklet, men et antall viktige aspekter vil likevel bli belyst.

### 3.1 Resultat, prosess, struktur og system

I dette avsnittet presenteres en modell for hvordan helsetjenestene skaper **resultater** i sine **prosesser**, som i sin tur avhenger av egenskaper i det overordnede **systemet** og dets **struktur**<sup>9</sup>. Strukturen innbefatter ikke bare den fysiske utformingen, ressursene som er tilgjengelige og informasjonshåndteringen som finnes, men også og kanskje enda viktigere systemets lederskap, læringsmekanismer, kultur og de incitamentene eller drivkrefter som finnes i systemet. Det er framfor alt denne strukturen den overordnede ledelsen kan påvirke.

#### 3.1.1 Resultater

God behandling og pleie handler om å oppnå gode resultater i helsevesenet. Resultatene er av forskjellige slag; helbredelse, sykdomsforløp, funksjonsnivå, arbeidslivsstatus, pasientopplevelse, smerte, pasientskader og så videre. Noe av dette fanges opp i satsningen på, og rapporteringen fra, de medisinske kvalitetsregistrene. Ettersom det er variasjoner mellom pasienter som kan skyldes genetiske, fenotypiske eller andre individuelle egenskaper, vil også resultatene fra kvalitetsregistrene variere. Behandlerne er også individer som seg imellom kan gjøre ting litt annerledes, og de gjør ikke selv alltid nøyaktig det samme i alle sammenlignbare situasjoner. Variasjon kan også skyldes forskjeller i arbeids- og behandlingsmetoder. Selve prosessen rundt behandling og pleie bestemmer i stor grad hvilke resultater som oppnås.

#### 3.1.2 Prosessen skaper resultatene

Prosessen, inklusive behandlingspersonalet, som en pasient som søker hjelp fra helsetjenesten med en sykdomstilstand møter, består av en rekke aktiviteter som i prinsippet gjentas for hver ny pasient med samme eller tilsvarende diagnose – men med variasjoner. **Prosessen** er det samlede settet med aktiviteter som skaper verdi for en viss pasientkategori<sup>10</sup>.

Det er viktig å reflektere over prosessens omfang. Hvis det for eksempel er tale om brystkreft har man kanskje tradisjonelt snakket om «fra en kul ble oppdaget i brystet til behandling og etterbehandling». Men i dag bør man kanskje se det i et mye videre perspektiv – kanskje «fra livsstil og en DNA-undersøkelse og mammografi (hvis hyppighet avhenger av genetiske egenskaper) helt til overlevelsestid eller død»[4]. I mange tilfeller kan forebyggende aspekter være av stor betydning. Mye taler for at vi går – og må gå – fra behandling til behandling i kombinasjon med helsefremmende tiltak, ikke bare på det retoriske plan.

### 3.1.3 Systemet og dets struktur bestemmer prosessen

Figur 3.1 viser hvordan systemet og dets struktur bestemmer forutsetningene for prosessen. Figuren er inspirert av en tilsvarende figur av Avedis Donabedian som tidlig reflekterte over kvalitet i helsetjenestene, men vi tillegger imidlertid begrepet struktur<sup>11</sup> en videre betydning enn i den opprinnelige modellen.

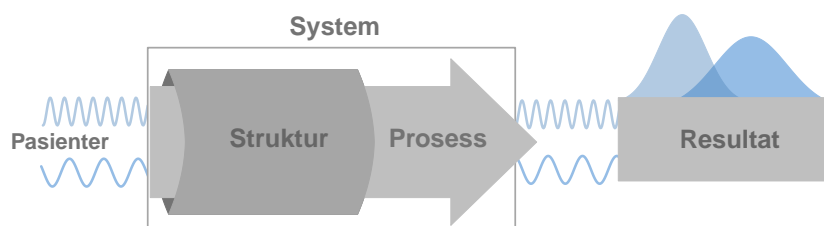
De samme menneskene som utfører aktivitetene i prosessen og som dermed utgjør en viktig del av den, er også de som har de holdninger og normer sentrale for strukturen. Strukturen er imidlertid mer enn delte holdninger og normer. I strukturen (eller kanskje heller «i det sosiale systemet hvis struktur det her er tale om») finnes også de personer i ledende stillinger (formelt og uformelt) som skaper mening og forståelse for hvordan man kan betrakte organisasjonsenheter og fenomener i systemet, og som dermed påvirker strukturen. Men også fysiske rammer har betydning – som betydningsbærere (såkalte forbedringstavler kan være en illustrasjon), som muliggjørere (eksempelvis et medisinsk-teknisk hjelpemiddel eller et IT-system, men også forbedringstavlen og møteplassen for forbedringsarbeid) og som

hindringer (eksempelvis lokalenes fysiske utforming og dårlig fungerende IT-systemer).

Det er ofte ledelsen som skaper den betydning og forståelse for virksomheten som vil danne basis for holdningene, normene og meningsdannelsen i organisasjonen. Denne meningsdannelsen utgjør grunnlaget for kulturen som utvikler seg i organisasjonen. Man må imidlertid huske på at det ikke bare er ledelsen slik den er beskrevet i organisasjonskartet som har betydning. De uformelle lederne kan ha vel så stor betydning for meningsdannelsen i organisasjonen. For å markere det dynamiske perspektivet – den stadige påvirkningen av strukturen fra handlinger og beslutninger og deres meningsskapende innhold – byttes begrepet organisasjon i dag ofte ut med **organisering**. Den uformelle strukturen i en organisasjon gjenskapes og forandres stadig avhengig av hva som hender i organisasjonen og hvordan disse hendelsene tolkes av organisasjonens medlemmer.

Om man for eksempel ikke gir påtale til en ansatt med pasientkontakt som har langermett frakk når han eller hun ifølge reglene skulle hatt korte ermer, uthules regelen. Dette skjer uansett hvem som gjør observasjonen og ikke påtaler det, men betydningen blir enda større om det er noen med høy legitimitet i organisasjonen som ser gjennom fingrene med det. Enda større påvirkningskraft har det om den langermede selv er en formell eller uformell leder i systemet.

En annen illustrasjon kan være et altfor ensidig fokus på kortsiktig kostnadsreduksjon som fører til kortsiktige besparelser og dårligere resultat for pasientene og på lang sikt til kostnadsøkning. Dette kan få negativ innvirkning på strukturen, tankesettet, systemet og dermed på prosessene som skaper resultat for pasienten. Beslutninger som prioriterer kortsiktige besparelser på bekostning av annet, gir tydelige signaler om hva som prioriteres i organi-



**Figur 3.1:** Variasjonen i resultatene er avhengig av variasjonen i pasientene og variasjonen i prosessaktivitetene. Hvor godt prosessen fungerer avhenger av strukturen i systemet prosessen er en del av.

variasjonen – hva som anses som viktig. Iblant gjør kanskje en ukontrollerbar virkelighet at man må ta en beslutning som kan oppfattes som problematisk av medarbeiderne – i slike tilfeller kreves en stor pedagogisk innsats og åpenhet for å unngå uønskede effekter på meningsdannelsen i organisasjonen og dermed på strukturen.

Disse eksemplene gir selvfølgelig et forenklet bilde av den stadig pågående meningsdannelsen i organisasjonen. Men det viser at det ikke bare er den formelle strukturen som gjelder, den uformelle delen av strukturen er trolig enda viktigere, og den er det mulig å påvirke positivt ved hjelp av dialog, støtte og rammebetingelser.

**Strukturen** er det som skaper eller begrenser handlingsrommet for dem som arbeider i organisasjonen, og gir deres handlinger en retning, blant annet gjennom gjeldende normer og uskrevne regler. Politikere og den overordnede ledelsen har muligheter for å påvirke strukturen i helsetjenestene gjennom dialog, støtte og rammebetingelser.

## 3.2 Resultat- og prosessmåling

For å kunne fange variasjonene i prosessen trenger man måltall. I dette delkapitlet drøftes et antall ulike typer måltall og hvordan de varierer.

### 3.2.1 Resultatmåling for enkeltpasienter

Resultatene i helsevesenet kan måles på forskjellige måter. Hvordan har selve behandlingen lyktes rent klinisk? Hva innebar det for pasienten? Hvordan opplevde pasienten sin status etter behandlingen? Hvor fornøyd var pasienten med behandlingen? Mens de to første spørsmålene kan være utgangspunkt for kliniske resultatmålinger, gir det tredje opphav til pasientrapporterte resultatmålinger (såkalte PROM – Patient Reported Outcome Measure) og det siste til måling av pasienttilfredshet.

### 3.2.2 Prosessmåling for enkeltpasienter

For hver enkelt pasient kan man framskaffe et antall mål på behandlingen pasienten har fått. Behandlingen av en pasient omfatter et stort antall aktiviteter som inngår i den individuelle behandlingsprosessen i tillegg til ventetider mellom de ulike aktivitetene. For hver aktivitet kan man danne ett eller flere måltall; hvordan ble aktiviteten utført, hvor lang tid tok den osv. Behandlingsprosessen for en pasient med hoftebrudd består blant annet av: ventetid på akuttten, tid til røntgen, faste før operasjonen, operasjonsmetode, valg av tiltak (hel eller halvprotese, eventuelt osteosyntese), tilgang til smertelindring etter tiltak, muligheter for rehabilitering osv. Målinger relatert til dette er relevante for den enkelte pasient.

Det er ikke alltid det er mulig å skille prosessmål fra resultatmål på en god måte, og det er heller ikke alltid nødvendig. For en diabetespasient kan man for eksempel betrakte HbA<sub>1c</sub> (blodprøve som måler gjennomsnittlig blodsukker de siste 8-12 ukene)<sup>12</sup> som både et prosessmål og et resultatmål.

Ved valg av relevante prosessmål finnes det en underliggende antakelse om hva som kan påvirke det endelige resultatet for pasienten – man antar at det finnes en (i alle fall statistisk) årsakssammenheng mellom prosessmål og det endelige resultatet<sup>13</sup>. Om det for eksempel finnes **evidens** for (eller kunnskap om) at X gir økt velbefinnende og bedre helse og at det derfor er viktig å informere pasienten om X før operasjonen, kan «informasjon om X er gitt» fungere som et prosessmål.

Oftest kreves det data fra både prosess- og resultatmålinger som omfatter mange pasienter i den aktuelle kategorien for å kunne si noe om det finnes en slik årsakssammenheng<sup>14</sup>. Dette vil drøftes videre i kapittel 8, men sammenhengen sies på populasjonsnivå å være statistisk/**probabilistisk**, hvilket betyr at den er deterministisk (det vil si bestemmer utfallet/resultatmålingen) for minst ett av (men som oftest flere av) enkeltindividene i populasjonen/pasientgruppen.

### 3.3 Måling av den overordnede prosessen

Det man i denne sammenhengen interesserer seg spesielt for, er ikke behandlings- og resultatmålene for enkeltpasienter, men måling av den overordnede prosessen. Hvordan fungerer prosessen for hele den pasientgruppen den skal skape verdi for? Fortsatt er det relevant å skille mellom **prosessmål**, som nå gjelder behandlingsmålene for samtlige pasienter i den gruppen prosessen er til for, og **resultatmål**, også

de for hele denne pasientgruppen. Ved valg av prosessmål finnes det en underliggende tanke om at disse målene kan indikere hvordan de endelige resultatmålene vil bli. Det er oftest først på et generelt prosessnivå at man kan si noe om det finnes en slik forbindelse mellom prosessaktiviteter (behandlinger osv.) og endelige resultater. Bli resultatene (statistisk) annerledes når en bestemt aktivitet utføres mer eller mindre bra, kanskje ikke er med i det hele tatt, eller er byttet ut med en annen aktivitet? Hvordan ser årsakssammenhengen ut – er det en forbindelse mellom prosessmålene og resultatmålene? Får ulike pasientgrupper ulik behandling ikke på grunn av selve sykdommens art men avhengig av faktorer som i sammenhengen burde betraktes som uvedkommende?

Når man for eksempel snakker om «gjennomsnittlig ventetid før operasjon» er dette en av mange indikatorer på hvor godt prosessen fungerer for en bestemt pasientgruppe med individuelle ventetider. Et mål på variasjon i gruppens ventetider er også en viktig indikator for hvor godt prosessen fungerer. Kanskje må man istedenfor å betrakte en middelvei se hele fordelingen, det vil si samtlige ventetider, for å få et bilde av prosessen (relevant for den aktuelle ventetiden) for denne pasientgruppen. Det kan jo være at visse pasienter får mye dårligere behandling enn andre. Det er ofte først når man ser helheten at det er mulig å danne seg en oppfatning om dette – og dermed få en indikasjon på forbedringspotensialet. En middelvei for hele pasientgruppen vil ikke gi noe informasjon om at ulike undergrupper av pasienter har forskjellige resultater dersom dette er tilfelle. Det innebærer at middelveiene må analyseres videre for å finne variasjon mellom undergrupper og kildene til denne variasjonen.



### 3.4 Måling over tid

At målingene foregår over tid, er spesielt viktig både på pasient- og organisasjonsnivå. Sykdomsforløp utvikler seg over tid, og i organisasjonen gir målinger over tid et bilde av hvordan man har forbedret seg. Er variasjonen man ser tilfeldig? Avtegner det seg mønstre? Har det hendt noe – positivt eller negativt – for resultatene? Disse og lignende spørsmål skal studeres nærmere i neste kapittel.

### 3.5 Måleproblematikk

All måling kan være beheftet med feil. Målinger kan være systematisk feil, det vil si at de ikke måler det de er ment å måle, eller at de har stor variasjon som ikke avhenger av det som måles. Mange slike feil oppstår på bakgrunn av feil i inndata og ikke-rapporterte data.

Dersom det forekommer direkte målefeil eller større tap av data, eller hvis innsamlede data ikke er adekvate for det aktuelle problemet, kan dette aldri kompenseres med avanserte beregnings- og problemløsningsmetoder. «Søppel inn gir søppel ut» er en noe flåsete men ofte treffende beskrivelse av situasjonen.

Et viktig aspekt ved innsamling av data er at de som medvirker til innsamlingen får klare tilbakemeldinger slik at de ser hvilken nytte man kan ha av data. Dersom det gis gode tilbakemeldinger og de som medvirker til innsamlingen innser nytten, reduseres risikoen for feil i forbindelse med datafangsten. Feil kan oppstå uansett hvor gode ambisjoner man har. Kanskje man venter med rapportering til en senere anledning da viktige aspekter kan ha forsvunnet, man kan ubevisst skrive feil eller kanskje glemme å rapportere. Når slike feil oppstår, er det sjelden bare personene som har slurvet; oftest er det systemfeil som gjør at datafangsten blir av dårligere kvalitet.

Det er også en tendens til å samle inn data som kan være «kjekt å ha», det vil si data som ingen egentlig vet hvordan de skal benytte eller om de i det hele tatt vil bli benyttet. Dette kan innebære at datahåndteringen blir tung og kostbar, og at man ikke får tid nok til data som uten tvil er verdifulle (en slags fortreningsseffekt). Datakvantiteten går ofte ut over kvaliteten på de innrapporterte data.

### 3.6 Oppsummering

Viktige kilder til variasjon er:

- pasienten,
- de individer, team og mikrosystemer<sup>15</sup> som leverer aktiviteter og behandling til pasienten,
- behandlingsmetoder, tiltak og innsats,
- integreringen av aktiviteter og delprosesser og hvordan de samvirker i prosessen,
- interaksjonen mellom ulike prosesser eller delsystemer som konkurrerer om ressurser,
- intervensjoner (endringer og forbedringer),
- måleprosessen, og
- hvordan det overordnede systemet og dets struktur påvirker prosessen.

Det er det overordnede systemet og dets struktur ledelsen kan påvirke. Om man ikke arbeider aktivt med forbedringer kan man regne med å få samme resultat som før. Om man gjør som man alltid har gjort, får man samme svar/resultater som man alltid har fått. Vil man ha et annet resultat, krever det forandring i systemet. Basert på diskusjoner med William Edwards Deming, uttrykker Paul Batalden<sup>16</sup> dette fenomenet som at «hvert system er perfekt utformet for å gi de resultater det faktisk har gitt».



## Kapittel 4

# Variasjon i et ledelsesperspektiv

Hvordan kan ledere på politisk og administrativt nivå se på den variasjonen som finnes i helsetjenester og kvalitetsregistre? Hva er det mulig å påvirke, og hva er det ikke mulig å påvirke? Det er slike spørsmål og drøfting av hvordan en eventuell påvirkning kan skje, som er av interesse i dette kapittelet.

### 4.1 Grunnleggende forutsetninger

Hva kan man fra et overordnet ledelsesnivå gjøre når man ser hvordan resultater fra helsesektoren genereres og varierer samt hvordan de identifiserte kildene til variasjon påvirker resultatene? Som allerede konstatert, kan en leder i prinsippet bare påvirke det overordnede systemets struktur. Man kan påvirke hvordan man i organisasjonen ser på arbeidet sitt og hvordan man ser på forbedringsarbeid<sup>17</sup>. For å kunne håndtere og drøfte variasjon på en måte som resulterer i forbedring, må visse grunnleggende forutsetninger være på plass.

#### 4.1.1 Visjon og mål

Et viktig punkt er å gi en **visjon** om en framtid som stimulerer til ulike typer forbedringsinnsats. Å formulere tydelige mål sammen med virksomheten – mål som medarbeiderne på ulike nivåer kan få mulighet til å påvirke og iden-

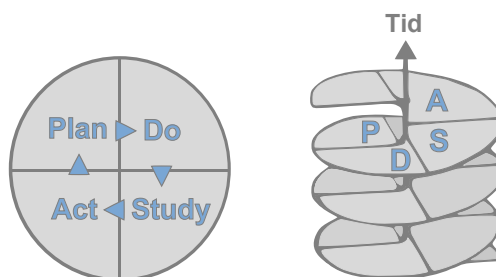
tifisere seg med – er viktig for at forbedringene skal få en retning. Figur 4.1 illustrerer aspekter ved forbedringsarbeidet som en leder kan påvirke; den såkalte **forbedringssirkelen** drives av organisasjonens visjon og strategien for å nå denne visjonen.

Som allerede understreket, er det ofte lett å bli enige om *hva*, vanskelighetene oppstår når man skal bli enige om *hvordan* man skal komme dit. Samtaler omkring vurderinger og normer kan gi en god pekepinn på dette. Slike samtaler kan gjerne ta utgangspunkt i de temaer som ble belyst i kapittel 3.

Det er viktig at man innad i en organisasjon sammen skaper en tydelig visjon for hvor organisasjonen er på vei, en visjon som kan engasjere både medarbeidere og ledere. At ledelsen virkelig vil noe med visjonen, bør komme tydelig fram, og strategier for framdrift må skapes i fellesskap. Disse strategiene bør omfatte en infrastruktur som støtter opp under forbedringsarbeidet. Det er også viktig at ledelsen viser vei og støtter et kreativt forbedringsarbeid.

#### 4.1.2 Forbedringsstrategi

En forbedringsstrategi inneholder en felles tilnærming; fra hvordan man definerer problemer og måler til hvordan man driver selve forbedringsarbeidet i tråd med den nevnte forbedringssirkelen som gjerne kalles Demings PDSA-sirkel<sup>18</sup>, hvor akronymet står for «Plan-Do-



**Figur 4.1:** Forbedringssirkelen drives av organisasjonens visjon og strategien for å nå denne visjonen. Betraktes tidsdimensjonen blir sirkelen en spiral.

Study-Act», Denne lærings sirkelen både kan og bør tolkes på to ulike nivåer:

1. For hver aktivitet har man en gang planlagt (mer eller mindre systematisk) hvordan den skal utføres. Når man så utfører det som er planlagt, må man følge opp resultatene. Ble det som man hadde tenkt, hva kan man lære av resultatene, og er det grunn til å prøve å forbedre resultatene? Når man ut fra innsamlede fakta ser behov for forbedring, kommer man til den andre tolkningen av forbedringssirkelen.
2. Etter å ha analysert fakta rundt et identifisert problem eller en identifisert mulighet skal man planlegge en løsning («Plan»). Planen skal deretter gjennomføres («Do») før resultatet vurderes («Study»). Dersom resultatet er godt, tar man lærdom av dette og iverksetter tiltak («Act») som gjør endringen permanent. Dersom planen som ble gjennomført, til tross for at man brukte all sin innsikt, ikke fungerte, er det også mye å lære. Med den nyervervede innsikten må problemet eller muligheten analyseres på nytt, man må reflektere over nye innfallsvinkler og eventuelt skape nye tankemodeller for å kunne planlegge en ny løsning. PDSA-sirkelen for forbedringsprosjekter er oftest en iterativ prosess som gjentas flere ganger.

Det finnes også andre tilnæringer til den andre typen bruk av PDSA-sirkelen enn dens

egne prosess. Ved å observere omgivelsene aktivt, sammenligne seg med andre og lese ny forskningslitteratur kan man oppdage nye muligheter som det kan være god grunn til å prøve ut. Man kan begynne i hvilken del man måtte ønske av lærings sirkelen, som i realiteten er en spiral når man inkluderer tidsdimensjonen som illustrert til høyre i Figur 4.1. Trinnet som er benevnt «Act» kalles ofte for «Learn». I dette legger man to betydninger: den ene er å lære samt reflektere over hva man har sett, og den andre er å ta lærdom av dette i den betydning at det faktisk skal påvirke framtidige handlinger. Planlegging kan omfatte aktiv utprøving. Det er organisasjonens utvikling som er i fokus.

Det finnes et annet aspekt ved betydningen av å vektlegge forbedring, dvs. organisasjonens utvikling langs tidsaksen, fremfor et nivå eller rangering på et enkelt tidspunkt. «Kreativ bokføring» kan få nivået til å se bra ut, år etter år. Ved et ensidig fokus på ventetider kan en organisasjon for eksempel flytte på pasienter og gjøre omrokkeringer i køen, og dermed fremstå på et høyere nivå uten at organisasjonen har blitt noe bedre. Dette gjør det vanskeligere å kunne vise til vedvarende **forbedringer**. Man burde altså ha mer fokus på forbedringstakten – helst over en lengre tidsperiode – og mindre fokus på nivåer og rangeringer. Den relevante sammenligningen for en organisasjon er ikke nødvendigvis andre lignende organisasjoner, den kan vel så gjerne være organisasjonen selv på et tidligere tidspunkt.

Inspirert av bevegelsen i retning av mer evidensbasert medisin som har påvirket helsevesenet de siste tiårene, har forskere innen organisasjonsteori og ledelse forsøkt å reflektere over hva det tilsvarende ville bli for ledelse av organisasjoner. Når det gjelder organisasjoners utvikling, er det ikke like «lett» som for medisinske resultater å gjøre vitenskapelige forsøk av typen randomiserte kontrollerte studier. Alle organisasjoner er unike, og deres utvikling og funksjon er ikke bestemt av en underliggende genetikk<sup>19</sup>, men snarere av deres utvikling over tid<sup>20</sup>. Dette forhindrer ikke at fakta om hvordan det går i organisasjonen fremdeles er veldig viktig. Hvordan lykkes organisasjonen i å skape verdi for de innbyggerne og pasienter den er til for? Har innsatsen som er gjort for å skape enda bedre verdi, lykkes? Med andre ord: har forbedringsarbeidet lykkes? Her kan offentliggjøring fra medisinske kvalitetsregistre by på et faktagrunnlag som tidligere ikke var like tilgjengelig, om det i det hele tatt fantes.

## 4.2 Illustrasjon av prosessforbedringer

Forbedringsarbeidet rettes ikke utelukkende mot aktivitetene i prosessen – viktige aspekter er hvordan prosessene faktisk defineres og formuleres samt hvordan de ulike aktivitetene i prosessen organiseres. I begynnelsen av forrige kapittel ble den forebyggende delen omtalt. Illustrasjonen var en individualisert screeningprosedyre i begynnelsen av en brystkreftprosess hvor vi også gikk inn på hvordan prosessen omfatter de siste stadiene. Generelt er det mye som kan gjøres med hensyn til å la prosesser omfatte også forebyggende og helsefremmende tiltak. På visse områder har man kommet et godt stykke på vei, for eksempel med kampanjer mot røyking.

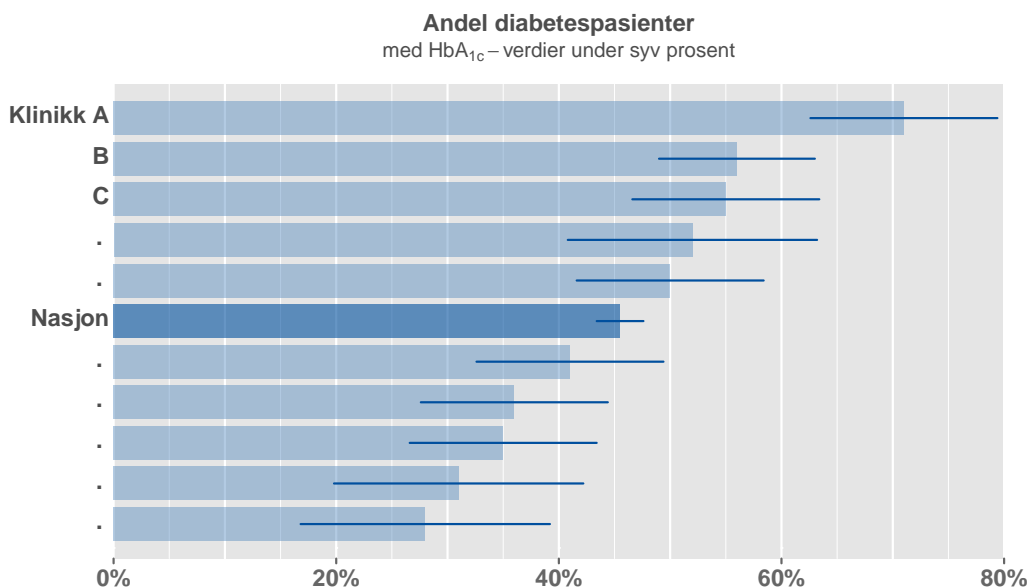
Det er ikke bare prosessenes omfang man kan fundere over. Ledelsen har også mulighet til å påvirke organiseringen av prosessenes aktiviteter og ressursbruk. I Skåne i Sverige har det vært gjennomført forsøk med en brystkreftprosess basert på konseptet «one-stop-shop». Hele kjeden, fra mammografi til biopsi av mistenkte svulster og laboratorieundersøkelser av vevsprøver, kan gjøres som en kontinuerlig strøm på ett og samme sted<sup>21</sup>.

Dette er et eksempel på hvordan man går bort fra den tradisjonelle organiseringen i helsevesenet, nettopp for å sikre optimal utnyttelse av ressurser som inngår i prosessen. I konvensjonelle behandlingsskjeder får de berørte kvinnene vente i et muligens langdrygt køsystem til ressursene blir tilgjengelige. Med en løsning å la «one-stop-shop» organiseres ressursene omkring kvinnene som skal behandles og flyten blir mer effektiv. Av ulike grunner er imidlertid dette forbedringsprosjektet nå avsluttet, men tilsvarende forsøk med samme konsept gjennomføres nå flere steder.

Et eksempel er helsetjenesten i Västra Skaraborg (Sverige ligger generelt langt fremme på området), der et team besøker multisyke eldre for å håndtere hele deres sykdoms- og livssituasjon i stedet for at pasientene må reise fram og tilbake til sykehus for å besøke ulike spesialistfunksjoner. Mye lidelse blir unngått samtidig som helsevesenet sparer ressurser på dette<sup>22</sup>.

Tradisjonelt har man i mange virksomheter, så effektivt som mulig, prøvd å utnytte hver enkelt av de ressursene som trengs i prosessen hver for seg. Dette fører imidlertid som oftest til suboptimalisering og spesielt dersom det finnes variasjon i systemet, dannes det køer og lange ventetider. Totalt og generelt blir systemet ikke spesielt effektivt. Systemer som derimot utvikles med henblikk på flyteeffektivitet og kundefokus, kan mye lettere effektiviseres for de prosessen er til for – nemlig pasientene.

Det har vært mange såkalte gjennombrudds-



**Figur 4.2:** Andelen barnediabetikere med HbA<sub>1c</sub>-verdi under syv prosent for et utvalg fiktive diabetesklinikker. Klinikk A har et avvikende resultat som gir grunn til å studere klinikkens rutiner og behandlingstilbud relativt andre.

prosjekter i den svenske helsetjenesten som kan gi inspirasjon til forbedringsarbeid, spesielt kan Nationella Diabetesregistrets forbedringsgrupper nevnes, som med tydelig utgangspunkt i et kvalitetsregister har iverksatt forbillig forbedringsarbeid i nasjonal skala.

### 4.3 Illustrasjon av læring fra kvalitetsregistre

I fortsettelsen illustreres det hvordan en læringsprosess basert på medisinske kvalitetsregistre kan se ut. Fra et perspektiv utenfra trekkes resonnementet så langt man kan med «variasjonsbriller» på. For å gå videre mot et reelt forbedringsarbeid kreves mer spesifikke kunnskaper på det aktuelle området. Dette må derfor bare betraktes som et mulig utgangspunkt, ikke som en fasit eller konklusjon.

Figur 4.2 illustrerer variasjonen mellom behandlingsinstitusjoner, i dette tilfellet fiktive klinikker for behandling av barnediabetes. Andelen barnepasienter som tilfredsstillt kriteriet om en målt HbA<sub>1c</sub>-verdi under syv prosent,

bør være så høy som mulig. Man kan med god grunn gå ut fra at de fleste som interesserer seg for barnediabetes ville blitt nysgjerrig på hva det er som gjør at verdiene fra en av klinikkene er så avvikende. Av figuren ser vi at den relevante variasjonen høyst sannsynlig ikke bare er et resultat av tilfeldig eller naturlig variasjon. Andelenes/klinikkenes konfidensintervaller (horisontal linje på søylene) sier noe om usikkerheten i målingene – noe vi vil komme tilbake til. For illustrasjonens skyld antas det at forskjellen mellom to klinikker kan vises å være statistisk signifikant når klinikkenes to konfidensintervaller ikke overlapper hverandre langs den horisontale akse. At en forskjell er statistisk signifikant kan her (noe forenklet) tolkes til å bety at forskjellen er større enn det man kan forvente av tilfeldigheter alene. Det er dog ikke alltid like lett som her å skille relevante variasjonskomponenter fra en naturlig eller tilfeldig variasjon.

### 4.3.1 Er variasjonen naturlig?

Det er viktig å stille spørsmål om det er sannsynlig at resultatene kun er et uttrykk for en **tilfeldig variasjon** før man trekker slutninger fra data. I dette tilfellet er det relativt enkelt å konstatere at resultatene neppe er et sannsynlig utfall av naturlig variasjon. Dette ses av konfidensintervallene (som gir en indikasjon på den tilfeldige usikkerheten), hvor intervallet for Klinik A ligger klart over nesten alle andre. Det er altså neppe tilfeldig variasjon som gjør at Klinik A ligger så godt an.

Før man konkluderer med at Klinik A er dyktig må man undersøke om det avvikende resultatet kan være resultatet av ulike operative definisjoner. Noen klinikker kan for eksempel ha utelatt nydiagnostiserte diabetikere eller kun inkludert HbA<sub>1c</sub>-verdiene for de polikliniske pasientene. For illustrasjonens skyld antas det videre at de operative definisjonene er de samme ved alle klinikker.

Ved hjelp av data fra det aktuelle kvalitetsregisteret er det da faktisk mulig å jakte på Klinik As framgangsfaktorer. Figur 4.3 viser Klinik As gjennomsnittlige HbA<sub>1c</sub>-verdier for årene 2000 til 2010 relativt det nasjonale gjennomsnittet. Et positivt avvik (altså større enn null) angir et årlig snitt over den nasjonale, mens et negativt angir årlig snitt under det nasjonale; for eksempel ligger Klinik A to prosentpoeng under det nasjonale gjennomsnittet i år 2000. Man kan uten særlig risiko for feilkonklusjon konstatere at det skjedde en eller annen forandring rundt 2006. Introduserte klinikken da nye behandlingsmetoder som andre klinikker umiddelbart bør kopiere?

Data fra medisinske kvalitetsregistre kan gi indikasjoner på mulige forskjeller som signaliserer forbedringsmuligheter, men før et forbedringsforslag foreligger må man gjennom en søkeprosess; hva er det de gjør ved Klinik A? Man må først eliminere risikoen for at man be-

nytter ulike operative definisjoner eller ulike inklusjonskriterier for pasientene. De identifiserte forskjellene mellom Klinik A og de øvrige klinikkene er kun en lovende start på den videre prosessen – det er fortsatt mye arbeid som gjenstår før man kan iverksette eventuelle tiltak for å heve nivået til de øvrige klinikkene.

## 4.4 Å håndtere variasjon

Det er tre aspekter som er spesielt viktig å reflektere over i forbindelse med tolkningen av resultater fra medisinske kvalitetsregistre:

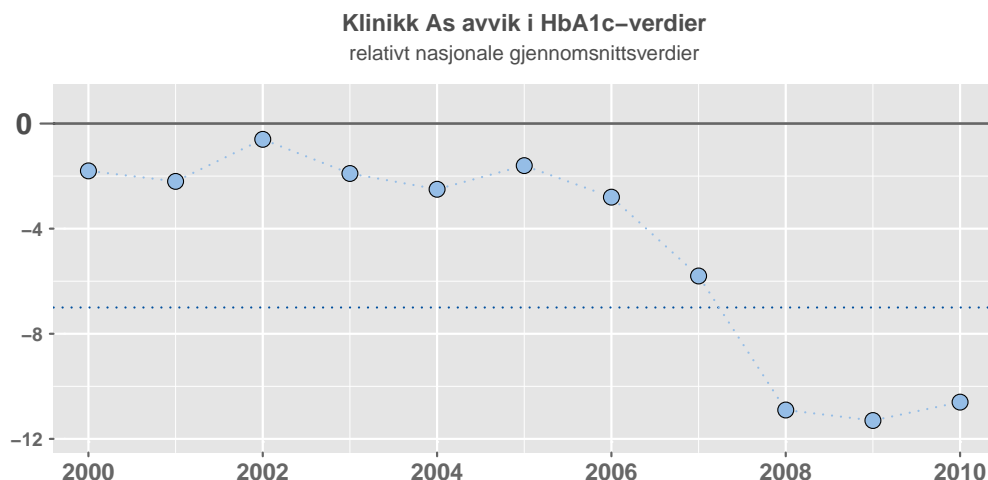
1. Finnes det andre organisasjoner, som er sammenlignbare, som er betydelig bedre enn oss – bortsett fra det som kan forklares med tilfeldig variasjon og en annen case-mix<sup>2</sup>?

I slike tilfeller er det all grunn til å undersøke andres tilsvarende prosesser – gjør de ting på en annen måte som man kan lære av? Kan man med nødvendige tilpasninger til de aktuelle forholdene gjøre noe tilsvarende? Still spørsmål til egen organisasjon, men husk at det er persone som eier prosessen og som har prosesskunnskapene som må trekke konklusjonene.

Selv der det ikke er mulig å trekke direkte lærdom av andre (for eksempel på grunn av en annen case-mix<sup>2</sup> eller ulik organisering) kan resultatene indikere at det er mulig å bli bedre; er det for eksempel mulig å stimulere prosesssteamene for den aktuelle prosessen til å innlede et kreativt forbedringsarbeid?

2. Har en annen organisasjon, utover en effekt av tilfeldig variasjon, forbedret seg mer enn oss? (Case-mix har ikke samme avgjørende betydning.)

Hvordan organiserer og systematiserer andre sitt forbedrings- og utviklingsarbeid, og kan



**Figur 4.3:** Klinik As gjennomsnittlige HbA<sub>1c</sub>-verdier for årene 2000 til 2010 relativt de årlige nasjonale gjennomsnittene langs nullinjen. Et negativt avvik betyr at HbA<sub>1c</sub>-snittet til klinikken er bedre/lavere enn landsgjennomsnittet. Striplet horisontal linje angir en kontrollgrense (eller et konfidensintervall) som indikerer at nedgangen sannsynligvis ikke skyldes tilfeldigheter.

man lære noe av dem? Kan man bli inspirert til egen kreativ organisering av forbedringsarbeidet ut fra de aktuelle forutsetningene?

3. Utmerker våre resultater seg når det gjelder absolutt nivå eller forbedringstakt?

Det er viktig å få oppmerksomhet også for de resultatene man har oppnådd. Altfor mange forteller at de «virkelig får høre det om noe har gått galt, men ingen oppmerksomhet når noe er veldig bra». Man må se og stimulere de som gjør en god innsats. Dette påvirker ikke bare de som får positiv oppmerksomhet, men også strukturene (tankesettet og forståelsen for hva som egentlig er viktig) i hele organisasjonen. Strukturene påvirker i sin tur prosessene og deres framtidige resultater.

Viktige aspekter er altså å gi oppmerksomhet og å stille spørsmål – men alle spørsmål er ikke like gode. Det er åpne – ikke dømmende og lukkede – spørsmål som kan få mennesker til å reflektere og sette i gang kreative prosesser. Åpne spørsmål kan resultere i prosesser og resultater spørsmålsstilleren aldri hadde

tenkt på. Dette er spesielt viktig i organisasjoner i helsetjenestene som har mange kompetente medarbeidere fra ulike yrkesgrupper. Med stimulerende spørsmål kan man komme veldig langt. Offentliggjøring fra medisinske kvalitetsregister skal betraktes som et signalsystem – en mulighet til å skape en meningsfull dialog mellom den overordnede ledelsen og medarbeiderne. Mye kan oppnås gjennom stimulerende spørsmål som viser at man virkelig bryr seg. Et relevant spørsmål er for eksempel: «Finnes det andre organisasjoner som er bedre enn oss eller som har forbedret seg mer enn oss, som vi kan lære av?»



## Kapittel 5

# Fenomenet variasjon

Fenomenet **variasjon** er kjent, men ofte dårlig forstått og ikke sjelden feid under tepet. Begrepet er problematisk på samme måte som tidsbegrepet; alle vet hva det er, men det er vanskelig å definere. I dette kapitlet utredes variasjonsbegrepet samtidig som grunnen legges for en dypere forståelse av hvordan man fra ulike ståsteder kan håndtere variasjonen i medisinske kvalitetsregistre samt bruke denne forståelsen som et utgangspunkt for kontinuerlig utvikling- og forbedringssarbeid.

### 5.1 Variasjon

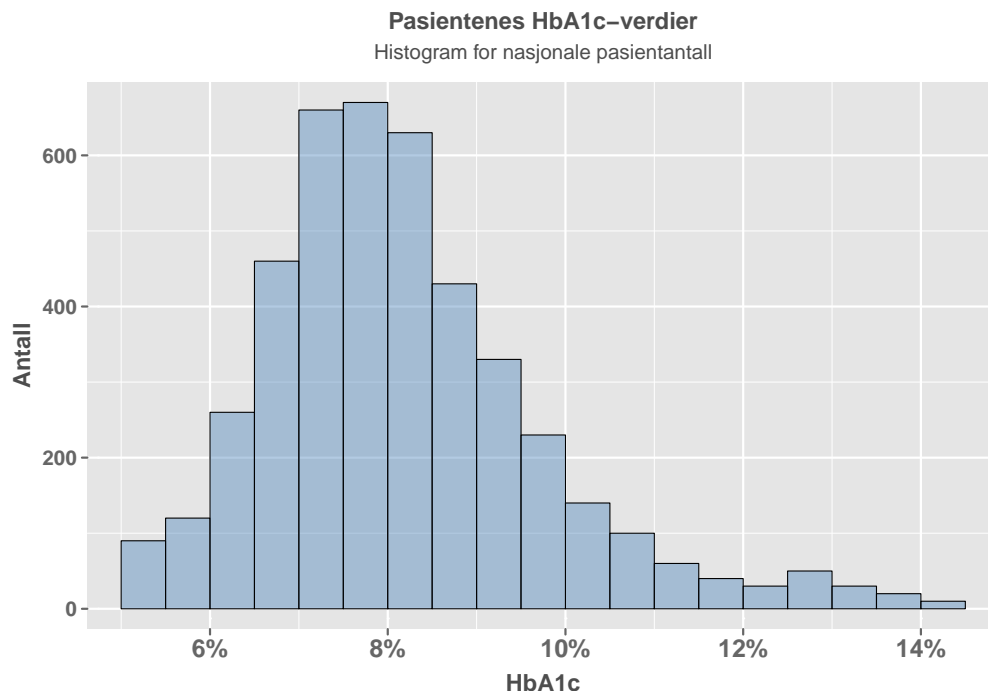
Variasjon er en grunnleggende egenskap ved verden slik den forstås i dag – fra Big Bang, partikkelfysikk, kjemi og biologisk evolusjon til språk, kunnskapsutvikling og sosiale systemer. Ikke minst i helsesektoren er det viktig å forstå variasjon. Alle pasienter er ulike; sykdomsforløp varierer; måleresultater varierer – både over tid for den enkelte pasient og mellom pasienter. Variasjonen avhenger også av ulike målemetoder og hvor godt de kan gjengi det de er ment å måle (**validitet**). Behandlingspersonalet tolker situasjoner forskjellig, og selv etter å ha tolket to situasjoner som hovedsakelig like, handler de forskjellig; samme behandling gir forskjellig effekt hos forskjellige pasienter.

Man kan se for seg kilder til variasjon, som hele tiden er aktive og skaper den variasjon

man kan observere – blant annet i medisinske kvalitetsregistre. Figur 5.1 illustrerer hvordan HbA<sub>1c</sub>-verdiene for pasienter med diabetes type 1 kan variere ved en diabetesklinikk. Gjennomsnittsverdien ligger rundt 65 mmol/mol, men det er stor variasjon i de individuelle målingene rundt denne verdien. En **gjennomsnittsverdi** kan være informativ, men enkelte pasienter, kanskje til og med ganske mange pasienter, kan ha en betydelig dårligere verdi. Dette er det viktig å ha i mente når man vurderer gjennomsnittsverdier.

#### 5.1.1 Synkron og diakron variasjon

Figur 5.2 illustrerer et eksempel på pasientvariasjon, i antall dager på venteliste for de fiktive klinikkens pasienter, i forhold til variasjonen mellom forskjellige behandlingstidspunkter. Denne variasjonen kalles **synkron variasjon**<sup>23</sup> og beskriver hvordan enheter varierer i forhold til hverandre på samme tidspunkt eller i samme tidsperiode. Betydningen av tidsdimensjonen har allerede vært diskutert; forbedring, det vil si en positiv forandring over tid, er vel så viktig som absolutte nivåer. I Figur 4.3 varierte resultatene for Klinik A for blodsukker verdi hos barn over tid, og man kunne konkludere med at noe hadde hendt rundt år 2006 da en forbedring syntes å ha inntruffet. Variasjon som forekommer langs tidsaksen kalles gjerne **diakron variasjon**.



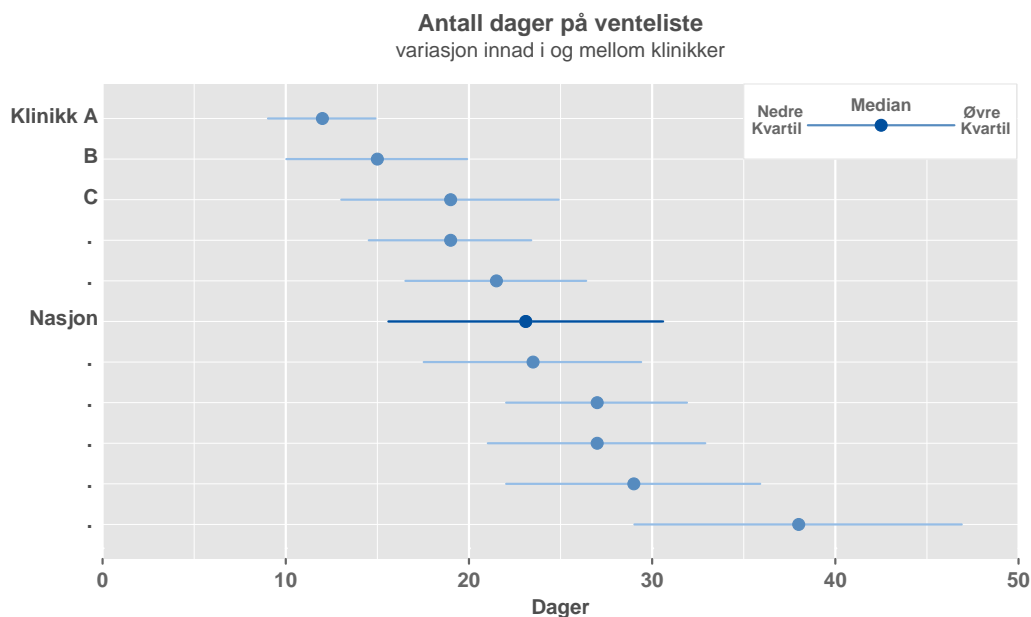
**Figur 5.1:** Variasjonen i HbA<sub>1c</sub>-verdiene for diabetespasienter er her illustreres med et histogram hvor de ulike rektanglene tilsvarer andelen pasienter i hvert intervall.

Som nevnt tidligere bør ikke variasjon mellom behandlingsinstitusjoner kun sees i lys av hvor sikker man kan være på en middelverdi eller medianverdi, man bør også vurdere spredningen innenfor den enkelte institusjon – det vil si variasjon mellom pasienter i samme behandlingsinstitusjon. Kanskje har noen behandlingsinstitusjoner bedre kontroll på sin prosess og dermed mindre spredning mellom pasientene. I så fall kan andre institusjoner kanskje lære noe av denne institusjonen. I Figur 5.2 kan man for hver klinikk, i tillegg til klinikkenes medianverdier (for antall dager på venteliste for en behandling), se spredningen innad i hver klinikk. Man ser for eksempel at Klinikk A ikke bare har en god medianverdi, de har også liten spredning blant pasientene med tanke på hvor lenge de må stå på venteliste for en viss type behandling.

Selv om det altså er nesten utelukkende den synkrona variasjonen, variasjonen mellom behandlingsinstitusjoner, som offentliggjøres fra

medisinske kvalitetsregistre, er det faktisk den diakrone variasjonen, variasjonen over tid, som i de fleste tilfeller er mest interessant. Ved å se bakover i tid på variasjon i resultater fra en prosess, vil man kunne uttale seg om framtidige resultater fra samme prosess. Det å bare vite hvordan det har gått tidligere er ikke så interessant om det ikke skulle kunne brukes til å si noe om hvordan resultatene kan bli i framtiden. Da er det faktisk den diakrone variasjonen som er interessant – har man for eksempel en positiv utvikling eller er det bare tilfeldige variasjoner man ser? Figur 4.3 ga en indikasjon på at det rundt år 2006 synes å ha vært iverksatt et vellykket tiltak, og man kan rette oppmerksomheten mot prosessforbedringer basert på denne informasjonen; hva var det man begynte å gjøre annerledes på denne tiden?

Figur 5.3 viser antall krybbedødsfall (SIDS eller plutselig spedbarnsdødsfall) i Sverige over tid, dvs. diakron variasjon. Her får man indikasjon på at tiltaket som ble i verkstatt i



**Figur 5.2:** Illustrasjon av spredningsinformasjonen for hver behandlingstjenestestasjon samt variasjonen mellom disse institusjonene. Midtpunkt angir median for antall dager på venteliste, mens venstre og høyre ytterpunkt angir første og tredje kvartil.

1992, å spre informasjon til nybakte foreldre (blant annet om ryggeleie), synes å ha gitt et godt resultat<sup>24</sup>. Absolutt sikre kan man i utgangspunktet ikke være, det kan ha vært noe annet som inntraff på samme tidspunkt og som gjorde at barnedødeligheten falt i hele landet. Dersom man etter grundige overveielser ikke finner andre relevante faktorer, kan det være en rimelig antakelse å tro at tiltaket hadde effekt. De horisontale linjene i figuren, som beskriver området eller antatte grenseverdier for tilfeldig variasjon, skal vi komme tilbake til. I de kommende avsnitt fortsetter drøftingen av tolkning av diakron variasjon.

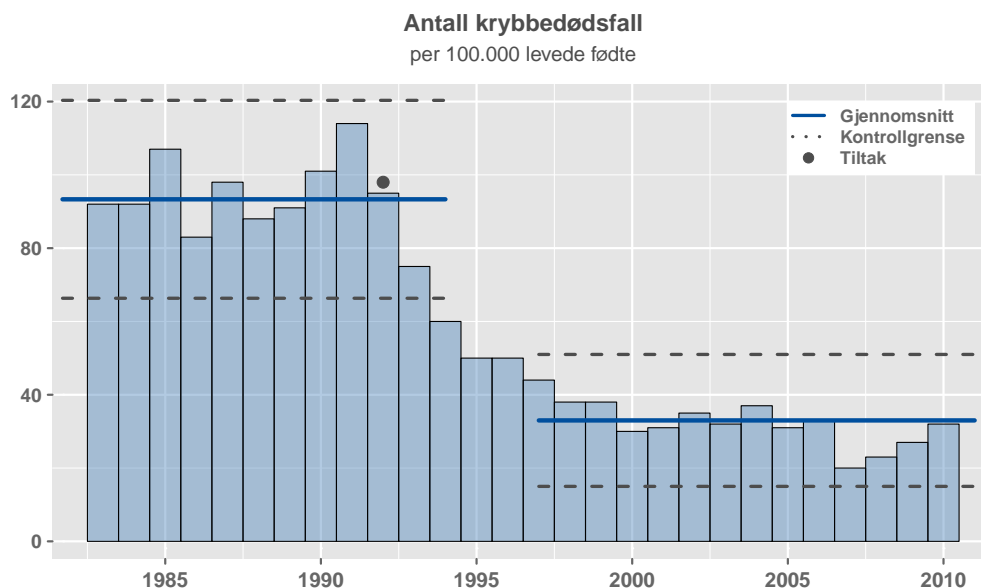
## 5.2 To typer studier

At man har samlet inn data og vil studere dem, skyldes rimeligvis at man vil anvende de til å fatte beslutninger av et eller annet slag – nå eller i framtiden. Hvilken type beslutning det gjelder bestemmer hvordan man bør se på variasjonen i data. Studier av data kan deles inn

i to typer; enten skal man ta en beslutning basert på resultater fra en viss tidsperiode eller så skal data brukes til å forutsi framtidige resultater. Her inkluderes også det å kunne bestemme om et tiltak har gitt ønsket resultat slik at den framtidige prosessen kan antas å ha blitt bedre enn den var før tiltaket ble iverksatt.

For å predikere er det nødvendig å anta at samme system av årsaker til variasjon kommer til å være gjeldende også i framtiden, og at dette systemet er noenlunde stabilt (det vil si at ingen enkelt variasjonskilde er dominerende). Historiske data blir dermed relevante når man prøver å forutsi hvilke data som kommer til å bli generert i framtiden, og hvilken variasjon disse data kommer til å ha. Den første typen studier kalles gjerne **beskrivende studier**, og den andre typen **analytiske studier**.

Dersom man fokuserer på en viss periode og vil benytte data fra denne, kan det typisk utføres gjennom en beskrivende studie. Dersom man da for eksempel tar med alle operasjoner av en bestemt type i den spesifikke perioden



**Figur 5.3:** Antall krybbedødsfall i Sverige. I 1992 begynte Sosialstyrelsen å dele ut brosjyre om krybbedød til alle nybakte foreldre. Der foreskrives blant annet ryggeleie for første gang, noe som anses å ha bidratt til en dramatisk reduksjon i antall dødsfall.

vil man kunne få en rimelig god oversikt over perioden, og det er unødvendig å uttale seg om hva som vil skje i framtiden. Man trenger ikke spekulere i årsakene til at det har blitt utført et visst antall operasjoner. Dersom formålet er å skape underlag for betaling for det som er oppnådd, såkalt «Pay for performance»<sup>25</sup>, er det nok med en enkel beskrivelse.

For teamet som har utført operasjonene og for teamets ledelse er situasjonen kanskje annerledes. Hva sier historiske data om det framtidige resultatet – hvordan ser årsakene til variasjonen ut? Den samme underliggende prosessen som har skapt resultater historisk vil skape resultater i framtiden. Derfor burde historiske resultater være indikative for framtidige resultater fra samme prosess. Hvilke resultater vil behandlingsprosesser, køer, skjemastrukturer og planleggingsgrunnlag gi i framtiden dersom ingenting forandres? Kan det antas at årsakssystemet, det vil si systemet av kilder til variasjon, vil være noenlunde stabilt?

Kanskje finnes det store/dominerende, gjerne kalt spesielle<sup>26</sup>, kilder til variasjon hvis virkning som ikke kan forutsies, og som man kanskje burde prøve å eliminere eller gjøre noe med? Det kan være en ny medarbeider som har fått utilfredsstillende utdanning, en ny målemetode som gjør historiske data urelevante, en inadekvat behandlingsmetode som er introdusert, osv. Dersom det ikke er funnet slike spesielle kilder til variasjon, må man kanskje vurdere å endre hele systemet slik at resultatene blir bedre i framtiden.

I blant kan en dominerende kilde til variasjon være ønskelig; det kan, som for krybbedød i Figur 5.3, dreie seg om et iverksatt tiltak – spørsmålet er om dette faktisk har gitt en spesiell forbedring i den aktuelle prosessen. For spedbarnsdødelighet iverksatte Sosialstyrelsen i Sverige et spesifikt tiltak som ble fulgt av en ønsket forbedring og en stabil prosess på et langt bedre nivå. Lignende tiltak ble også gjennomført i Norge med tilsvarende resultater<sup>27</sup>.

For at man i noen grad skal kunne forutsi framtiden fra historiske data, er det nødvendig å gjøre en antagelse om at systemet av variasjonsårsaker også i framtiden vil generere den type variasjon som er historisk observert. Som allerede nevnt er «dagens system er perfekt utformet for å gi de resultater som er observert». Men det er en sannhet med modifikasjoner. Hvor forutsigbare disse resultatene er, kan bare bedømmes ved hjelp av de data som er observert – ser det ut som systemet av årsaker gjen-skapes over tid? Dette er temaet i neste avsnitt.

### 5.3 En analytisk tilnærming

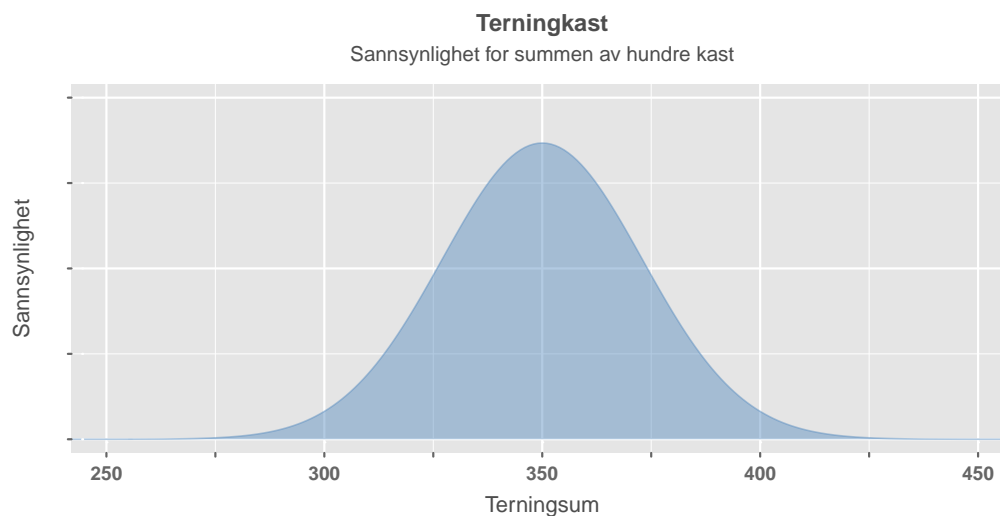
Begrepet «analytisk» kommer av at man skaper en modell av hvordan variasjonen kan beskrives statistisk ved hjelp av analyse av historiske data, og gjør antakelsen om at modellen vil gjelde for framtidige resultater. Dersom det kommer inn spesielle årsaker til variasjon på en ukjent eller uforutsigbar måte, vil ikke dette fungere. Man forutsetter altså at variasjonen er tilfeldig og at det er de samme tilfeldighetsmekanismene som produserer data. Om en kilde til variasjon skal regnes som spesiell eller ikke, er avhengig av sammenhengen. Dersom dens bidrag i forhold til annen variasjon er lite, trenger man sjelden å ta spesielle hensyn til den – den kan betraktes som tilfeldig<sup>28</sup>.

Ledelsen av en organisasjon bør fokusere på framtidige resultater og ikke betrakte resultatene som allerede er oppnådd. Det er kun de fremtidige resultatene man kan påvirke. Historiske resultater er et grunnlag for å beskrive framtiden dersom intet vesentlig endres. Ledelsen bør ta utgangspunkt i en analytisk tilnærming, analysere de historiske observasjoner og – med en viss variasjon – prøve å forutsi framtiden og iverksette tiltak for å oppnå bedre framtidige resultater?<sup>29</sup> Neste avsnitt vil utdype forståelsen av spesiell og tilfeldig variasjon.

### 5.4 To typer årsaker til variasjon

Som allerede indikert, finnes det to hovedtyper av variasjon; **tilfeldig variasjon**, som ofte skyldes mange små og noenlunde innbyrdes uavhengige bidrag uten at noen enkelt variasjonskilde dominerer, og **spesiell variasjon** som kan gi dominerende og i visse tilfeller uforutsigbare bidrag. Kaster man en vanlig, men perfekt sekskantet terning hundre ganger, vil summen av alle resultatene bli et tall mellom 100 og 600. Riktignok kan man ikke si nøyaktig hva resultatet vil bli, men det er for eksempel svært usannsynlig<sup>30</sup> at summen blir 100, altså at terningen hver eneste gang, ett hundre ganger, lander med siden som har ett øye opp. Det finnes mange små årsaker som bidrar til det resultatet man får i hvert enkelt kast og siden til sluttsummen. Dette er en veldig enkel illustrasjon av det man kan kalle tilfeldig variasjon. I dette tilfellet forstår man godt hvordan systemet av årsaker påvirker det enkelte terningkast i den forstand at man kan gi hver og en av de seks sidene lik sannsynlighet for å vende opp. Man kan da matematisk beregne sannsynligheten for alle de mulige utfallene av kastenes sluttsum. Dette er illustrert i Figur 5.4. Summen til de fleste seriene med hundre kast kommer til å ligge rundt sannsynlighetsfordelingens midtpunkt rundt 350, og nesten alle serier vil ende opp med en sluttsum i intervallet fra 300 til 400.

Man kan med relativ stor sikkerhet gjøre denne typen forutsigelser så lenge systemet av årsaker, alle de uendelig mange små årsakene til terningens bevegelser, er stabilt, det vil si at terningen er intakt og forholdene eller rammebetingelsene er stabile. Men hvis terningen sprakk eller biter begynte å falle av, er ikke systemet av årsaker lenger stabilt. Kanskje ville det bli større variasjon i resultatene, og det ville ikke lenger være like lett å forutsi hva som

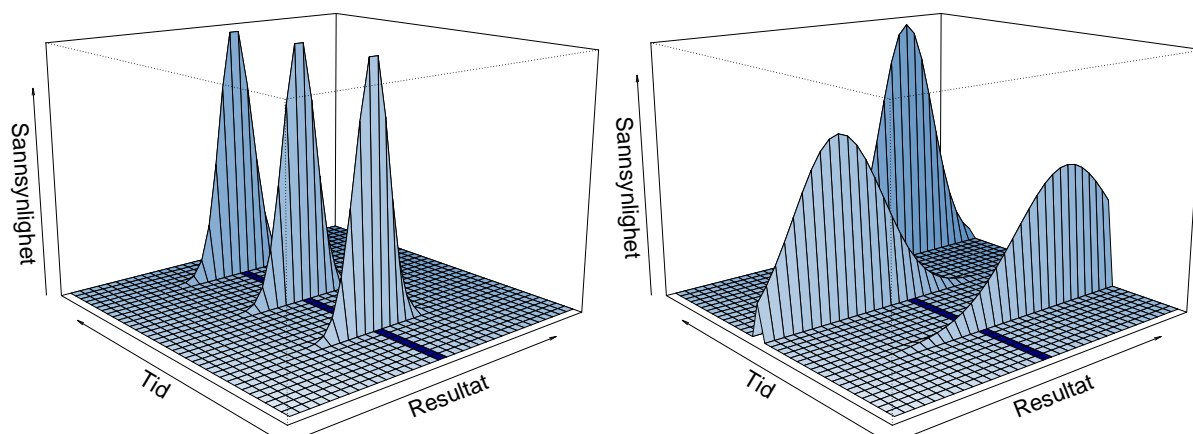


**Figur 5.4:** Sannsynlighetsfordeling for summen av hundre terningkast. Kurven viser for eksempel at det er størst sannsynlighet for å få en sum rundt 350 øyne og liten sannsynlighet for å få en sum over 400.

vil skje i neste runde. En spesiell årsak til variasjon har kommet inn og forstyrret det naturlige systemet av årsaker. En årsak til variasjon har blitt dominerende, den er ukjent, og effekten av den lar seg vanskelig forutsi. Dette er en **spesiell** kilde til variasjon. En slik kilde kan påvirke nivået og variasjonen i resultatene på en uforutsigbar måte.

På samme måte varierer resultatene i en behandlingsprosess, som illustrert av diabetespasientene i Figur 5.1. Det finnes mange årsaker til variasjon som er stabile og naturlige, men iblant inntreffer det hendelser som forstyrrer prosessen. Det kan for eksempel være en ny medarbeider som ikke kjenner rutinene og dermed skaper kaos i behandlingsprosessen, eller noen som til tross for sin erfaring gjør feil<sup>31</sup>. Det kan også være en ny behandlingsmetode som er innført eller en annen case-mix<sup>2</sup> som er aktuell. Historiske data blir ikke lenger like relevante, annet enn som grunnlag for å avgjøre om for eksempel en ny behandlingsmetode er bedre enn den gamle. Stort sett kan man forutsi resultatet av behandlingen, men iblant kommer det inn spesielle årsaker til variasjon som forstyrrer.

Figur 5.5 illustrerer hvordan to ulike prosesser varierer på tre ulike tidspunkter langs tidsaksen. Den venstre prosessen er stabil i den betydning at resultatet (eller utfallet) ved de tre tidspunktene er sentrert rundt samme middelværdi. Videre er også variansen (bredden på sannsynlighetens fordelingskurver) ved hvert tidspunkt relativt liten og konstant. I den høyre prosessen er ikke kurvenes midtpunkt, det vil si forventningsverdien (eller forventet utfall) for resultatet, den samme ved de tre tidspunktene – prosessen er ikke stabil. Videre er variasjonen i resultatene, i tillegg til å være relativt stor, svært ulik fra ett tidspunkt til et annet. Dersom det var operasjonstider disse tre prosessene illustrerte, ville den til høyre prosessen kanskje illustrere hvordan ulike operasjonsteam har forskjellige måter å handle på – både middelværdiene og spredningen blir forskjellig. Det er vanskelig å forutsi hva som vil skje framover dersom man ikke vet nøyaktig hvilket operasjonsteam det er snakk om. I den venstre prosessen er det mye lettere å forutsi hvilke resultater og hvilken variasjon man kommer til å få i fremtiden. I den høyre prosessen har man bedre «kontroll på situasjonen» da



**Figur 5.5:** En stabil prosess (sentrert rundt det samme resultat for alle tidspunkt) med liten varians (liten variasjon / smal fordeling langs resultataksen) til venstre og en ustabil prosess (sannsynlighetsfordelingene for de tre tidspunktene er sentrert rundt ulike resultatpunkt) med større og varierende varians (bredere og ulike utstrekninger langs resultataksen for sannsynlighetsfordelingene) til høyre.

den tilfeldige variasjonen er langt mindre.

En prosess er i **statistisk likevekt** eller statistisk stabil dersom det ikke finnes spesielle årsaker til variasjon – ved å eliminere spesielle kilder til variasjon kan man gjøre en prosess (statistisk) forutsigbar – det vil si forutsigbar innenfor de grenser som bestemmes av den naturlige eller tilfeldige variasjonen. Walter Shewhart<sup>32</sup> satte opp et antall kriterier for at en prosess skal kunne anses å være i statistisk likevekt. Det mest kjente av disse var kontrolldiagrammet som er illustrert i Figur 5.6.

Et kontrolldiagram er konstruert slik at dersom det finnes store, spesielle kilder til variasjon, vil man relativt raskt få varsel om dette, gjennom at man får observasjoner som ligger utenfor visse forutbestemte kontrollgrenser.

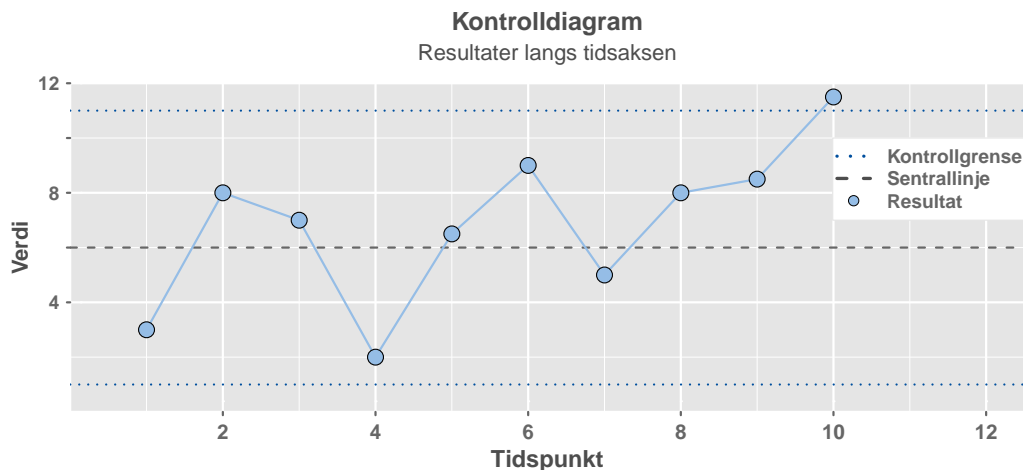
Hvis systemet av årsaker til variasjon er stabilt, bør det gå lang tid før man observerer verdier i diagrammet som ligger utenfor kontrollgrensene. Man ønsker å unngå «falsk alarm». Dessverre må man balansere to sannsynligheter opp mot hverandre: en høy følsomhet for spesiell variasjon og en lav risiko for falsk

alarm, noe som resulterer i et kompromiss mellom de to ønskede egenskapene.

Offentliggjøring av resultater fra medisinske kvalitetsregistre skal fungere som et **signal-system** og en bro mellom ledelsen og de delene av virksomheten som er direkte verdiskapende. Det kan da være viktigere å fange opp en eventuell forandring enn å være absolutt sikker på at en forandring faktisk har funnet sted. Dette drøftes nærmere i neste kapittel.

Hvordan grensene i et kontrolldiagram settes, kommer an på situasjonen, det vil si hvordan man forstår systemet av årsaker som har generert variasjonen. Et vanlig valg som anbefales av Shewhart, er  $3 \cdot \sigma$ -grenser (middelerdiene av det aktuelle målet pluss/minus tre ganger standardavviket for målet). Dette innebærer at det vil gå relativt lang tid mellom falske varsler (det vil si punkter utenfor kontrollgrensene selv om prosessen er statistisk stabil). Likevel vil diagrammet relativt raskt signalisere forekomst av en spesiell kilde til variasjon etter at en slik har oppstått.

Et kontrolldiagram gir også mulighet for å se om et tiltak har gitt ønsket effekt. Har re-



**Figur 5.6:** En prosessindikator verdi plottet over tid. Den tiende uken utløses et varsel, sannsynligvis er en spesiell kilde til variasjon påvist. Avstanden fra senterlinjen (som er prosessindikatorens gjennomsnittlige verdi over en lengre tidsperiode) til kontrollgrensene er ofte satt til tre ganger standardavviket for indikatoren [2].

sultatene faktisk forandret seg slik man trodde da tiltaket ble iverksatt? Kan tiltaket betraktes som en spesiell kilde til variasjon som virker positivt på de resultatene man vil oppnå? Dette illustreres i Figur 5.3. Omrentlige kontrollgrenser<sup>33</sup> for andelen plutselige spedbarnsdødsfall er illustrert i figuren, først for perioden før tiltaket fram til begynnelsen av 90-tallet, da det åpenbart inntraff en spesiell kilde til variasjon – den naturlige tolkningen er at det er tiltaket som har gitt effekt for den videre perioden.

Merk at man har betont håndtering av variasjon i tid, diakron variasjon. Tilsvarende resonnerment for synkron variasjon innebærer at man setter opp grenser for hva som kan betraktes som naturlig variasjon mellom de ulike enhetene som studeres. Det har her vært foreslått å anvende  $3 \cdot \sigma$ -grenser også her, men dette er – og må være – være svært situasjonsavhengig.

I offentliggjøring fra medisinske kvalitetsregistre finnes sjeldent eller aldri kontrollgrenser, men man ser ofte konfidensgrenser – noe som vil bli drøftet nærmere i neste kapittel.

En viss naturlig variasjon kan iblant framstå som spesiell. For visse behandlingsmetoder kan for eksempel en persons genotype ha betydning – et eksempel er Warfarin-behandling (blodfor-

tynnende legemiddel). Responser på behandlingen avhenger av genotype, noe som kan resultere i indikasjon på en spesiell kilde til variasjon – i dette tilfellet genotype. Dette gjelder også disposisjonen for brystkreft, som eksempelet i kapittel 3.1.2. Denne typen variasjon stammer fra det naturlige årsakssystemet. Naturlig variasjon kan altså inneholde komponenter av spesiell variasjon – men ofte spesiell på en måte som er statistisk mulig å modellere. Man kan gjennom observasjon anslå andelen av en viss genotype, og dermed forutsi effekter av denne typen variasjon.

## 5.5 Når har noe hendt?

Som tidligere nevnt er offentliggjøring av resultater fra medisinske kvalitetsregistre ment å fungere som et signalsystem – et «GPS-system» for helsesektoren som forteller hvor man er, slik at man kan gå videre mot visjonen. Det blir da viktig å være oppmerksom på mulige endringer og trender selv om det er risiko for at tilfeldighetene kan spille oss et puss. Å basere viktige beslutninger bare på tilfeldige variasjoner, kan være ødeleggende. Men å innlede en dialog omkring resultatene i en virksomhet er utelukkende positivt, og med et slikt utgangs-



punkt trenger man ikke være like oppmerksom på risikoen for at variasjonen som observeres, bare er tilfeldig. La oss studere resultatet i Figur 4.3, der HbA<sub>1c</sub>-verdiene for Klinikk A varierer i tiden. Gitt at det finnes andre indikasjoner på at klinikken avviker positivt, er det rimelig å anse nedgangen i 2007 som svært interessant og noe som burde føre til utvidet dialog og læringsmuligheter.

Generelt kan man, som en tommelfingerregel, si at når åtte punkter i et kontrolldiagram ligger under eller over det som er forventet, har noe uvanlig (utover den tilfeldige variasjon) inntruffet. Allerede etter fem punkter, som i Figur 4.3, kan det være relevant å starte en dialog omkring resultatet. (I et tilfelle som Klinikk A ville man imidlertid kanskje ikke trenge mer enn observasjonen fra 2007 for å bli nysgjerrig på hva som kan ha hendt da, særlig om man samtidig har en så sterk indikasjon som i Figur 4.2).

Ut fra ovenstående resonnement kan man av Figur 5.3 (spedbarnsdødelighet), selv uten noen form for kontrollgrenser, se at det har skjedd noe i perioden 1991–1998. Åtte observasjoner med avtagende verdier er en sterk indikasjon på at det har hendt noe som ikke kan forklares med tilfeldigheter. I dette konkrete tilfellet gir imidlertid kontrolldiagrammet varsel mye tidligere.

Som nok en tommelfingerregel kan man si at dersom en institusjon ligger over eller under nasjonalt gjennomsnitt fem ganger etter hverandre, er sannsynligheten for at dette kun skyldes tilfeldig variasjon ca. 5-6 prosent – altså ikke særlig sannsynlig. Det er da all grunn til å begynne å reflektere over dette. Ligger verdien på «rett» side, er det grunn til å forsøke å stimulere til ytterligere forbedringer, og om den er på «feil» side, bør man forsøke å skape en dialog omkring resultatene – en dialog som gjerne kan begynne med åpne spørsmål som stimulerer til refleksjon og siden til forbedring.

Det er imidlertid viktig at man i dialogen er tydelig på at tilfeldige variasjoner kan være en årsak til det som observeres.



## Kapittel 6

# Å beskrive variasjon

Så langt har ulike typer variasjon blitt drøftet. I dette kapitlet oppsummeres de viktigste begrepene før flere begreper for å beskrive variasjon og lignende fenomener, for eksempel usikkerhet, diskuteres.

### 6.1 Typer variasjon

I forrige kapittel var det flere eksempler på variasjon. **Tilfeldig variasjon** var en viktig kategori, og den viktigste typen av slik variasjon er resultatet av mange små, innbyrdes uavhengige årsaker der ingen av årsakene gir et dominerende bidrag. En prosess som bare framviser tilfeldig variasjon er i **statistisk likevekt**. Iblant er det man ser en blanding av ulike delsystemer: det kan for eksempel gjelde en så enkel og vanlig forekommende forskjell som den mellom kvinner og menn eller, mer generelt, ulike genotyper. Det er ikke alltid mulig å se hvilken genotype en person tilhører, men den kan likevel gi opphav til en spesiell kilde til variasjon. I denne sammenheng kan det være **naturlig variasjon** – altså en variasjon som forekommer naturlig og består av tilfeldig variasjon eventuelt i kombinasjon med variasjon som skyldes pasientenes/individenes iboende egenskaper (kjønn, genotype osv.). En kilde til variasjon som gir et dominerende bidrag, kaller man **spesiell variasjon**. Den gir et bidrag som er klart merkbart sammenlignet

med de andre variasjonskildene i prosessen. I tillegg til den typen spesiell variasjon som for eksempel ulike genotyper kan gi bidrag til, finnes det andre kilder til variasjon som kan være mer eller mindre uønsket: en ny medarbeider som ikke kjenner rutinene og derfor iblant gjør feil, en sykepleier som blir forstyrret under utdelingen av legemidler, ulike behandlingsmetoder som på grunn av pasientens tilstand ikke er forsvarlige, og så videre. Iblant kalles variasjon som ikke er forbundet med pasienten og dens tilstand, for **uønsket variasjon**. Dette er variasjon som man i det lange løp vil unngå eller redusere så mye som mulig.

Men det finnes også variasjon som er positiv. Det er viktig at helsesektoren prøver å forbedre helsetjenestene og kanskje finne en bedre behandling eller en bedre organisering av prosessen. Dette kan danne utgangspunktet for forbedring og dermed en positiv spesiell effekt. Denne typen spesiell variasjon kan man trekke lærdom av. Når tiltak iverksettes med sikte på forbedring, er det også viktig at man måler og faktisk oppnår forbedring – at tiltaket på en positiv måte avstedkommer et spesielt bidrag – noe man i så fall må prøve å bevare og trekke lærdom av (jfr. PDSA-sirkelen i Figur 4.1).

Variasjon mellom enheter, helseinstitusjoner eller for den saks skyld pasienter, har allerede blitt omtalt. Når man ikke trekker noen tidsdimensjon inn her, altså ser bort fra variasjon

over tid, kaller man dette **synkron variasjon**. Variasjon over tid, enten den er tilfeldig eller spesiell, omtales som **diakron variasjon**. En forbedring viser seg kanskje først som en diakron variasjon – forbedringen har dannet utgangspunkt for en spesiell variasjon over tid (se Figur 5.3). Når man deretter sammenligner ulike helseinstitusjoner som har samme type prosess, vil kanskje den forbedrede helseinstitusjonen gi opphav til en spesiell synkron variasjon, altså en variasjon som skiller seg fra de andre.

## 6.2 Prosesser i statistisk likevekt

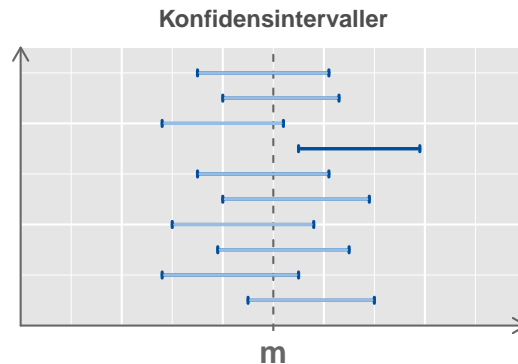
En prosess som utelukkende har naturlige kilder til variasjon, omtaler man som en prosess i **statistisk likevekt**, og man kan til en viss grad forutsi hva som kommer til å skje i framtiden. Man kan ikke forutsi framtiden helt nøyaktig, for fortsatt foreligger den tilfeldige eller naturlige variasjonen – men denne typen variasjon er omtrent som terningkastene omtalt i kapittel 5.2. Man kan ikke si noe nøyaktig om den enkelte pasient, men man kan si ganske mye om et større antall pasienter.

En annen betegnelse for statistisk likevekt er statistisk stabilitet. Til en (stokastisk<sup>34</sup>) prosess som noenlunde er i statistisk likevekt kan man relatere en (sannsynlighets-) **fordeling** som gir informasjon om hvordan de observerte måleverdiene vil fordele seg på den numeriske skalaen (som illustrert i Figur 5.4). For å gi en grov beskrivelse av en fordeling kan man angi fordelings **middelverdi** som er gjennomsnittsverdien dersom man har et veldig stort antall verdier/pasienter. Fordelingens **median** er verdien som har like mange observasjoner til høyre som til venstre for seg, mens fordelings spredning/varians kan beskrives ved hjelp av fordelings **standard-**

**avvik** som beskriver hvor mye observasjonene vil avvike fra middelveidien. For å beskrive spredning kan man også anvende **kvartiler**, som i Figur 5.2; nedre kvartil angir fordelingsverdien hvor 25 prosent av de øvrige verdiene ligger nedenfor/under denne kvartilverdien, mens øvre kvartil angir fordelingsverdien hvor 25 prosent av de øvrige verdiene ligger over kvartilverdien.

## 6.3 Usikkerhet, sannsynlighet og middelveidi

Som allerede konstatert gir en prosess som er i statistisk likevekt eller er statistisk stabil, likeartede resultater i framtiden som de resultater den allerede har gitt. Det betyr at man har et estimat eller en approksimasjon på prosessens fordeling, som gir informasjon om prosessens fremtidige måleverdier. Men selv om man har observert måleverdier fra prosessen, betyr ikke det at man kjenner prosessens fremtidige verdier – det finnes en viss usikkerhet selv om prosessen er statistisk stabil. Usikkerheten beskrives gjerne med det man kaller **konfidensintervall**. Om man for eksempel er interessert i prosessens middelveidi, kan man anslå denne ved hjelp av middelveidien av de observerte resultatene fra prosessen, men i tillegg kommer altså den usikkerheten som skyldes tilfeldighetene i resultatet. For å beskrive denne usikkerheten angir man et intervall som man regner med vil inneholde den reelle middelveidien i prosessen. Et bredt intervall vil ofte inneholde den reelle middelveidien, mens et smalt intervall sjeldnere vil inneholde denne middelveidien. Konfidensintervallets bredde henger sammen med det som kalles intervallets **konfidensgrad** eller **signifikansnivå**. Konfidensgraden, som typisk ofte er 95%, kan **tolkes**<sup>35</sup> til å være sannsynligheten for at intervallet dekker den reelle middelveidien. Et intervall med høy konfidensgrad vil derfor væ-



**Figur 6.1:** Ti 95%-konfidensintervaller for verdien  $m$ . Ni av de ti konfidensintervallene, som tilsvarer 90% av intervallene, inneholder her den korrekte  $m$ -verdien. Intervallet i mørkeblått dekker imidlertid ikke verdien.

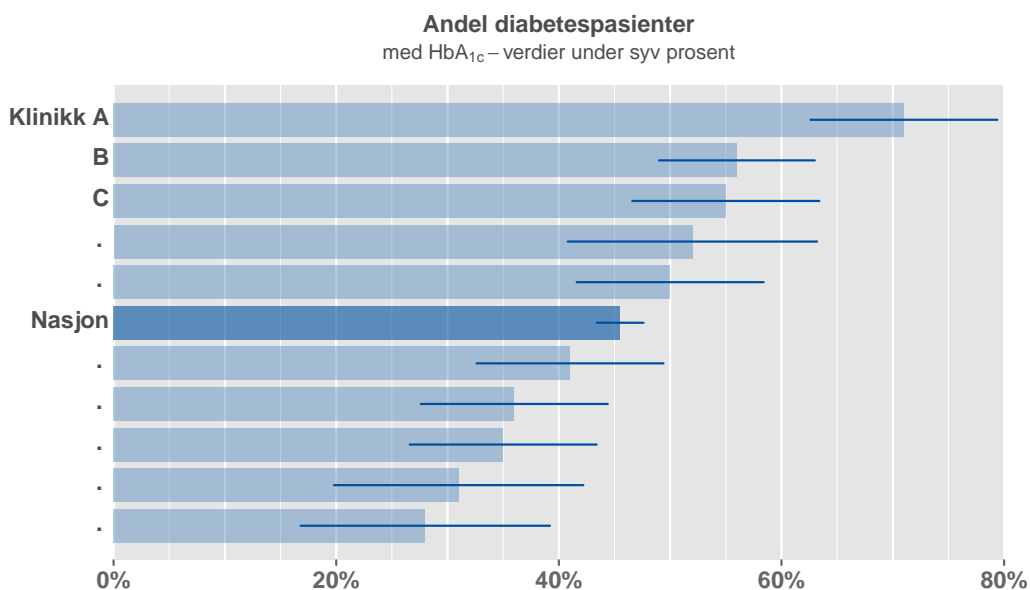
re bredt, mens et intervall med lav konfidensgrad vil være smalt. Dersom man anvender en konfidensgrad på 95% kan man, i det lange løp i tilsvarende situasjoner, anta at middelveiden ligger innenfor intervallet i 95% av tilfellene. Dette er illustrert i Figur 6.1. Der er den sanne middelveiden  $m$  ukjent, men basert på ti ulike observasjoner av prosessens verdier er det konstruert ti konfidensintervaller med en konfidensgrad på 95%, noe som betyr at man kan anta eller forvente at intervallene inneholder den reelle middelveiden  $m$  i 95% av tilfellene. I figuren er det ett av de ti konfidensintervallene som ikke inneholder den korrekte  $m$ -verdien. Dette intervallet utgjør 10% av det totale antallet av intervaller i figuren.

I Figur 4.2, som er gjengitt i Figur 6.2, er intervallene rundt den oppnådde andelen er beregnet. Konfidensgraden angir hvor sikker man kan være på at intervallet faktisk dekker den verdien som er relevant for prosessen. Dersom to enheter er like flinke og har identiske prosesser, er det bare tilfeldig variasjon som gjør at resultatene vil variere. Hadde man hatt mange enheter med samme egenskaper, vil konfidensintervallet dekke prosessens «reelle» verdi i 95 prosent av tilfellene, slik at fem prosent av intervallene ikke dekker verdien. Dette betyr at man, hvis det ikke er noen forskjeller mellom enhetene, kan forvente at rundt ett av tjue in-

tervaller ikke vil dekke prosessens «reelle» verdi. At man **forventer** dette betyr at det iblant vil være færre og iblant flere av intervallene som ikke dekker prosessens verdi. I Figur 6.1 er det for eksempel 10% av intervallene som ikke dekker den reelle  $m$ -verdien.

Hvis de ti enhetene/klinikkene i Figur 6.2 hadde hatt nøyaktig likeverdige prosesser, kunne man forvente at 95 prosent av konfidensintervallene, altså nitten av tjue intervaller, ville dekke den typiske middelveiden for prosessen<sup>36</sup>. Det ville derfor ikke være særlig klokt å treffe kostbare beslutninger på grunnlag av et resultat hvor ett av ti intervaller, som i Figur 6.1, ikke dekker middelveiden. Men å ta det som utgangspunkt for å stille spørsmål og starte en dialog med den operative virksomheten er likevel et fornuftig tiltak.

I Figur 6.2 ser man derimot at flere klinikker har konfidensintervaller som ikke dekker nasjonaens middelveid (som angitt av den mørkeblå søylen). Det er da grunn til å tro at forskjellene mellom klinikkene ikke kan forklares av tilfeldig variasjon. Spesielt har Klinikk A et konfidensintervall som ligger langt over middelveiden. Da har alle de andre klinikkene god grunn til å lure på hva det er denne enheten gjør som synes å gi så mye bedre resultater. I forbedringsarbeidet er det spesielt viktig å prøve å forstå hva de som er best gjør, og re-



**Figur 6.2:** Andelen pasienter som har HbA<sub>1c</sub>-verdi under syv prosent for et utvalg fiktive diabetesklinikker. Klinikk A har et sterkt avvikende resultat som gir grunn til å undersøke om klinikken har rutiner eller behandlingstilbud som andre kan lære av. Det er flere klinikker med konfidensintervaller ikke dekker middelverdien/landsgjennomsnittet. Flere av klinikkene har også konfidensintervaller som ikke dekker hverandre.

flektere over hvordan man selv kan forandre seg slik at man får like gode eller bedre resultater. **Hensikten** med offentliggjøring av resultater fra de medisinske kvalitetsregistrene er ikke for institusjonene å sammenligne seg med et langsgjennomsnitt, men å se til de beste og kontinuerlig arbeide for å bli like bra eller bedre, samtidig som de beste også arbeider for å bli bedre. Det er klinikkenes utvikling langs tidsaksen, altså den diakroniske utviklingen, som er av størst interesse.

enkle grafer som gir institusjonene mulighet til å følge egne resultater og kvalitetsindikatorer langs tidaksen.

## 6.4 Å illustrere variasjon

For å analysere variasjon med henblikk på å avdekke forbedringsmuligheter, finnes det et stort antall metoder og verktøy som er utviklet innenfor kvalitetsteknikk, prosesskontroll, statistikk og epidemiologi. Dette vil ikke bli berørt her, men litt overordnet kan man si at det for praktikere som skal utføre forbedringsarbeidet, er viktig å visualisere resultatene over tid med

## Kapittel 7

# Fallgruver

En nybegynner uten god nok forståelse av variasjon, kan lett gå seg bort i de statistiske irrgangene og havne i fallgruver som han eller hun kanskje ikke en gang oppdager. Å bli lurt av tilfeldighetene og tro at tilfeldigheter kan påvirkes, eller å tro at spesielle årsaker bare skyldes tilfeldig variasjon, er to aspekter som allerede har blitt diskutert. I dette kapitlet illustreres et par statistiske fallgruver, og mot slutten av kapitlet omtales også en del andre, men ikke mindre problematiske, fallgruver. Det er viktig å innse at det man vanligvis er interessert i, er hvordan klinisk interessante fenomener varierer – numeriske mål kan fange opp slike variasjoner mer, middels eller mindre bra. Det vil også bli diskutert hvordan naiv bruk av data kan føre til dårligere datakvalitet.

### 7.1 Sentralitets- og spredningsmål

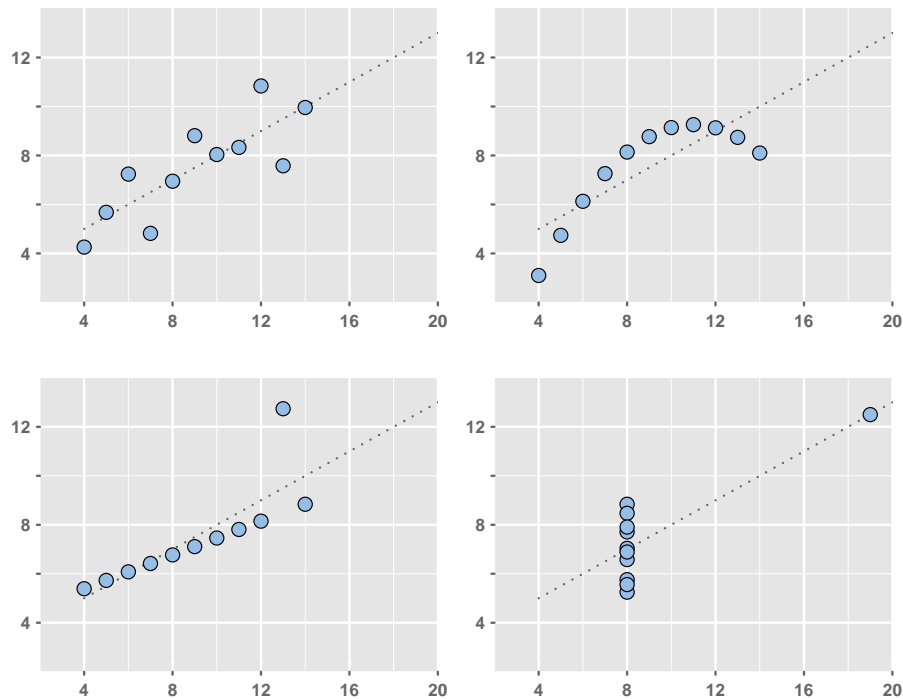
I mange sammenhenger presenteres middelveier, standardavvik og, når man studerer par av variabler, **korrelasjonskoeffisienter**<sup>37</sup> og/eller tilsvarende regresjonskoeffisienter (typisk fra lineær regresjon). Det kan synes som en fyldig beskrivelse av den variasjon som finnes, men dette er dessverre ikke tilfelle. Den grafiske presentasjonen av data i Figur 7.1 kalles gjerne Anscombes kvartett (etter Francis Anscombe) og gir fire eksempler der middel-

verdiene og standardavvikene for  $x$  og  $y$  er like i alle tilfellene. Og som om ikke det var nok, er også korrelasjonskoeffisientene (og regresjonslingning) like i alle fire tilfeller. Til tross for felles middelveier, standardavvik og korrelasjonskoeffisient gir den grafiske illustrasjonen av data opphav til svært forskjellige bilder.

Dette konstruerte eksempelet viser hvor viktig det er å illustrere data grafisk, for med den grafiske illustrasjonen blir det temmelig åpenbart at det må konkluderes helt ulikt i de fire tilfellene. Dersom det kun er de sammendragende målene som beskrives, vil man kanskje behandle alle de fire tilfellene på nøyaktig samme måte selv om de er fundamentalt forskjellige. Dersom man presenterer datasettene i fire histogrammer, kan også konklusjonene bli feil ettersom histogrammene kan se like ut selv om de bakenforliggende tidsseriene er svært ulike.

### 7.2 Faren med sentralitetsmål

Altfor ofte tar man beslutninger basert på middelveier. Faren for å gjøre feil er da relativt stor. En form for feilslutninger ble omtalt i forrige avsnittet. Et annet enkelt, men slående, eksempel kan man se innen kapasitetsplanlegging; man vet typisk hva det gjennomsnittlige behovet er (for eksempel med hensyn til innkommende pasienter, sengeplasser og hvor mye kapasitet hver enkelt av dem krever) og



**Figur 7.1:** Fire datasett som viser svært forskjellige mønstre selv om de har samme middelværdi, standardavvik og korrelasjonskoeffisient.

kan lett gå i fellen med å planlegge den totale kapasiteten slik at den tilsvarer det gjennomsnittlige behovet. Men det man da har glemt, er variasjonen, som ut fra det man tidligere har konstatert, kan komme både innenfra og utenfra. Jo større variasjon, desto større forskjell kan det være mellom gjennomsnittlig behov og det reelle daglige kapasitetsbehovet. Dette kan medføre lange ventelister og er et fenomen man alt for ofte kan observere innen helsevesenet.

Denne type problemstillinger er ikke bare relatert til ventetider, det er et generelt problem. Man bør alltid ta variasjonen med i beregningene. Dette er også en aktuell problemstilling med tanke på offentliggjøring fra medisinske kvalitetsregistre hvor det ofte bare er middel- og medianverdier som oppgis eventuelt sammen med utsagn om sannsynligheter. Generelt er det dessverre for lite fokus på å beskrive variasjonen i ulike datasett.

## 7.3 Statistisk skjevhet

Man ønsker at dommerne i rettssalen skal være upartiske og ikke ha noen forutinntatt mening om skyldspørsmålet. Bare de faktiske forhold skal legges til grunn. På samme vis vil man trekke konklusjoner av det data kan fortelle om virkeligheten uten at verken forutinntatte meninger eller andre former for ubevisste feilkilder bidrar til at konklusjonene blir feil. Ofte brukes det engelske ordet «**bias**»<sup>38</sup>, men her anvendes det norske begrepet **forventnings-skjevhet**, eller bare **skjevhet** (som i statistisk skjevhet).

### 7.3.1 Utvalgsskjevhet

Hvis man vil uttale seg om egenskapene til en populasjon (for eksempel en pasientgruppe) men ikke har mulighet til å undersøke samtlige individer eller enheter i denne populasjonen, er det vanlig at man i stedet tar et tilfeldig utvalg



fra populasjonen. De relevante egenskaper hos individene bestemmes, og man antar at populasjonen, med en statistisk feilmargin som kan beregnes, har de samme egenskapene som individene i utvalget. Dette stiller imidlertid store krav til utformingen av utvalget – hvem er det som i realiteten har hatt anledning til å bli med i utvalget? Hvem er det som virkelig kan og vil være med? Det kan oppstå store feil om utvalget ikke er **representativt** for populasjonen.

Denne type problemstilling er i utgangspunktet ikke så aktuell for medisinske kvalitetsregistre ettersom målet er å samle data for samtlige pasienter som er gjennom den aktuelle prosessen per år. Men jo større bortfall man har, desto viktigere er det å reflektere over hvorvidt de som har falt bort, virkelig har de samme egenskaper som dem som blir observert. Når registreringsfrekvensen er lav, hvem er det da som ikke registreres? Risikoen for feilslutninger på grunn av bortfall kan, paradoksalt nok, øke når man – noe som i og for seg er veldig positivt – inkluderer flere pasientrapporterte resultater. Det er alltid en fare for at de pasientprosessene som ikke er registrert er systematisk annerledes enn de som er registrert.

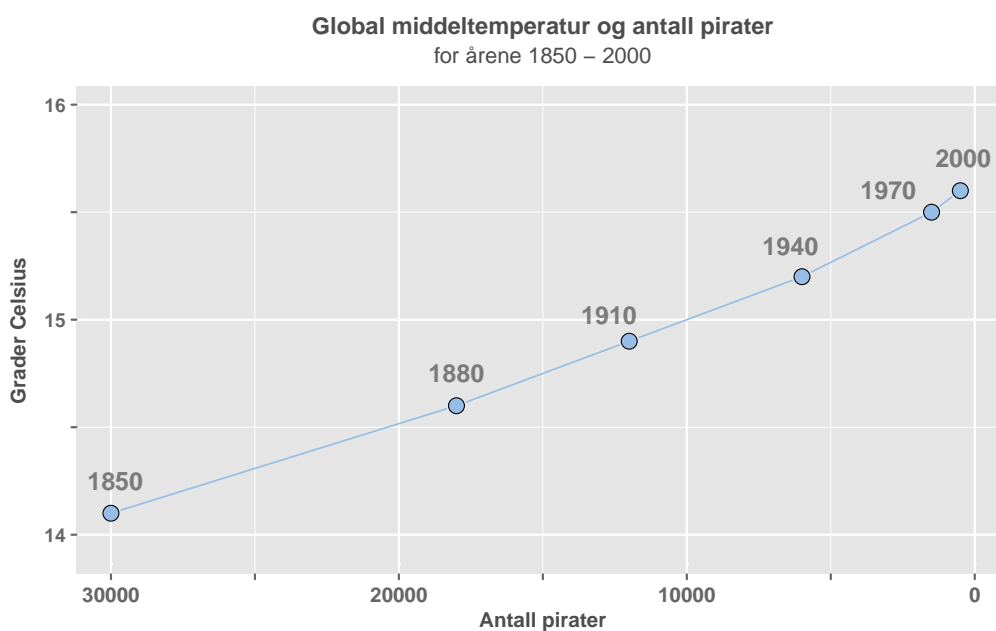
Problemet med ikke-representative utvalg er også til stede i medisinsk litteratur, der man ikke alltid har vært like nøye med at utvalget virkelig skal representere den relevante populasjonen. Resultater som har framkommet basert på middelaldrende menn, har blitt anvendt på kvinner, barn og eldre med kanskje en helt annen metabolisme og/eller andre egenskaper som er viktige i sammenhengen. I den senere tid har man imidlertid viet dette større oppmerksomhet. Her er resultater fra kvalitetsregistre viktige for å kompensere for mangler som kan finnes i medisinske forskningsresultater.

### 7.3.2 Den glemte nevneren

Det florerer med merkelige medisinske «sannheter» på folkemunne. Alle har en bekjent som har blitt frisk etter en viss type behandling, for eksempel «behandling» med homeopatiske preparater. Men hvordan kan man vite om det finnes en årsakssammenheng mellom behandlingen og helbredelse? Det at hendelse B (helbredelse) inntraff etter hendelse A (behandling) betyr ikke at hendelse A forårsaket B. Kausalitet eller årsakssammenheng er et intrikat og ofte diffust begrep. Det er flere fallgruver å gå i, og problemstillinger knyttet til placeboeffekten vil ikke bli berørt her. To av de vanligste fallgruvene vil derimot bli drøftet; «den glemte nevneren» og i neste delkapittel forbindelsen mellom årsak og virkning.

Den første fallgruven, gjerne omtalt som «den glemte nevneren», skyldes at man ofte bare hører snakk om dem som faktisk har blitt friske etter en behandling. Hvordan gikk det med alle de andre som har fått den samme behandlingen? Og alle de som har blitt friske uten denne behandlingen? Det er viktig alltid å være på vakt når det ikke finnes eller oppgis noen opplysninger om hvor mange som har fått behandlingen. Det er ikke nok å beskrive hvor mange som har blitt friske (tallet som skal stå i telleren når man betrakter en andel) – man må også vite hvor mange som har fått behandling (tallet som skal stå i nevneren). **Andelen** som har fått behandling og blitt friske er viktig, og det relevante spørsmålet er om denne andelen er større enn andelen som har blitt friske uten å ha fått behandlingen.

Denne typen skjevhet kalles **overlevelses-skjevhet** (survival bias), det vil si at det kun finnes (eller er overvekt av) rapporter om de positive resultatene – i dette tilfellet de som har fått behandling og blitt friske. Formulert slik er det lett å gjennomskue, men skjevheten er vanlig innenfor forskning – da omtalt som **publiseringsskjevhet**. Dette betyr at



**Figur 7.2:** Global middeltemperatur og antatt antall pirater i verden for årene 1850-2000. Middeltemperaturen har økt mens antall pirater har gått ned.

det blant alle publiserte artikler er en overvekt av studier som viser en positiv (statistisk signifikant) effekt – alle studier som ikke påviser noen effekt havner gjerne i skrivebordsskuffen. Det finnes metoder for å «avsløre» publiseringsskjevhet med såkalte metaanalyser, men temaet vil ikke bli berørt her.

## 7.4 Årsakssammenheng

Det er vanskelig å være sikker på årsakssammenhenger. En litt useriøs, men forhåpentligvis pedagogisk god, illustrasjon er gitt i Figur 7.2; i dette tilfellet vet man godt at det ikke finnes noen direkte kausal sammenheng mellom global middeltemperatur og estimert/antatt antall pirater i verden, men likevel virker det på figuren som om det er en sammenheng mellom de to variablene ettersom figuren viser en sterk korrelasjon. Et naivt utsagn basert på dette er at man kan senke den globale gjennomsnittstemperaturen ved å innføre flere pirater. Her er det lett å avvise utsagnet som «sjøsprøyt»,

men generelt er det en nokså vanlig feil å forveksle assosiasjon (korrelasjon er en lineær assosiasjon) med kausalitet. Som dette useriøse eksempelet illustrerer, er det lett å lage utallige diagrammer som ved første øyekast kan virke svært overraskende eller feilaktig «troverdige». Et mantra å ha bakhodet er at «**korrelasjon impliserer ikke kausalitet**».

Hvis man tar utgangspunkt i to tidsserier som er vesentlig økende eller fallende over tid kommer de til å korrelere. Utviklingen til den globale temperaturen har vært økende, mens antall pirater i verden har vært fallende. Det eneste de har til felles, er at de øker eller avtar med tiden, og tiden vises ikke i diagrammet. Tiden er her det som på norsk kan kalles en **konfunderingsfaktor**<sup>39</sup> (confounder på engelsk). Det er ikke umulig at det noen ganger kan finnes konfunderingsfaktorer bak relevante sammenhenger – variabler som gjør «funnet» tiltalende, men som er helt uten verdi for å forstå årsakssammenhengen. For å unngå problemene med konfunderingsfaktorer er det veldig viktig å gjennomføre eksperimentet på en måte som gjør at konfundering ikke inn-

treffer. Likevel kan det hende at du uforvarende endrer konfunderingsfaktoren samtidig som du endrer den relevante variabelen. Da blir det umulig å skille mellom en relevant effekt og en effekt som skyldes konfunderingsfaktoren.

Ikke sjelden brukes korrelasjonskoeffisienten som indikasjon på årsakssammenheng. Som illustrert er ikke dette alltid en god indikasjon; i Figur 7.2 er det høy korrelasjon mellom global gjennomsnittstemperatur og antall pirater.

Det er for øvrig ikke bare konfunderingsfaktorer som kan ligge bak en høy verdi på en korrelasjonskoeffisient. Det kan også være en enkelt måleverdi som er blitt feil, og som påvirker resultatet uforholdsmessig mye.

## 7.5 Måling av andeler som innfrir et kriterium

Å måle andeler som oppfyller et kriterium medfører en risiko det bør tas hensyn til i tolkning og analyse; et eksempel er HbA<sub>1c</sub>-målinger for barn med diabetes, som bør være under åtte prosent. Hvor stor andel av pasientene oppfyller «minimumskriteriet»? Grensen er satt mer eller mindre vilkårlig<sup>40</sup>, og det er risiko for at man i sine anstrengelser for å nå målet gjør det verre for et flertall av pasientene.

Problemetstillingen med å rapportere andeler kan illustreres ved å betrakte et konstruert eksempel med HbA<sub>1c</sub>-målinger. Anta at man anvender grenseverdien åtte prosent og rapporterer andelen som ligger under verdien. Anta videre at det er to likeverdige diabetesklinikker, klinikk A og B, som har nøyaktig samme fordeling av pasienter med hensyn til HbA<sub>1c</sub>-verdiene; fordelingen er som illustrert til venstre i Figur 7.3, hvor en viss andel av pasientene har HbA<sub>1c</sub>-verdi under åtte prosent. Anta så at Klinikk B iverksetter et tiltak som er ment å forbedre situasjonen (målt med andelen pa-

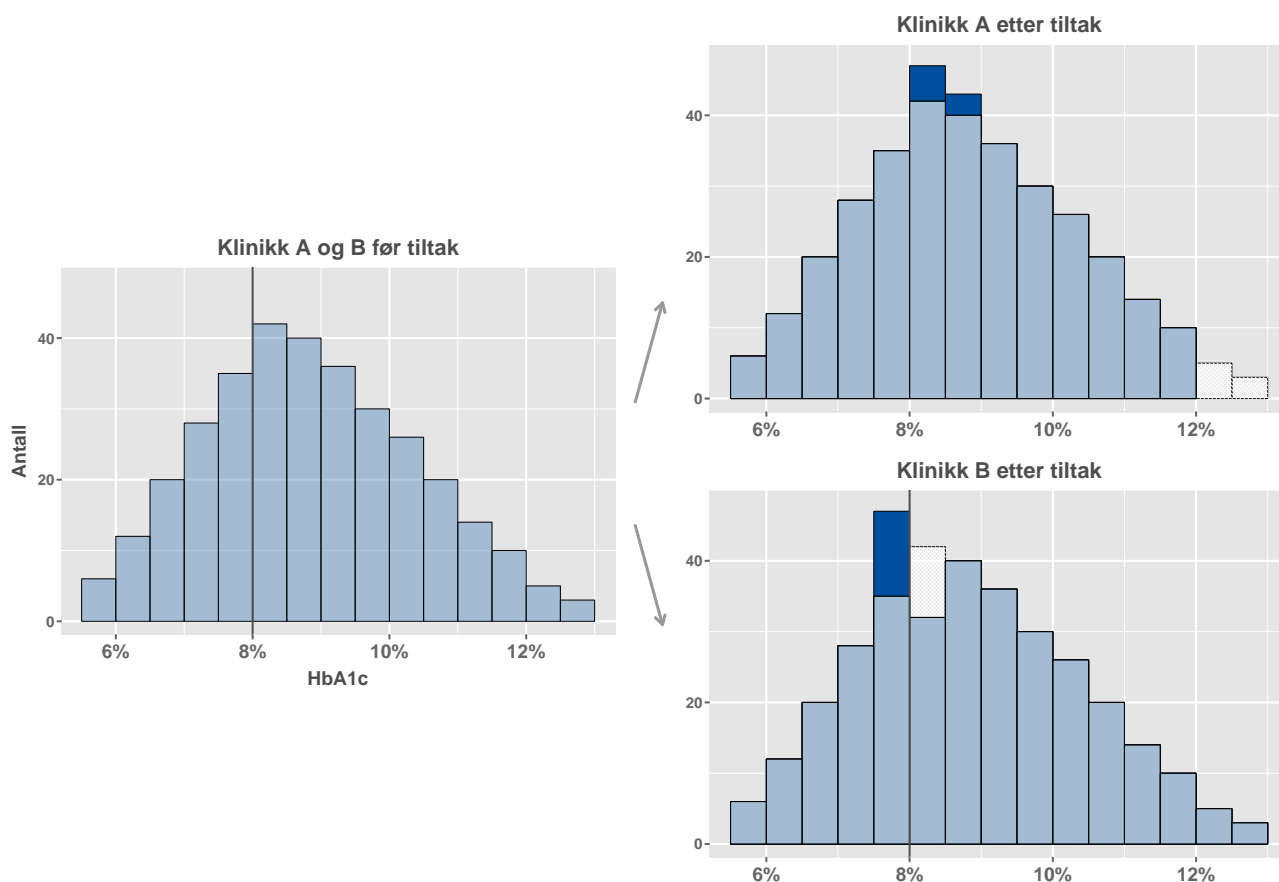
sienter under grenseverdien). På et senere tidspunkt, etter tiltaket, rapporterer Klinikk A og B sine resultater. Klinikk B har økt andelen pasienter under grenseverdien, mens Klinikk A ikke har forbedret seg med hensyn til rapporteringsmålet. Har Klinikk B gjennomført et vellykket forbedringsarbeid?

Anta at de rapporterte resultatene til Klinikk A og B etter gjennomført tiltak er som illustrert til høyre i Figur 7.3. Det er ikke innlysende at Klinikk B har gjennomført et vellykket forbedringsarbeid. Mye taler faktisk for at Klinikk A har en bedre strategi for pasientene enn Klinikk B. Resonnementet knyttet til rapportering av en andel fører til «premiering» av en intervensjon eller tiltak som sannsynligvis er mindre god. Klinikk A har forbedret resultatene for de mest utsatte pasientene (de med høye HbA<sub>1c</sub>-verdier) og oppnådd en gjennomsnittlig reduksjon som er større enn den Klinikk B har oppnådd. Klinikk B har kun marginalt redusert verdiene til fire pasienter som i utgangspunktet lå rett over grensen slik at de ved siste resultatrapportering akkurat kommer under den «magiske» grensen. Belønning av det beste arbeidet basert på andeler kan sende systemet mindre heldige signaler.

### 7.5.1 Betydningen av operative definisjoner

Illustrasjonen med diabetesklinikkene viser et generelt og viktig aspekt, nemlig betydningen av å ha klare **operative definisjoner**. Flyttes grensene, vil den operative definisjonen av begrepet «andel godkjente» endre seg, og man kan ende opp med en størrelse som over tid mister sin betydning. Dersom en ledergruppe drøfter en størrelse der definisjonene endrer seg, kan det bære veldig galt av sted om man ikke er observant.

Ett annet aspekt ved operative definisjoner illustreres av måling av ventetider; her kan



**Figur 7.3:** Fordeling av  $HbA_{1c}$ -verdier for pasientene ved Klinikk A og B, før og etter et tiltak er gjennomført. De hvite feltene er flyttet til de mørkeblå etter tiltak.

mange faktorer være relevante – når settes klokken i gang, når stoppes den, hvilke pasienter regnes med, hvilke prosesser blir berørt osv. Det er med andre ord mulig å «manipulere» systemet for å oppnå et gitt mål.

## 7.6 Datakvalitet og tillit til data

Problematikken knyttet til **datakvalitet** har allerede blitt nevnt. Graden av skepsis kontra tillit til statistiske data henger ofte sammen med hvilken avstand den som vurderer dataene har til kilden for datainnsamling. Når man er nær punktet for datainnsamling er skepsisen

stor, man vet at det kan være problemer med registrering og feiltolkninger. På litt avstand er skepsisen mindre, og man har stor tillit til data – i mange tilfeller kanskje litt for stor tillit. Dette er forståelig; det er vanskelig å ha tilstrekkelig kjennskap til problematikken «på gulvet» samtidig som man trenger data for å treffe eller legitimere beslutninger. På større avstand fra datainnsamlingspunktet, det vil si hos utenforstående som vurderer og kritiserer, vil skepsisen igjen typisk øke og tiltroen avta. En som gjør en kritisk vurdering, har ikke samme behov for å tro på data for eksempel av legitimitetshensyn – snarere tvert imot. Det er ofte en ekstern granskers oppgave å være kritisk.

## 7.7 Mer enn numerisk variasjon

I kvalitetsregistrene er variasjon oftest uttrykt numerisk, men mye av den vensentlige variasjonen kan ikke fanges opp av slike tall. De numeriske verdiene som brukes er oftest kun en enkel indikator for den variasjonen som faktisk finnes i de bakenforliggende fenomenene av interesse. Å beskrive hvor syk en diabetespasient er ved å angi en HbA<sub>1c</sub>-verdi, kan være enkelt og lett anvendelig, men dette tallet beskriver ikke hele det sykdomsfenomenet man – og framfor alt pasienten – er interessert i. Ikke en gang om det legges til mange andre relevante variabler og interaksjoner mellom disse, vil man få et dekkende bilde av sykdommen hos en enkelt pasient. Informasjonen man har, er alltid mer eller mindre fragmentert.

## 7.8 Et motstykke til fallgruvene

Til tross for de fallgruvene som finnes, må man ikke havne i håpløshetens felle, det vil si kapitulere for alle de feilmuligheter som kan smyge seg inn når man forsøker å bruke statistiske data. Som tidligere understreket, bør rapportering fra medisinske kvalitetsregistre betraktes som et signalsystem, som et grunnlag for å fortsette dialogen i organisasjonene og som noe som kan fungere som et felles utgangspunkt. Ved hjelp av rapporteringen kan man sikte seg inn på de områdene som kan være interessante å jobbe videre med, som disuktert i kapittel 4. Rapporteringen fra medisinske kvalitetsregistre må betraktes som et utgangspunkt for videre analyser. Disse analysene må gå langt dypere enn det som er gjort her, og i tillegg til aspektene av mer allmenn karakter samt aspekter som krever særskilte fagkunnskaper innenfor det aktuelle området.



## Kapittel 8

# Refleksjoner og anbefalinger

Så langt har både muligheter og behov for å benytte data fra medisinske kvalitetsregistre for å utvikle helsetjenestene, blitt diskutert. Variasjon og sentrale relaterte begreper har blitt drøftet, og endelig har også noen fallgruver blitt omtalt. Dette kapitlet inneholder refleksjoner omkring dette samt noen anbefalinger for fremtiden.

### 8.1 Muligheter

Medisinske kvalitetsregistre gir store muligheter til læring og forbedringer i helsesektoren. Til tross for de omtalte og mer intrikate fallgruver er det ingen grunn til å miste motet; mulighetene overskygger definitivt problemene!

Først og fremst er rapporteringen fra medisinske kvalitetsregistre et **signalsystem** som kan brukes som et faktabasert utgangspunkt for en dialog mellom den overordnede ledelsen og de som aktivt og direkte påvirker de verdiskapende prosessene.

Når man vurderer resultatene bør man velge ut områder av stor betydning og områder noen eller enkelte enheter har gode resultater, både faktisk og statistisk, i forhold til andre – finnes det positive aspekter ved enkelte enheter som kan replikeres eller forsterkes slik at helsetjenestene blir enda bedre? Positive trender bør fanges opp og medføre tydelige tilbakemeldinger og stimulerende dialog. Alle enheter

bør søke å lære av «de beste i klassen» og støtte/oppmuntre/stimulere sine medarbeidere til å søke stadig forbedring og en kontinuerlig læringsprosess.

Variasjon mellom institusjoner i helsesektoren må vektlegges og medføre spørsmål av typen omtalt i kapittel 4. Alle systemer kan forbedres – at man i dag arbeider med rapportering fra medisinske kvalitetsregistre og har fått aksept for den typen oppfølging som kvalitetsregistrene gir anledning til, er et **meget viktig framskritt**. Man har fjernet seg et godt stykke fra de problemene Codman, som for hundre år siden ivret for åpne sammenligninger av resultater i helsevesenet og havnet i unåde blant sine kolleger, opplevde da han forsøkte seg på det man i dag gjør. Se appendikset kalt *Historisk tilbakeblikk* for mer om dette.

I fortsettelsen drøftes konsekvenser av fallgruver og hva de innebærer for tolkningen av rapporterte resultater fra kvalitetsregistrene.

### 8.2 Resultatbasering og forbedringsarbeid

Forskning viser at belønning for visse prestasjoner kan føre feil av gårde<sup>41</sup>. Dette gjelder spesielt hvis de **operative definisjonene**<sup>42</sup> av de måltall som anvendes, ikke er tydelige. Det finnes avskrekkende eksempler på kreative tolkninger av definisjoner av mål som kan

påvirke det økonomiske resultatet i betydelig grad. Spesielt ved studier av ventetider er det lett å finne forskjellige måter å sette pasientene i køer som ikke regnes med eller vente med å «sette i gang klokken» til et så sent tidspunkt som mulig osv. En ikke altfor uvanlig reaksjon når ventetider diskuteres er at man setter inn et støt for pasienter som nærmer seg grensen, og kanskje forsømmer pasienter med større behov, men med mer tid igjen til de når den «magiske» grensen. Det kan også settes inn ekstra personalressurser der det er mange som er i ferd med å overskride grensen. Risikoen er stor for at man med disse metodene for å nå målverdiene faktisk skaper mer arbeid og blir tvunget til å løpe fortore eller sette inn mer ressurser i stedet for å arbeide smartere.

Dette kan føre til forskyvningseffekter; mange i helsevesenet forteller for eksempel at de ikke får tid til å gjennomføre gjenbesøk for tidligere pasienter ettersom det er krise med hensyn til ventetidene for de førstegangsbesøkende. Det er grunn til å være bekymret for disse forskyvningseffektene; pasienter som har kort tid igjen til en magisk tidsfrist, får komme til før pasienter som fra et **behandlingsynspunkt** burde gå foran.

Det kan også være verdt å nevne at en ikke urimelig tolkning av de lange køene og ventetidene, er at de er en gjenværende effekt av at det i tidligere uformelle systemer for ressursfordeling kan ha vært fordelaktig å ha lange køer. Enheter kan ha benyttet pasientkøene til å synliggjøre nødvendighet og lite ressurser. At dette forhåpentligvis ikke lenger er tilfellet, trenger ikke spille noen rolle; strukturer som har fått sette seg over lang tid er det ikke helt enkelt å endre på. Samtidig har en del ledere rapportert at de har fått mer ressurser når de mot slutten av året risikerte å få mange pasienter som passerer den maksimalt aksepterte ventetiden. Signaleffektene dette har på systemet er det god grunn til å ta på alvor.

Risikoen er også stor for at optimering i ett delsystem skaper problemer i et annet. Ujevn flyt, altså stor variasjon i et delsystem, kan lett skape problemer for andre deler av systemet. Lengde på køen er ikke bare en funksjon av forventet behov og forventet kapasitet – variasjonen er også en relevant faktor som kan ha en høyst påtakelig effekt<sup>43</sup>. Om man i en del av systemet setter inn ekstra ressurser for å unngå altfor mange pasienter på feil side av den magiske grensen, kan det skape problemer i etterfølgende delprosesser. Gjennom et slikt tiltak kan det skapes køer i andre deler av systemet, slik at færre pasienter totalt får hjelp selv om man har klart å holde seg innenfor den magiske tidsgrensen.

Et annet argument mot belønninger basert på data av typen som primært bør hentes fram for å forbedre virksomheten, er at dersom risikoen for ulike typer forvrengning eller manipulering (bevisst eller ubevisst) av data øker, øker også risikoen for at datakvaliteten blir redusert, slik at det blir vanskeligere å trekke korrekte konklusjoner i forbedringsarbeidet. Det oppstår en risiko for at man rokker ved tolkningen av de viktige operative definisjonene. Å ha en form for «Pay-for-Performance» basert på resultater fra data som også skal anvendes i forbedringsarbeid og i forskningsarbeid, kan være svært ødeleggende for validiteten til forbedrings- og forskningsarbeidet. Lignende resonnementer føres og behandles blant annet i [22].

Til tross for alt dette er det mulig å forsvare bruken av registerdata i forbindelse med godtgjørelser og belønning – ikke sjelden hører man uttalelser som «først nå, når resultatene er koplet til økonomisk godtgjørelse, begynner man å interessere seg for disse resultatene. Det er kjempestimulerende». Og slik kan det være – at ledere på ulike nivåer faktisk spør om og er interessert i resultatene, er ekstremt viktig. Spørsmålet er imidlertid hva denne nyvakte in-



teressen for viktige problemstillinger forteller om de strukturene som har eksistert og som sikkert fortsatt finnes mange steder i helsesektoren i dag<sup>44</sup>.

### 8.3 Supplering av rapporteringen

I fortsettelsen presenteres her noen forslag til hvordan rapporteringen fra medisinske kvalitetsregistre kan suppleres i påvente av at konseptet blir videreutviklet.

#### 8.3.1 Forbedringstakt

Det handler generelt om å finne et mål for **forbedringstakt** og ikke bare betrakte det siste års nivåer og rangeringer. I det lange løp er det de som forbedrer seg raskest på grunnlag av egne og andres erfaringer som kommer til å nå det høyeste nivået. De som gjør forbedringer i det daglige arbeidet, har evne til å ta til seg nye erfaringer og gjennomføre kreative forandringer og innovasjoner, er de som kommer til å nå lengst.

#### 8.3.2 Spredningen har betydning

Som illustrert tidligere kan fokus på, og rapportering av, andeler virke mot sin hensikt. I forrige kapittel bruktes andelen pasienter som hadde nådd en viss HbA<sub>1c</sub>-verdi som eksempel. I mange andre situasjoner, ja kanskje i de aller fleste situasjoner, kan det være vel så viktig å se hele fordelingen. Det finnes som regel ulike delpopulasjoner av pasienter, kanskje til og med på grunn av ulik behandling. Dette blir først synlig når man ser hele fordelingen, da kan man gå videre og undersøke årsakene.

Å ta med spredning, det vil si et mål på variasjonen, er viktig med tanke på kapasitetsplanlegging. Dette er allerede diskutert, både i dette og forrige kapittel. I alle former

for flytsystemer bestemmes gjennomløpstiden av køsituasjoner, som i sin tur i stor grad er forårsaket av variasjon, både internt og utenfra<sup>45</sup>. Å basere beslutninger på middelværdier blir i slike tilfeller direkte uforsvarlig da det i siste instans kan gå utover enkeltpasienter. Problemstillingen er derfor ikke unik for kapasitetsplanlegging, en ikke-illustrert spredning er en generell problemstilling; hvordan er for eksempel variasjonen mellom pasienter som behandles ved samme enhet? Kanskje er en del av variasjonen helt naturlig med tanke på at hver pasient i en viss forstand er unik, men det kan også skyldes ulik behandling. Faktisk kan også variasjon for en og samme pasient i ulike sammenhenger være av betydning<sup>46</sup>.

#### 8.3.3 Konfidensgrader

I rapporteringen fra medisinske kvalitetsregistre er konfidensgraden i dag ofte satt til 95 prosent. Dette innebærer at dersom alt er tilfeldig variasjon, vil for eksempel likevel ett av tjue konfidensintervaller typisk ikke dekke den felles middelværdien eller den felles andelen. Det kan altså forventes at enkelte enheters konfidensintervaller ikke dekker den felles middelværdien selv om alle prosessene skulle være like. Å basere kostbare beslutninger på slikt kan være direkte uforsvarlig. Men om slike opplysninger tas som utgangspunkt for en dialog omkring de resultatene som er oppnådd, vil ikke konsekvensene være av samme art; kanskje vil det tvert imot være positivt – i det minste om dialogen skjer i den positive ånden som alt forbedringsarbeid er avhengig av.

#### 8.3.4 Variasjon og pasientsammensetning

Variasjon som forekommer innenfor en enhet, kommer sjeldent godt fram i rapporteringen fra medisinske kvalitetsregistre. Dette gjør det vanskeligere å få øye på case-

mix-problematikken som er høyst reell og **svært viktig** for sammenligning av enheter. En såkalt case-mix-problematikk har man altså når man i samme diagnosegruppe har pasienter med forskjellig sykdomsbilde og alvorlighetsgrad, og dette må det tas hensyn til. Det er ikke sikkert at samme behandling skal gis til pasientene i de ulike undergruppene, og derfor kan for eksempel «andelen som har fulgt standard behandlingsplan eller retningslinjer» bli misvisende. Et annet aspekt kan være at den beste legen får de dårligste resultatene etter som de vanskeligste tilfellene med dårlig prognose krever den dyktigste legen. En statistisk analyse med justering for **case-mix** vil aldri kunne bli komplett, man kan kun justere for de forskjeller som er registrert, så problemet med forskjeller i pasientsammensetning kommer til en viss grad alltid til å lure i bakgrunnen.

Som nevnt tidligere og i avsnittet over (men det er verdt en gjentagelse og poengtering), er case-mix-problematikken ikke bare relevant med tanke på variasjon innenfor enheter, den er kanskje vel så viktig for evaluering av variasjonen mellom enheter. Ulike pasientsammensetninger er en av hovedgrunnene til at fokuset bør rettes på forbedringstakt – ikke på absolutte størrelser!

## 8.4 Datakvalitet

Ved tolkning av kvalitetsindikatorer er det viktig å huske på at datakvaliteten ofte er varierende. Resultaters gyldighet er svært avhengig av hva som rapporteres inn fra enhetene. Mangel på gode rapporteringsrutiner er dessverre en aktuell problemstilling, men offentliggjøringen av resultater vil forhåpentligvis bøte noe på dette i form av økt oppmerksomhet på data generelt.

Det er viktig at den overordnede ledelsen skaper forutsetninger for god datafangst med IKT-systemer som støtter forbedringsarbeidet

i stedet for å bli tidstyver eller reelle hindre. For å sikre god datakvalitet er det nødvendig at de som registrerer data, kan se nytten av det forbedringsarbeidet de gjør, for eksempel ved direkte tilbakemeldinger og rapporteringssystemer som til en hver tid er tilgjengelig og oppdatert på nett.

## 8.5 Læring og forbedringsmuligheter

For organisasjonsledere i helsesektoren vil en effektiv metode for å stimulere til læring i organisasjonen være å stille opplyste, velfunderte spørsmål som for eksempel tar utgangspunkt i variasjonen som kommer frem i rapporteringen fra medisinske kvalitetsregistre. Man må være oppmerksom på positive trender og variasjon som indikerer gode resultater. Det handler om å søke forbedringsmuligheter. Det handler om å stimulere til framvekst av arenaer for forbedringsarbeid selv om de på kort sikt kan stjele tid fra arbeidet med pasientene. Prosessforbedringer kan og vil komme de framtidige pasienter til gode. Dette er selve meningen med eller kjernen i forbedringsarbeidet. Det er om å sikre alle pasienter – både de man har her og nå, og de som vil komme i fremtiden.

## 8.6 Spørsmål til refleksjon

Lederskap for forbedring og fornyelse krever at lederen gjennom sitt lederskap stimulerer til arbeidsmetoder der forbedring og fornyelse blir en naturlig del, og sammen med medarbeiderne skaper strukturer som støtter en slik arbeidsmetode. Noen viktige spørsmål på veien til å etablere slike arbeidsmetoder kan være til hjelp. Spørsmålene kan rettes både til lederne selv, i en selvrefleksjon, og til medarbeiderne:

- Har vi felles mål og et tydelig fokus på hvor vi ønsker oss i organisasjonen, og hvem vi faktisk er til for?
- Har vi fokus på bedre helsetjenester?
- Har vi en langsiktig strategi for å nå disse målene?
- Følger vi opp hvordan vi kommer videre ved både å anvende lokale mål og sammenligne oss med andre organisasjoner?
- Har vi arenaer i organisasjonen som gir oss mulighet til å ta til oss vellykkede eksempler fra andre organisasjoner?
- Kontinuerlig forbedring berører ikke kun de operative organisasjonsenhetene. Like viktig er det at forbedringer i organisasjonens struktur forenkler de operative enhetenes forbedringsarbeid. Har vi et tydelig forbedringsarbeid?
- Hvordan utnytter og støtter vi i organisasjonen de positive kreftene som ligger i at pasientene og deres pårørende deltar i forbedringsteam?
- Er vi tydelige på at vi kan få store kostnadsbesparelser ved å fjerne kvalitetsmangler?
- Forbedringsarbeidet handler mye om å ha evne til å arbeide ut over profesjons- og organisasjonsgrenser. Legger vi til rette for forbedringsteam som er tverrfaglige og som går ut over organisasjonsgrensene?
- Bygger vi tydelig inn forutsetninger for forbedringsarbeid i organisasjonen?
- Skaper vi evne til forbedringsarbeid?
- Skaper vi gode muligheter for å ta imot nye ideer og innovasjoner i organisasjonen?

Mer presise spørsmål som gjelder håndteringen av resultater fra medisinske kvalitetsregistre, handler om hvordan man bruker informasjonen som **signalsystem** og bruker rapportene for på en positiv måte å stimulere til **forbedring**. Signalsystemet skal brukes for å støtte utvikling – ikke for å legge skylden på noen! Variasjon kan oppstå av mange naturlige årsaker som man ikke

kan påvirke: tilfeldig variasjon, case-mix<sup>2</sup>, forutsetninger som hindrer utvikling, og så videre. Noen viktige spørsmål er:

- Finnes det andre organisasjoner som er bedre enn oss, som vi kan lære av?
- Har andre organisasjoner forbedret seg mer enn oss, og hva kan vi lære av dem?
- Utmerker våre resultater seg – når det gjelder absolutt nivå eller forbedringstakt?

Mye kan oppnås gjennom stimulerende spørsmål. Å vie resultatene oppmerksomhet – både gode resultater og resultater som tyder på at det kan finnes forbedringsmuligheter – er en av de viktigste funksjonene til ledelsen.

## 8.7 Avslutning

All ære til dem som fått i gang offentliggjøring og rapportering av resultater fra medisinske kvalitetsregistre. Selv om det åpenbart finnes mangler i dagens presentasjoner og systemer, har det blitt tatt et **svært viktig skritt** mot enda bedre helsetjenester i Norge. Rapporteringen fra kvalitetsregistrene kan og bør fungere som en bro mellom politiske og administrative ledere og de som arbeider i den kliniske hverdagen/virkeligheten. Rapportene skal betraktes som et signalsystem som kan utløse fruktbar dialog og være et felles utgangspunkt for det videre arbeidet for å skape **enda bedre helsetjenester!**



## A. Historisk tilbakeblikk

Allerede på begynnelsen av attenhundretallet begynte enkelte å argumentere for en klarere oppfølging i helsevesenet [15] – i Frankrike Pierre Charles Alexandre Louis, i England William Farr, John Snow og Florence Nightingale, og i Sverige Fredrik Theodor Berg, verdens første professor i pediatri og grunnlegger av Statistiska Centralbyrån. Allerede så tidlig som i 1856 ble verdens første landsdekkende sykdomsregister opprettet av Ove Guldberg Høegh (1814–1863) i Norge. Registeret gjaldt spedalske. På begynnelsen av nittenhundretallet tok amerikaneren Ernest Amory Codman med stor overbevisning til orde for åpen rapportering av resultatene i helsevesenet. Han grunnla et eget sykehus, kalt *The End Result hospital*, der han åpent rapporterte alle problemer som hadde oppstått på grunn av behandlingen. Hans teori var at om man ikke gjorde dette, ville det aldri bli mulig å oppnå viktige forbedringer. Han hadde tidligere argumentert så sterkt for åpen resultatrapportering at han ble avskjediget fra sykehuset han da arbeidet på, og ble tvunget til å gå av som leder for en forening han selv hadde vært med å stifte (en forgjenger til dagens Joint Commission). Codman opprettet også USAs første register over resultater i helsevesenet, et register over hans spesialfelt: skulderoperasjoner. Selv om det da var gått mer enn et halvt hundre år siden det norske lepraregisteret ble opprettet, var også dette et av verdens første kvalitetsregistre.

Nå, knapt hundre år etter Codmans pionerarbeid har man kommet et stykke på vei: Sveriges første kvalitetsregister ble opprettet i 1975,

og i 2006 startet deres årlige offentliggjøring *Öppna Jämförelser* som er en sammenstilling av allment interessante data fra mange ulike registre og andre datakilder.

Mye av tankegodset som den svenske modellen er basert på, altså selve konseptet om kontinuerlig læring fra kvalitetsregistre, er hentet fra et annet kunnskapsfelt enn helsevesenet, nemlig fra produksjonsindustrien. I svensk industri ble det tidlig opprettet databaser for tilbakeføring av erfaringer fra drift til produkt- og prosessutvikling, noe som i dag er svært vanlig for eksempel i bilindustrien for å forbedre produktene. Erfaringene fra industrien og dens kvalitetsfokus har bidratt til en kunnskapsutvikling som har kommet helsevesenet til gode. Industrien satte tidlig verdiskapingen i sentrum. Som en av bevegelsens pionerer, Walter A. Shewhart [23], uttrykte det, eksisterer organisasjoner for å tilfredsstillere behov som finnes i samfunnet, noe som også gjelder innenfor helsetjenestene. William Edwards Deming [5] videreførte og populariserte Shewharts tanker. Sammen med ideer fra Joseph Juran og fra «scientific management» fikk de stor betydning for det som kan kalles det «japanske industriunderet», som begynte på 1950-tallet og varte til ut på 1980-tallet. Tankene spredte seg tilbake til Vestens industri. På grunn av ulike tolkninger av de japanske forbedringsstrategiene ga dette opphav til et antall innbyrdes noe ulike bevegelser som gikk under betegnelser som for eksempel «Total Quality Management», «Six Sigma» og «Lean Production».

Deming sammenfattet sine tanker tidlig på

1990-tallet og reflekterte blant annet over hvilket kunnskapsgrunnlag en leder må ta til seg for å kunne stimulere sin organisasjon til kundefokusert forbedringsarbeid. De kunnskapsområdene Deming foreslo omtales i dag som «forbedringskunnskap» og har fått stort gjennomslag i helsesektoren gjennom arbeidet til blant annet Donald Berwick og Paul Batalden (se hjemmesiden til Institute for Healthcare Improvement). Områdene innenfor forbedringskunnskap er forståelse for variasjon, psykologi, kunnskapsteori og systemforståelse. I dette dokumentet har man først og fremst forsøkt å legge et grunnlag for forståelse av variasjon.

Forbedringskunnskap har i dag funnet veien inn i mange helserelaterte utdanninger, og det er utviklet en nasjonal plattform for forbedringskunnskap i samarbeid mellom mange av yrkesorganisasjonene i helsesektoren. I 2009 iverksatte Helse- og omsorgsdepartementet en nasjonal satsning på medisinske kvalitetsregistre med mål om å samordne og modernisere hele helseregisterfeltet i Norge. Rapporten *Gode helseregistre – bedre helse. Strategi for modernisering og samordning av sentrale helseregistre og medisinske kvalitetsregistre 2010-2020* ble utarbeidet i forprosjektet til satsingen og dannet grunnlag for etablering av Nasjonalt helseregisterprosjekt i 2011.

I Norge er det i dag (per 2013) 45 medisinske kvalitetsregistre med nasjonal status, men antallet er, i takt med satsingen på kvalitetsregistre innenfor de fleste medisinske fagfelt, stadig stigende. Med en generøs definisjon av et kvalitetsregister er det kanskje snakk om et par hundre medisinske kvalitetsregistre utover de nasjonale, men det finnes ingen eksakt oversikt. De eksisterende registre er ulikt utformet både når det gjelder omfang, innhold, datakvalitet, hjemmelsgrunnlag og tekniske løsninger.

I motsetning til Sverige og Danmark er Norge fortsatt i startfasen når det gjelder kom-

petanse, organisering og *systematisk* bruk av kvalitetsregistre. I Norge ser man nå spesielt til Sveriges *Öppna Jämförelser*, og vi er i ferd med å etablere og drifte en lignende systematisk rapportering fra kvalitetsregistrene. For mer om dette arbeidet og kvalitetsregisterfeltet generelt henvises det til nettsidene til Nasjonalt servicemiljø for medisinske kvalitetsregistre på:

[www.kvalitetsregistre.no](http://www.kvalitetsregistre.no)

## B. Sluttnoter

1. For en innføring i **kvalitetsutvikling** se [2], mens forbedringsarbeid beskrives i [9]. **Forbedringskunnskap** er også beskrevet i [6, 16]. Forskningsgrunnlaget for forbedringskunnskap gjennomgås i artikkelserien *Quality and Safety in Healthcare* i tidsskrivet BMJ fra mars 2011. En utvidet utgave av den svenske rapporten *Ännu bättre vård* planlegges før øvrig utgitt i 2014. SKL har på sin hjemmeside dokumenter som er relevante for emnet, for eksempel [13] og [8]. Informasjon om den svenske modellen finnes blant annet i [18] som førte til en omfattende satsing på kvalitetsregistre i Sverige. Høsten 2012 gjorde dessuten Myndigheten För Vårdanalys en vurdering av Öppna jämförelser.
2. En såkalt **case-mix**-problematikk er aktuell når man i samme diagnosegruppe har pasienter med forskjellige sykdomsbilder og alvorlighetsgrad (eller andre sentrale faktorer). Det er for eksempel ikke sikkert at samme behandling skal gis til pasientene i de ulike undergruppene av diagnosegruppen. Case-mix-problematikk gjør tolkning av kvalitetsindikatorer til en svært intrikat øvelse. Det er slettet ikke sikkert at en institusjon med et «bedre» resultat enn en annen institusjon virkelig *er* bedre, det kan faktisk være helt motsatt; kanskje blir de mest alvorlig syke pasientene (som påvirker en resultatindikator i negativ retning) sendt til behandlingsstedet som antas å være mest kompetent (ulike enheter har gjerne ulike spisskompetanser). Ofte kan også alvorlig syke pasienter overføres fra større institusjoner til mindre etter endt behandling, slik at små institusjoner kommer dårligere ut på en resultatoversikt. En statistisk analyse med justering for case-mix blir aldri komplett, man kan kun justere for de forskjeller i pasientene som er registrert, slik at problemet med (uobserverte) forskjeller i pasientsammensetninger alltid til en viss grad kommer til å lure i bakgrunnen.
3. Forbedringstakt vil drøftes nærmere senere. Man unngår mange problemer med ulik pasientsammensetning ved å anvende egne tidligere resultater som sammenligningsgrunnlag.
4. Både betydning av faktisk å måle og skape et velinformert beslutningsgrunnlag, og betydningen av å forstå at ikke alt lar seg måle, understrekes av Stanford-professorene Jeffrey Pfeffer og Robert Sutton i [21] hvor Einstein-sitatet er hentet fra.
5. Se for eksempel [www.helseregistre](http://www.helseregistre) for mer informasjon om de ulike typer helseregister.
6. Se for eksempel [www.helsenorge.no](http://www.helsenorge.no) (Helsetjenster  $\mapsto$  Kvalitet i helsetjenesten  $\mapsto$  Om kvalitet og kvalitetsindikatorer) for mer om kvalitet og kvalitetsindikatorer.
7. Kriterier for nasjonal status og evalueringsmodell for medisinske kvalitetsregistre kan finnes på: [www.kvalitetsregistre.no](http://www.kvalitetsregistre.no) (Kvalitetsregistre  $\mapsto$  Kriterier nasjonal status)
8. Dokumentet kan finnes (søk på «Håndbok» i søkefeltet) på: [www.kvalitetsregistre.no](http://www.kvalitetsregistre.no)
9. Modellen kapittelet tar utgangspunkt i, stammer fra Avedis Donabedian på 60-tallet, se [7]. Når det gjelder de individuelle pasientprosessene, har dette blitt behandlet systematisk innenfor rammen av mikrosystemteorien, der for eksempel pasienters individuelle målinger i ulike dimensjoner beskrives i form av et verdikompass, se [16].

10. Prosessbegrepet drøftes i [2, 10, 11, 3]. Flytterspektivet, som er nært forbundet med prosessbegrepet, drøftes av [14]. Det som ofte betegnes som nettverk og «solution shops» er i vårt perspektiv også prosesser. Det viktige er at man har en viss kontinuitet i virksomheten og kan lære av sine erfaringer.
11. Strukturbegrepet som anvendes her, har mye til felles med Antony Giddens begrep i hans struktureringsteori og Pjotr Sztompkas begrep i hans teori om «Social becoming», men har også visse likheter med Actor-Network-begrepet i Bruno Latours teori, der artefakter spiller en stor rolle i strukturens dynamikk.
12. Glykosert hemoglobin A<sub>1c</sub>, eller HbA<sub>1c</sub> er en blodprøve som viser gjennomsnittlig blodsukker for ca de siste åtte ukene før prøven er tatt. Når HbA<sub>1c</sub> måles finner man andel hemoglobin som har bundet sukker til seg. Svaret oppgis som regel i prosent (men her brukes benevnelsen mmol/l, eller mmol/mol, for en enkeltmåling av blodsukker ved et bestemt tidspunkt). Mennesker uten diabetes har HbA<sub>1c</sub>-verdier mellom fire og seks prosent, mens en diabetiker sies å ha et velregulert blodsukker hvis verdien er under syv prosent og et «dårligere» regulert blodsukker for verdier over. Helsedirektoratet anbefalte i september 2012 at HbA<sub>1c</sub>-verdier større enn 6,5% bør brukes som det primære diagnostiske kriterium for diabetes i Norge.
13. Det er vanskelig å si hva som fungerer og hva som ikke fungerer for en enkelt pasient, se [17].
14. Årsaks- og virkningssammenhenger er noe man håndterer daglig og skaper seg oppfatninger om, men det finnes ingen klar og entydig vitenskapelig definisjon av kausalitetsbegrepet. Filosofen David Hume drøftet de filosofiske aspektene ved dette. Noe av dette omtales også i kapittel 7.
15. Mikrosystemet er den minste funksjonelle enheten som til sammen står for et antall aktiviteter i en prosess (se for eksempel [16]) og dermed skaper verdi for pasienten. Selv om enheten har stor betydning innenfor helsetjenestene, følges Donabedians inndeling (se [7]) der resultatene gis av en prosess som bæres oppe av ett eller flere mikrosystemer.
16. Systemforståelse er en viktig del av forbedringskunnskap. I første vedlegg gis en kort historisk oversikt, der blant annet William Edwards Demings og Paul Bataldens innsats beskrives.
17. Dette diskuteres i [2].
18. Opprinnelsen til PDSA-sirkelen kan spores tilbake til Pragmatisk filosofi og Walter Shewharts tolkning av denne allerede i 1939, se [23].
19. Man innser i stadig større grad at dette kanskje heller ikke alltid er tilfellet i medisinske studier. Det understrekes stadig oftere at genetiske forskjeller gjør at behandlingen må tilpasses den enkelte. En mer omfattende forståelse for epigenetikk forsterker dette ytterligere. Og kanskje enda viktigere er det å gripe fatt i leveste og kontekstuelle faktorer som i høy grad er bestemmende for sykelighet og helse.
20. Kunnskapsbasert behandling framholdes i [21] som et viktig prinsipp i all behandling. Det unike med hver organisasjon gjør at man ved studier av intervensjoner er henvist til metoder som er anvendbare for n=1; det er ikke bare slik at organisasjonen er unik, ofte tolkes intervensjonene ulikt, og selve intervensjonsprosessen vil (som den må) bli håndtert forskjellig fra organisasjon til organisasjon. Men metoder for n=1 blir stadig vanligere også for medisinske forskningsstudier. Mennesker er unike og kan inndeles i stadig mindre grupper med like geno- og fenotyper, og derfor hører man stadig oftere at man søker etter individuelt tilpasset medisinsk behandling. Også innenfor medisinen nærmer man seg altså det spesielle problemet som n=1 innebærer. Slike metoder stiller krav til forståelsen av variasjon over tid, som drøftes i kapittel 5 og eksempelvis [17].
21. Se for eksempel [14].
22. Västra Skaraborg er beskrevet i [12].





35. Å tolke en konfidensgrad som en «virkelig» sannsynlighet er strengt tatt ikke helt korrekt. Et **konfidensintervall** er et intervall hvor man tror en ukjent verdi kan ligge, og det kalles et konfidensintervall fordi man har tillit eller konfidens (confidentia er latinsk for tillit) til at verdien ligger der. Konfidensintervallet angir en nedre og en øvre grense for en estimert størrelse som er beheftet med usikkerhet eller variasjon. Lengden av intervallet antyder hvor godt estimatet er; et langt intervall signaliserer større usikkerhet enn et kort og er derfor mindre informativt. Generelt kan dataobservasjoner betraktes som realiserte verdier av en eller flere stokastiske variabler/prosesser. Begrepet **stokastisk** betyr her **tilfeldig** men statistisk beregnbar, hvor det med tilfeldig menes det motsatte av deterministisk (A medfører alltid B). Fra dataobservasjoner av en stokastisk prosess kan det dannes et intervall som med en viss grad av sikkerhet eller konfidens inneholder den reelle men ukjente størrelsen av interesse. Intervallet er i utgangspunktet stokastisk og hvis man kunne observere verdiene til de stokastiske variablene mange ganger under identiske forhold, vil en konfidensgrad på 95% bety at man – i det lange løp – kan anta at 5 av 100 intervaller vil inneholde den ukjente størrelsen.
36. Skal man være veldig nøye, burde man også ta hensyn til den variasjon som finnes i de nasjonale verdiene. En mer presis beregning skulle da gi en anelse større konfidensintervaller, men her feies det under teppet av pedagogiske årsaker. Det samme er gjort med tanke på multippel testing.
37. **Korrelasjon** er et mål på styrken og retningen på den lineære avhengigheten mellom to variabler. En lineær sammenheng kan uttrykkes på formen  $Y = a \cdot X$ . En empirisk observert korrelasjon er *ikke* en tilstrekkelig forutsetning for å avdekke en kausal sammenheng mellom to variabler.
38. Det Store medisinske leksikon omtaler begrepet **bias** som: «brukes i statistisk og epidemiologisk forskning når resultater eller slutninger systematisk avviker fra det korrekte. Bias kan oppstå på grunn av feil eller unøyaktigheter ved utvalg av undersøkelsesobjekter, valg av undersøkelsesmetode eller vurdering av resultater».
39. Uttrykket **konfunderingsfaktor** brukes i epidemiologisk forskning om en faktor som samvarierer med de undersøkte årsaks- og virkningsvariablene uten selv å inngå i den undersøkte årsakskjeden. Man kan tenke på en konfunderingsfaktor mellom to variabler  $X$  og  $Y$  som en felles årsak  $Z$  til de to variablene.
40. Det ligger nok mye forskning bak for eksempel HbA<sub>1c</sub>-grensen, men en pasient med verdi 7.8% føler seg ganske sikkert ikke mye bedre enn en med 8% – i enkelte tilfeller kan det gjerne være omvendt.
41. Dersom det er mulig å rokke ved de operative definisjonene kan det oppstå ubehagelige effekter (se eksempelvis [22]).
42. Dette var en av Edward William Demings kjepphester – om man ikke har gode operative definisjoner når man måler, hvilke konklusjoner kan man da egentlig trekke? Ulike personer kan ha ulike definisjoner. Se appendikset kalt *Historisk tilbakeblikk* for mer om dette.
43. Se for eksempel [14].
44. Ledelsens engasjement for ett aspekt kan meget vel fortrenge andre aspekter som er langt viktigere for organisasjonens utvikling uten at ledelsen har ønsket en slik omprioritering.
45. Se for eksempel [14].
46. Så fort systemer framviser ikke-linearitet, og det gjør egentlig mer eller mindre de fleste systemer, blir et resonnement basert på middelveidier feil. Pasienters individuelle variasjon kan være det som påvirker sikkerheten og kanskje også behandlingsresultatet. Et eksempel på dette kan være blodsukkerverdien ved diabetes.

# Bibliografi

- [1] Bernt Alm, Per Möllborg, Laslo Erdes, Rolf Pettersson, Nils Åberg, Gunnar Norvenius, and Göran Wennergren. Sids risk factors and factors associated with prone sleeping in sweden. *Archives of disease in childhood*, 91(11):915–919, 2006.
- [2] B. Bergman and B. Klefsjö. *Kvalitet från behov till användning*. Studentlitteratur AB, 5. oplag edition, 2012.
- [3] B. Bergman, D. Neuhauser, and L. Provoost. Five main processes in healthcare: a citizen perspective. *BMJ quality & safety*, 20, 2011.
- [4] Science Business and Karolinska Institutet. *Health for All, Care for You, Unlocking the value of Personalised Healthcare in Europe*, 2010.
- [5] W. E. Deming. *Out of the crisis*. MIT Cambridge, 1982.
- [6] W. E. Deming. *The new economics: for industry, government, education*. MIT press, 2000.
- [7] Avedis Donabedian. The quality of medical care. *Science*, 200(4344), 1978.
- [8] Anders Edström, Carina Svensson, and Jesper Olsson. Att mäta för att veta. *Praktiska råd och tips*, 2005.
- [9] M. Elg, L. Witell, and V. Gautheureau. *Att lyckas med förbättringsarbete—förbättra, förändra, förnya*. Studentlitteratur, 2007.
- [10] A. Hellström. *On the Diffusion and Adoption of Management Ideas: Findings from six empirical studies in the quality field*. Chalmers University of Technology, 2007.
- [11] A. Hellström, S. Lifvergren, and J. Quist. Applying process management in healthcare—investigating implementation difficulties. In *International Annual EurOMA 2009 Conference Proceedings of the 16th Annual International EurOMA Conference, Göteborg*. Offox Press, 2009.
- [12] S. Lifvergren, U. Andin, A. Hellström, T. Huzzard, and B. Bergman. Toward a sustainable healthcare system: From learning to teaching clinical microsystems. *Emerald*, Vol. 2, 2012.
- [13] S. Lifvergren, O. Rentzhog, M. Edin, B. Klintberg, J. Olsson, and S. Tullberg. Six sigma i sjukvården. *Sveriges Kommuner och Landsting, Stockholm*, 2008.
- [14] N. Modig and P. Åhlström. *Vad är lean. En guide till kundfokus och flödeseffektivitet*, Stockholm, 2011.
- [15] A. Morabia. *A history of epidemiologic methods and concepts*. Birkhauser, 2004.
- [16] E. Nelson, P. Batalden, J. Mohr, M. Godfrey, L. Headrick, J. Wasson, et al. Microsystems in health care: Part 1. learning from highperforming front-line clinical units. *Joint Commission Journal on Quality and Patient Safety*, 28(9):472–493, 2002.

- [17] D. Neuhauser, L. Provost, and B. Bergman. The meaning of variation to health-care managers, clinical and health-services researchers, and individual patients. *BMJ quality & safety*, 20, 2011.
- [18] Sveriges Kommuner och Landsting. Översyn av de nationella kvalitetsregistren. *Guldgruvan i hälso-och sjukvården*, 2010.
- [19] N. Øyen. Spedbarn skal sove på rygg – en epidemiologisk suksesshistorie. *Norsk Epidemiologi*, 17(2):121–125, 2007.
- [20] N. Øyen, T. Markestad, L. Irgens, K. Helweg-Larsen, B. Alm, G. Norvenius, and G. Wennergren. Combined effects of sleeping position and prenatal risk factors in sudden infant death syndrome: the nordic epidemiological sids study. *Pediatrics*, 100(4):613–621, 1997.
- [21] J. Pfeffer and R. Sutton. *Hard facts, dangerous half-truths, and total nonsense: Profiting from evidence-based management*. Harvard Business Press, 2006.
- [22] D. Pink. *Drivkrafter, den förvånande sanningen om vad som motiverar oss*. Book-House edition, 2010.
- [23] Wa. Shewhart and W. Deming. *Statistical method from the viewpoint of quality control*. Dover Publications, 1939.
- [24] N.N. Taleb. *The Black Swan: Second Edition - The Impact of the Highly Improbable Fragility*. Random House Publishing Group, 2010.

# Register

- 3 ·  $\sigma$ -grenser, 25, 26  
Årsakssammenheng, *Se* Kausalitet
- Actor-Network-teori, 50  
Anscombe, Francis, 33  
Anscombes kvartett, 33  
Assosiasjon, 36
- Batalden, Paul, 11  
Behandlingskjede, 15  
Beskrivende studier, 21  
Bias, *Se* Skjevhet  
Black swan, 51
- Case-mix, 1, 17, 44, **49**  
Critical thinking, 51
- Datafangst, 11  
Datakvalitet, 11, 38  
Deming, William Edwards, 11, 50, 52  
Deterministisk, **10**, 51, 52  
Deweys, John, 51  
Diabetes, 10  
Diakron variasjon, 19, 21, 30, 51  
Donabedian, Avedis, 7, 8, 49, 50
- Epigenetikk, 50  
Evidens, **10**, 50
- Flytperspektiv, 15, 43, 50  
Forbedring, 14  
Forbedringsarbeid, 1, 13, 15, 16  
Forbedrings sirkel, **13**, 14  
Forbedringstakt, 2, 14, 18, **43**, 49  
Fordeling, *Se* Sannsynlighetsfordeling  
Forebyggende tiltak, 8, 15  
Forskyvningseffekt, 42
- Forventningsskjevhet, *Se* Skjevhet
- Genotype, 26, 29  
Gjennomsnittsverdi, **19**, 30, 33, 51  
Glemt nevner, 35
- Håndbok, 6, 49  
HbA<sub>1c</sub>, 10, 16, 17, 19, 20, 27, 37, 39, 43, **50**  
Heterogenitet, 51  
Holdninger, 8  
Hume, David, 50
- Indikator, 10  
Individuell pasientprosess, 9, 49, 50  
Informasjonshåndtering, 2, 7  
Infrastruktur, 2, 13
- Juridisk hjemmel, 5
- Kausalitet, 10, 35, **36**, 50, 52  
Kilder til variasjon, 11  
Konfidensgrad, **30**, 31, 52  
Konfidensintervall, 16, 30, 31, 43, 51, **52**  
Konfunderingsfaktor, 36  
Kontrolldiagram, 25, 27  
Korrelasjon, **33**, **52**  
Krybbedød, 20  
Kunnskapsbasert, *Se* Evidens  
Kvalitetsforbedring, 5  
kvalitetsindikator, 5  
Kvalitetsregister, **5**, 7, 13, 15–20, 41, 43
- Læringsprosess, 16  
Ledelse, 1, 2, 13
- Måleproblematikk, 3, **11**  
Måling over tid, **11**, 14

- Måltall, 9  
 Median, 30  
 Medisinsk kval.register, *Se* Kvalitetsregister  
 Middelværdi, *Se* Gjennomsnittsverdi  
 Mikrosystemteori, 49, 50
- Nasjonal status, 6  
 Nasjonalt servicemiljø, 6, 48  
 Naturlig variasjon, *Se* Tilfeldig variasjon  
 Nettbasert, 2
- Offentliggjøring, 15, 18, 32, 45  
 One-stop-shop, 15  
 Operativ definisjon, 17, 37, 52  
 Organisasjonsteori, 15  
 Organisering, 8  
 Overlevelsesskjevhet, 35
- Pasientforløp, 8  
 Pay for performance, 22, 51  
 PDSA-sirkel, *Se* Forbedringssirkel  
 Pirater, 36  
 Placeboeffekt, 35  
 Plan-Do-Study-Act, *Se* Forbedringssirkel  
 Probabilistisk, 10, 51  
 Prosess, 7  
 Prosessforbedring, 15  
 Prosessmåling, 9  
 Publiseringsskjevhet, 35
- Randomiserte kontrollerte studier, 15  
 Representativt utvalg, 35  
 Ressursbruk, 1, 15  
 Resultat, 7  
 Resultatmåling, 9
- Samtykke, 5  
 Sannsynlighetsfordeling, 10, 23, 30, 51, 51  
 Shewhart, Walter, 25, 50, 51  
 Signalsystem, 18, 25, 41, 45  
 Signifikans, 16  
 Signifikansnivå, *Se* Konfidensgrad  
 Skjevhet, 34  
 Spesiell variasjon, 23, 23, 29  
 Standardavvik, 30, 33, 51
- Statistisk skjevhet, *Se* Skjevhet  
 Statistisk stabilitet/likevekt, 25, 29, 30, 51  
 Stokastisk, 51, 52  
 Stokastisk prosess, 51, 52  
 Struktur, 7, 8, 50  
 Struktureringsteori, 50  
 Suboptimalisering, 15  
 Synkron variasjon, 19, 30, 51  
 System, 7, 8  
 Systemforståelse, 50  
 Sztompkas, Pjotr, 50
- Taleb, Nassim, 51  
 Tidsdimensjon, 11, 19, 20  
 Tidsserie, 27, 51  
 Tilfeldig, 23, 30, 52  
 Tilfeldig variasjon, 17, 23, 29, 31
- Uønsket variasjon, 29  
 Usikkerhet, 30  
 Utvalgsskjevhet, *Se* Skjevhet
- Validitet, 19  
 Variasjon, 19  
 Visjon, 13



## Forskjeller i helsetjenesten

Dokumentet omhandler hvordan man ut fra ulike perspektiver skal forstå, håndtere og trekke konklusjoner fra variasjonen som forekommer i de medisinske kvalitetsregistrene. Primært legges et overordnet perspektiv til grunn og teksten tar for seg hvordan man med et slikt (ledelses-)perspektiv kan stimulere til enda bedre helsetjenester. Det gjennomgående mantra er: «Den som ikke forbedrer seg, slutter snart å være god.»

[www.kvalitetsregistre.no](http://www.kvalitetsregistre.no)

ISBN: 978-82-93141-12-9

### Postadresse

Senter for Klinisk Dokumentasjon og Evaluering (SKDE),  
Postboks 6,  
9038 Tromsø

### Besøksadresse

Plan 5, Fløy C0,  
Univ.sykehuset Nord-Norge,  
Sykehusveien 38, Breikvia  
9038 Tromsø

### Kontaktinformasjon

Telefon: 777 55 800  
E-post: [post@skde.no](mailto:post@skde.no)  
Internett: [www.skde.no](http://www.skde.no),  
[www.kvalitetsregistre.no](http://www.kvalitetsregistre.no)